Improved Methods for Combining Point Forecasts for an Asymmetrically Distributed Variable*

Özer Karagedikli
SEACEN and CAMA

Shaun P. Vahey
University of Warwick and CAMA

Elizabeth C. Wakerly
Economic Prediction Systems

June 22, 2019

Abstract

Many studies have found that combining forecasts improves predictive accuracy. An often-used approach developed by Granger and Ramanathan (GR, 1984) utilises a linear-Gaussian regression model to combine point forecasts. This paper generalises their approach for an asymmetrically distributed target variable. Our copula point forecast combination methodology involves fitting non-parametric marginal distributions for the target variable and the individual forecasts being combined; and then estimating the correlation parameters capturing dependence between the target and the experts' predictions. If the target variable and experts' predictions are individually Gaussian distributed, our copula point combination reproduces the GR combination. We illustrate our methodology with two applications examining quarterly forecasts for the Federal Funds rate and for US output growth, respectively. The copula point combinations outperform the forecasts from the individual experts in both applications, with gains in root mean squared forecast error in the region of 40% for the Federal Funds rate and 4% for output growth relative to the GR combination. The fitted marginal distribution for the interest rate exhibits strong asymmetry.

Keywords: Forecast combination; Copula modelling; Interest rates; Vulnerable economic growth.

*We thank Anthony Garratt, Ivan Petrella, Christie Smith, Leif Anders Thorsrud, Knut Are Aastveit, Hilde Bjørnland and seminar participants at the Reserve Bank of New Zealand, Bank of Canada, Norges Bank, HEC Montreal, Warwick University, CIRANO Montreal, Monash University, ANU, and the December 2017 SCE Meetings, the March 2019 SNDE Meetings and the May 2019 EABCN conference on "Advances in Business Cycle Analysis" (held at the Bank of Spain) for comments. The not for publication appendix can be downloaded from shaunvahey.com. The views in this paper do not necessarily represent the views of SEACEN.

1 Introduction

Policymakers and (some) survey respondents often communicate the conditional asymmetric risks for macroeconomic variables. Prominent examples include the Federal Open Market Committee discussions of risks, and the forecast densities provided by both the Blue Chip Economic Survey of economists and the Federal Reserve Bank of Philadelphia's Survey of Professional Forecasters. There has also been renewed interest recently in quantifying the asymmetries in conditional risk assessments for US macroeconomic variables. Among others, Salgado, Guvenen and Bloom (2016), Giglio, Kelly and Pruitt (2016), Smith and Vahey (2016) and Adrian, Boyarchenko and Giannone (2018) provide recent examples.

A distinct literature has considered the scope for predictive combinations to improve the accuracy of point forecasts. Granger and Ramanathan (GR, 1984) proposed using a linear-Gaussian regression framework to combine point forecasts from individual experts. Their approach built on earlier work by (among others) Bates and Granger (1969) which proposed the use of weighted averages for forecast combination. In the GR framework, the decision maker performing the combination only has point forecasts from the experts to combine.

In this paper, we modify the GR point combination framework to a setting with an asymmetrically distributed target variable. Our decision maker fits well-calibrated non-parametric marginal distributions to the target variable and the individual forecast from each expert to construct the Probability Integral Transforms, PITS. (Among others) Rosenblatt (1952), Diebold, Gunther and Tay (1998), Ranjan and Gneiting (2010), Galbraith and van Norden (2012) and Rossi and Sekhposyan (2018) discuss properties of the PITS and probabilistic calibration. Our decision maker transforms the (univariate) PITS to be normally distributed—following the intuition of Berkowitz (2001)—and uses a pseudo regression (either with or without restrictions) to estimate the dependence parameters. This copula point combination approach nests the conventional GR regression in the sense that restricting the fitted marginal distributions delivers exactly the GR combination.

We analyse two applications in depth to illustrate our copula point combination methodology, with individual experts predicting macroeconomic variables through the Global Financial Crisis and its aftermath. In the first application, we predict the Federal Funds rate, where the experts produce point forecasts based on real-time measurements for the output gap and inflation. An advantage of our copula modelling strategy is that no auxiliary assumptions are required about the existence of a fixed (or time-varying) lower bound for interest rates. Considering a quarterly evaluation sample from 2000:2 to 2015:1, we find strong evidence of interest rate predictability, with our copula point combination outperforming the conventional GR combination in terms of root mean squared forecast error. The copula combination provides accurate predictions when interest rates are unusually low but positive after the crisis, benefitting from the bimodal marginal fitted distribution for the target variable.

In the second application, we consider a target macroeconomic variable with a fitted marginal distribution that is closer to symmetric. We analyse predictions for US real output growth, where the experts' predictions are based on copula time series models. Motivated by earlier work documenting the skewness in the conditional predictive distributions for quarterly US economic growth by Adrian, Boyarchenko and Giannone (2018), our two experts exploit the dependence between real activity and financial conditions. One expert utilises an autoregressive copula model with linear dependence and the other uses credit spreads with stronger left-tail dependence than at the mean—asymmetric dependence. Considering a quarterly evaluation sample from 1992:4 to 2017:1, we again find that our methodology produces lower root mean squared forecast errors than the conventional GR combination and the individual experts through the Global Financial Crisis and the recovery years.

Copulas are an appealing route to extend the GR regression methodology for point forecast combinations. They provide a flexible econometric modelling strategy, separating marginal distributions from dependence. Earlier econometrics papers in macroeconomics and finance using copula modelling methods include Patton (2006, 2012), Scotti (2011) and Smith and Vahey (2016). To estimate the parameters of the combination dependence structure, we follow Smith and Maneesoonthorn (2018) by deploying an "inversion" approach. Loaiza-Maya and Smith (2018) provide a recent multivariate application based on real-time US macroeconomic variables.

A useful feature shared by both GR and our framework is that the combinations only require point forecasts from the experts. A related but distinct literature focuses on opinion (prediction) pooling via density combination, where the experts report forecast densities. Jore, Mitchell and Vahey (2010), Geweke and Amisano (2011), Waggoner and Zha (2012), Billio, Casarin, Ravazzolo, and van Dijk (2013), Del Negro, Hasegawa and Schorfheide (2016) and Bassetti, Casarin and Ravazzolo (2018) provide recent macroeconomic and finance applications. In contrast, Jouini and Clemen (1996) and Meucci (2005) explore copula opinion pools, which combine forecast densities from experts, modelling dependence and the marginal distributions separately. In this paper, we combine point forecasts from experts in the Granger-Ramanathan tradition.

Unlike many applied studies considering forecast combination, recursively updating the point combination weights aids forecast performance in both applications. In applied work, simple equally weighted combinations often perform better in practice. For discussions of this "forecast combination puzzle" or "equal weights puzzle" see (among others)

Clemen and Winkler (1986), Timmermann (2006), Capistrian and Timmermann (2009), Bjørnland, Gerdrup, Jore, Smith and Thorsrud (2011), Conflitti, De Mol and Giannone (2015) and Diebold and Shin (2018). Existing explanations for the puzzle emphasise the estimation error in the weights, the potential gain from forecast combination (related to expert selection), truncation (discarding experts) and the potential for shrinkage. The scope for asymmetrically distributed macroeconomic variables constitutes a neglected issue in the forecast combination puzzle.

The remainder of this paper is structured as follows. In Section 2, we set out our copula point combination framework. In Sections 3 and 4, we consider the two applications, one based on forecasting the Federal Funds rate and the other on predicting US economic growth. We draw some conclusions and ideas for extending the analysis in the final section.

2 A Framework for Copula Point Combinations

In this section, we generalise the GR approach to deal with an asymmetrically distributed target variable utilising a copula modelling strategy.

2.1 The GR Combination Regression

The underlying idea is that the decision maker combines the forecasts produced by individual experts; see Barnard (1963) and Bates and Granger (1969). The regression approach, developed by GR, involves estimation of a linear-Gaussian model with the predictions of experts as regressors:

$$Z_t \;=\; a \;+\; b_1 Z^p_{1,t} \;+\; b_2 Z^p_{2,t} + \cdots + \; b_R Z^p_{R,t} \;+\; e_t, \qquad t = 1, \ldots, T, \tag{1}$$

where there are $r = 1, \ldots, R$ experts, Z_t is the target variable of interest (eg the Federal Funds rate), $Z^p_{r,t}$ is expert r's point prediction, and the combination disturbance term, e_t, is usually taken to be Gaussian (or a generalisation of that distribution). Although GR explored a frequentist route to inference, Bayesian methods are also feasible. Degrees of freedom issues arise when the number of experts is large relative to the number of time series observations. Nevertheless, if the focus of the decision maker lies in forecasting—rather than estimating the "true" weights—collinearity amongst the experts may not be a concern.

Common modifications explored in practice include: restricting the b coefficients to sum to one and dropping the intercept; inequality constraints on the parameters, $0 \leq b_r \leq 1$; various selection criteria based weights; model averaging and shrinkage; and restricting the weights to be equal for all experts.

A related combination strategy to improve probabilistic calibration in weather forecasting applications (with probabilistic forecasts) is known as Ensemble Model Output Statistics; see Gneiting, Raftery, Westveld and Goldman (2005). Schefzik, Thorarinsdottir and Gneiting (2013) develop Ensemble Copula Coupling to extend consideration to multivariate dependence in meteorology.

This paper is concerned with cases in which the target macroeconomic variable, Z_t, is asymmetrically distributed and the decision maker wishes to combine point forecasts. This includes applications where the experts do not report forecast densities.

2.2 The Copula Point Combination

We extend the GR approach by exploiting Sklar's Theorem to seperate the marginal distributions from the underlying dependence structure.

For a multivariate time series, $\boldsymbol{Z}$, comprising the target variable and the R forecasts from experts (where $S = R + 1$), applying Sklar's Theorem (Nelsen, 2006) suggests there exists a copula function C such that the joint distribution function can be written:

$$F(\boldsymbol{z}) = C(\boldsymbol{u}) \tag{2}$$

where $\boldsymbol{z} = (\boldsymbol{z}'_1, \ldots, \boldsymbol{z}'_T)'$, $\boldsymbol{u} = (\boldsymbol{u}'_1, \ldots, \boldsymbol{u}'_T)'$, $\boldsymbol{z}_t = (z'_{1,t}, \ldots, z'_{S,t})'$, $\boldsymbol{u}_t = (u'_{1,t}, \ldots, u'_{S,t})'$, $u_{s,t} = F_{z_s}(z_{s,t})$ with $s = r + 1$.

The target variable and the predictions from the experts are assumed to be stationary and in our applications the margins are taken to be time invariant. The copula function C itself is a distribution function on the unit cube $[0, 1]^S$, and all margins are strictly uniform. Dependence between elements of $\boldsymbol{Z}$ is captured by the copula function.

By differentiating the distribution function, Equation (2), we obtain the density of $\boldsymbol{Z}$ as:

$$f(\boldsymbol{z}) = c(\boldsymbol{u}) \prod_{t=1}^{T} \prod_{s=1}^{S} f_{z_s}(z_{s,t}) \qquad s = r + 1 \tag{3}$$

where $f_{z_s}(z_{s,t})$ is the marginal density of $z_{s,t}$, $F_{z_s}(z_{s,t})$ is the corresponding distribution function, and $c(\boldsymbol{u}) = \frac{\partial}{\partial \boldsymbol{u}} C(\mathbf{u})$ is usually referred to as a copula density, where $\mathbf{u} = (u_1, \ldots, u_S)'$.

From Sklar's Theorem, a copula density always exists but it is unknown and has to be selected in practice. Assuming a Gaussian copula facilitates likelihood-based estimation and corresponds to the linear dependence case considered in the GR regression.

Denoting the normal distribution function as $\Phi(\cdot)$, the Gaussian copula density is:

$$c(\mathbf{u}) = c_{\text{Ga}}(\mathbf{u}; \Omega) = |\Omega|^{-1/2} \exp\left\{ -\frac{1}{2} \boldsymbol{\zeta}'(\Omega^{-1} - I_S)\boldsymbol{\zeta} \right\}, \tag{4}$$

with $\mathbf{u} = (u_1, \ldots, u_S)$, $\boldsymbol{\zeta} = (\Phi^{-1}(u_1), \ldots, \Phi^{-1}(u_S))'$, and a correlation matrix Ω as dependence parameters.

2.3 Estimation and Prediction with Copula Point Combinations

In our applied work, the process of fitting a copula model involves four steps taken by the decision maker. First, fit a cumulative distribution function (CDF), $F_{Z_s}(Z_{s,t})$, to each of the S margins—for the experts' forecasts and the target variable. Second, compute the Probability Integral Transforms, PITS, for each time series from the (time-invariant) marginal distribution. That is, compute the time series of PITS, $F_{Z_s}(z_{s,t}) = Pr(Z_s \leq z_{s,t})$, for each of the $S = R + 1$ margins. Third, following Smith and Maneesoonthorn (2018), transform each PITS series (on the unit interval) using an inverse normal CDF to give a multivariate pseudo time series, with each variable individually normally distributed. The fourth step involves modelling the dependence structure. Given that the dependence in the GR regression is linear, the decision maker carries out inference via a linear-Gaussian regression on the pseudo time series representation:

$$\zeta_t \; = \; \alpha \; + \; \beta_1 \zeta^p_{1,t} \; + \; \beta_2 \zeta^p_{2,t} + \cdots + \; \beta_R \zeta^p_{R,t} \; + \; \epsilon_t, \qquad t = 1, \ldots, T \tag{5}$$

where ζ_t is the pseudo time series counterpart to Z_t, and the pseudo predictions of the $r = 1, \ldots, R$ experts to be combined are denoted $\zeta^p_{r,t}$. The pseudo combination error term, ϵ_t, is normally distributed.

With inference conditional on the fitted marginal, this limited information estimation strategy, sometimes known as "inference for margins", is slightly less efficient than full information; see Joe (2006). Inference can be carried out by Bayesian or frequentist methods as for the conventional GR regression.

Prediction by the decision maker with the copula combination involves reversing the process for the predictand. First, construct the predicted values for the pseudo target variable, $\hat{\zeta}$, based on Equation (5). Second, compute the PITS from the Gaussian marginal distribution, $F_\zeta(\zeta_t) = \Phi_\zeta(\zeta_t) = Pr(\zeta \leq \zeta_t)$. Third, use the inverse marginal distribution, $F_{Z_1}^{-1}(Z_{1,t}) = F_Z^{-1}(Z_t)$, to generate predictions from the copula combination with the same marginal distribution as the target variable.

In our applied work, all estimation and prediction steps are carried out recursively to mimic real-time forecasting by the decision maker.

2.4 Fitting Marginal Distributions

The decision maker must fit (time-invariant) margins to the sample data. A special case occurs where Z_t follows a Gaussian distribution, as do the forecasts from the experts, the Z_t^p's. Then, the operations through the CDFs and inverse CDFs cause the pseudo regression to coincide with the original GR regression. In this case the copula point combination forecast is the GR regression point forecast.

More generally, where the target variable follows an unknown asymmetric distribution, non-parametric fitting of the Empirical CDF (ECDF) to Z_t via a kernel smoother provides a pragmatic route to selecting marginals. (Parametric marginal fitting is also feasible.) In our applied work, we follow the non-parametric approach, fitting an ECDF to each $Z_{s,t}^p$ in turn. In effect, we treat the forecasts as explanatory variables in the pseudo regression. This requires S marginals to be fitted—one for each of the R experts, plus one for the target variable.

Alternatively, since the experts are predicting the target variable, the decision maker could use the same marginal distribution fitted to Z_t for the forecasts from the individual experts, the Z_t^p's. In our applications, this second homogeneous marginals approach produces slightly weaker forecast performance in terms of root mean squared forecast error. We report below the heterogeneous marginals case and report the homogeneous marginals case in the not for publication appendix.

3 Application: Federal Funds Rate Predictions

In the first application, we consider point forecast combinations by a decision maker for the quarterly Federal Funds rate for the target dates from 2000:2 to 2015:1.

3.1 Forecasts from the Experts

Our experts predict nominal US interest rates using real-time measurements for the output gap and inflation.[1] Both experts forecast with specifications discussed by Bernanke (2015). The experts are differentiated by the importance of the output gap, with the first expert favouring a greater weight on this consideration than the second. Both experts predict the end-quarter Federal Funds rate based on real-time measurements available within the current quarter.

The experts utilise a linear dependence approach to capture the contemporaneous relationship between the Federal Funds rate and output gap and inflation. The first expert uses the following specification:

$$FFR_t \;=\; INF_t \;+\; 1.0\,GAP_t \;+\; 0.5\,(INF_t - 2) \;+\; 2 \tag{6}$$

where FFR_t, INF_t and GAP_t denote the Federal Funds rate, inflation and the output gap in time t, respectively. The second expert has a smaller weight on the output gap than the first—a coefficient of 0.5 rather than 1.0 for the variable GAP_t.

For convenience in presenting the copula forecast combination results, we will refer to

[1] Taylor (1993) and Henderson and McKibbin (1993) provide analysis of monetary feedback rules. Taylor (1999) notes the effectiveness of this single equation approach as a predictive tool. Carlstrom and Jacobson (2015) note that many close variants have been used for forecasting in practice.

the first expert as Bernanke, and the second as Taylor. We emphasise that we do not imply that either Bernanke or Taylor would endorse forecasting interest rates in this automated fashion in practice. We note that Yellen (2012) argues for a specification similar to the Bernanke expert and that Taylor (1999) proposes a variant in which the coefficient on the output gap is greater than zero. Among others, Johanssen and Mertens (2016) model a lower bound using a shadow interest rate. That approach treats the price of money as an asymmetrically distributed variable but without separating the margins from the dependence.

As described in the data appendix, we assume that both experts use the "real-time" macroeconomic data provided by Bernanke (2015) to forecast. There are minor differences in the measures for inflation and the output gap used by each expert.

Our two experts make quarterly predictions for the target variable for 2000:2 to 2015:1. Figure 1 plots the real-time predictions from the experts, along with the target variable, the Federal Funds rate (black line). The Taylor expert (red line) performs poorly during the Greenspan boom, 2004 through 2006, when it predicts higher interest rates than seen in the data. In contrast, the Bernanke expert (blue line) tracks the Federal Funds rate fairly closely through the same observations. In the aftermath of the Global Financial Crisis, both experts predict spells of negative nominal interest rates, and as a result there are large forecast errors from both. Of the two, the Taylor expert predicts higher rates throughout the post-2008 sample, with interest rates in excess of zero forecast from early 2011. The predictions from the Bernanke expert do not imply until the end of the evaluation sample.

Given the slow recovery experienced by the US economy in the aftermath of the financial crisis, it is unsurprising that systematic forecast errors are made by both experts. There is scope for recalibration to improve the performance of the experts. We shall return to this issue subsequently.

3.2 Forecast Combination by the Decision Maker

The decision maker recursively combines the forecasts from the experts using an expanding window for parameter (weight) estimation from 1996:2 to 2014:4, with weight training from 1996:2 to 2000:1. Prior to combination, the decision maker must fit marginal distributions. In the heterogenous marginals case, the decision maker fits a distinct marginal distribution to the forecasts produced by each expert, and to the target variable. We assume that in practice the decision maker uses a kernel-smoothed non-parametric ECDF approach for all margins, recursively fitting each marginal distribution through the expanding window used for estimation. Non-parametric kernel fitting uses the MATLAB function ksdensity.[2]

The (smoothed) Probability Density Function (PDF) for the Federal Funds rate plotted in Figure 2 (black solid line) displays an asymmetric distribution, and is based on the sample of data considered for the last forecast target, 2015:1. Our forecast evaluations use recursively fitted marginal distributions based on expanding windows of data. Hence the fitted marginals for each variable vary by forecast origin. For the last origin, the PDF for the Federal Funds rate is bimodal with the peak around zero reflecting the lower interest rates from 2009:1.

The PDFs corresponding to the marginal distributions for the Taylor expert (red line) and the Bernanke expert (blue line) are also shown in Figure 2. Both display high-order moments that depart from a symmetric Gaussian distribution. Furthermore, the marginals for the experts display considerable differences from each other, and from the fitted marginal distributions for the target variable. In particular, the Bernanke expert has more mass in the lower tail than the Taylor expert, a feature consistent with the two variants of the Taylor rule in the aftermath of the financial crisis, as displayed in Figure 1. The three marginal distributions are relatively close in the upper tail. For the

[2]The not for publication appendix discusses robustness to the choice of bandwidth.

target variable, and the two experts, the null hypothesis of normality is rejected at the 1% significance level for the Shapiro-Wilk test.

3.2.1 Recalibrated Experts

Given the forecast combination setup, where the predictions from the experts are subsequently combined by the decision maker, the experts might prefer to recalibrate prior to combination. Madansky (1964) and Genest and Zidek (1986) discuss related concepts about group rationality and the externally Bayesian property of pooled opinions. Here, we consider recalibrated predictions based on the marginal distribution fitted to the target variable by the decision maker. Whether the decision maker performs the recalibration or the expert is not germane, given the setup of the combination step in the GR framework. In our application, to recalibrate the forecasts, the predictions from each expert were transformed using the (real-time) fitted marginal distribution for the target variable. The predictions from the recalibrated experts are plotted in Figure A1 in the not for publication appendices.

3.2.2 Estimation and Prediction by the Decision Maker

We now turn to the estimation of the dependence parameters, shown in Equation (5). For each forecast origin, the decision maker computes the PITS conditional on the respective fitted marginals and then constructs the pseudo time series, as discussed in the previous section, via the inverse CDF. The pseudo regression is then estimated by maximum likelihood. In this manner, the decision maker estimates the linear dependence parameters (of a Gaussian copula). The predicted value for each period is converted from the unit interval using the inverse (kernel-smoothed) ECDF to give the mean forecasts for the observable target variable, FFR_t.

As for the conventional GR regression, a number of restrictions on the parameters are feasible, including equal weights, which are explored in the next section.

3.3 Results

The forecast evaluations use data for 2000:2 to 2015:1, with all forecast combinations and predictions from the experts being real-time quarterly predictions. The decision maker estimates point combination weights using an expanding window from 1996:2 to $t-1$ to produce the forecast for observation t.

Figure 3 displays the predictions from the copula point combination (red solid line), together with the unrestricted GR regression benchmark (blue solid line) with recursive weights and its equal weighted counterpart (blue dashed line). In this application, the restricted GR outperforms the unrestricted GR by around 2%, in terms of root mean squared forecast error (RMSFE) over the whole evaluation sample. The copula point combination has a zero intercept, bounded weights (estimated recursively) on the unit interval and weights that sum to one across experts. (The unrestricted copula point combination performs around 20% worse than the restricted equivalent for the full sample in terms of RMSFE.) The copula methodology provides accurate point predictions of the Federal Funds rate (black line), especially in the aftermath of the Global Financial Crisis. As one might expect with "real-time" interest rate projections, the target variable falls faster than the combination prediction during the onset of the crisis, leading to persistent forecast errors. The Bernanke and Taylor expert predictions both share this characteristic; see Figure 1. Elsewhere, the copula point combination tracks the actual interest rate relatively well. Given the controversy around the relative strengths of equal and recursively estimated weights, it is interesting to note that the performance of the GR regression combination varies little with weight type.

Figure 4 displays the recursively estimated weights from the copula point combination (solid lines) and the benchmark GR regression (dotted lines). In each case, the weights on the Taylor expert (red lines) and the Bernanke expert (blue lines) exhibit considerable time variation, with the Bernanke expert receiving the greater weight for much of the evaluation sample, for both the benchmark and the copula point combinations. For the copula combination, however, the weight on the Bernanke expert exhibits relative stability prior to the financial crisis. For the GR regression, the weight on the Bernanke expert is low at the start of the evaluation sample, but reaches over 50% by the onset of the crisis. After 2009, the weights shift towards the Bernanke expert for the GR regression, but in contrast the copula combination suggests that the Bernanke weights decline. The relatively modest variation in the target variable after the Financial Crisis contributes to the weight imprecision in this era.

Table 1 displays end-evaluation relative RMSFE results for a variety of combinations and experts, with the conventional GR regression as a benchmark (with unrestricted recursively estimated parameters and naive experts). The equally-weighted conventional GR regression, row 3, has a 2% higher RMSFE than the benchmark, row 2. The copula point combination, row 4, provides the best RMSFE, with a performance gain of nearly 45% over the benchmark. Recall that this version of the copula point combination has restricted coefficients (zero intercept, bounded weights, weights sum to one).[3] The copula point combination with equal weights, row 5, performs less well, but still improves on the benchmark by around 34%. Hence, equal weights are not as effective as recursive weights with the copula point combination methodology.

The Taylor and Bernanke experts, rows 6 and 7, fail to outperform the GR regression. Individually, the experts are around 6% and 62% worse than the benchmark combination,

[3]The unrestricted variant (not reported in the table) lifts the relative RMSFE slightly closer to the benchmark at 0.72.

Table 1: Relative RMSFE, Federal Funds Rate

Expert/Combination		
Granger-Ramanathan (GR)	1.0000	
GR Equal Weight	1.0320	
Copula Combination (CC)	0.5569	∗∗
CC Equal Weight (EW)	0.6690	∗
Taylor	1.0570	
Bernanke	1.6199	
Recalibrated Taylor	1.0086	
Recalibrated Bernanke	0.6493	∗
GR Recalibrated Experts	0.6635	∗
GR Recalibrated Experts EW	0.7064	

As a rough guide, using the Harvey et al. (1997) small-sample adjustment of the Diebold and Mariano (1995) test, ∗ denotes 5% and ∗∗ 1% significance in the two-sided test against the GR benchmark, respectively.

respectively.

In contrast, the recalibrated experts, rows 8 and 9, perform considerably better than the naive experts. The recalibrated Taylor expert matches approximately the benchmark and the recalibrated Bernanke expert is around 35% better than the benchmark. Recalibration is efficient in terms of maximising the performance of each individual expert.

Turning to the GR regression results with the recalibrated experts, the recursive weight combination, row 10, performs well with around a 33% improvement over the benchmark. The equal weight GR regression with recalibrated experts, row 11, has a slightly weaker performance at around 30% better than the benchmark. Regardless of the weights, neither GR regression combination based on recalibrated experts matches the performance of the recalibrated Bernanke expert.

Digging deeper into the real-time performance, Table 2 displays recursively computed RMSFE (not relative to the benchmark) for the expanding evaluation sample from 2000:2

to the dates shown in the first column. For example, row 2 shows the RMSFE computed using data from 2000:2 to 2001:1 for various combinations. The consecutive columns (left to right) refer to the GR Regression, the GR Regression with Equal Weights (GR EW), the Copula Combination (CC) and the Copula Combination with Equal Weights (CC EW). The best performing combination for each row is marked in bold. (Figure A2 in the not for publication appendix plots the RMSFE for the four combinations shown in Table 2.) There are three main findings. First, the copula point combinations dominate the GR regressions, with (recursive) gains in performance over the benchmark occurring both before and after the financial crisis. The GR approach is the best combination method only for the sample to 2001:1; thereafter the copula combination is superior. Second, copula combination performs less well with equal weights than with recursive weights throughout, with little time variation after 2003 in the performance gap between the two. Third, the equal weights GR regression outperforms the recursive weights variant for much of the evaluation sample, although the recursive weights benchmark performs relatively well from 2013.

The recursively computed RMSFE for the recalibrated experts are displayed in Table 3, with the second column displaying the results from the copula combination to facilitate comparisons. The subsequent columns (left to right) refer to the recalibrated Taylor expert, the recalibrated Bernanke expert, the GR regression with recalibrated experts and recursive weights and the GR regression with recalibrated experts and equal weights. (Figure A3 in the not for publication appendix plots the RMSFE for the five combinations shown in Table 3.) For all recursive evaluations ending in the dates shown in the first column, the copula point combination dominates the GR regression with recalibrated experts, regardless of whether the latter uses recursive or equal weights. The GR regression with recursive weights and recalibrated experts performs better than the variant with equal

Table 2: Recursive RMSFE, Federal Funds Rate, by Combination

	GR	GR EW	CC	CC EW
2001:1	**0.5502**	0.8407	0.7289	0.8943
2002:1	1.0653	0.7737	**0.6593**	0.7977
2003:1	0.9080	0.6644	**0.5519**	0.6613
2004:1	0.7961	0.6126	**0.4873**	0.5849
2005:1	0.7533	0.6711	**0.4661**	0.5457
2006:1	0.6974	0.6216	**0.5046**	0.6226
2007:1	0.6583	0.5847	**0.4715**	0.5873
2008:1	0.6721	0.5992	**0.5270**	0.6422
2009:1	0.9323	0.9115	**0.8147**	0.9807
2010:1	1.2655	1.1286	**0.7737**	0.9325
2011:1	1.2433	1.2226	**0.7388**	0.8898
2012:1	1.2091	1.2209	**0.7075**	0.8521
2013:1	1.1647	1.2028	**0.6799**	0.8190
2014:1	1.1284	1.1765	**0.6556**	0.7893
2015:1	1.1408	1.1558	**0.6353**	0.7632

weights in the aftermath of the financial crisis but otherwise they are fairly evenly matched in general. The recalibrated Bernanke expert performs particularly well—nearly always superior to the GR regression with recursive weights and recalibrated experts. This expert also outperforms the GR combinations shown in the previous table.

Summarising the results based on recursive RMSFE, the copula point combinations perform consistently well, with strong evidence that equal weights gives weaker predictive performance. Whether the GR regression uses equal or recursive weights with naive experts makes little difference by the end of the evaluation, but there are times when the equal weights variant is the better of the two. When using recalibrated experts, the GR regression does better, but not as well as the recalibrated Bernanke expert typically, and not as well as the copula point combination.

Table 3: Recursive RMSFE, Federal Funds Rate, Recalibrated Experts and Combinations

	CC	Recal Taylor	Recal Berna	GR Recal	GR Recal EW
2001:1	0.7289	**0.3732**	0.7571	0.3865	0.4809
2002:1	**0.6593**	1.3931	0.8045	1.0306	1.0386
2003:1	**0.5519**	1.1458	0.8796	0.9748	0.9064
2004:1	**0.4873**	1.0452	0.7636	0.8501	0.8060
2005:1	**0.4661**	1.2151	0.7062	0.7784	0.8049
2006:1	**0.5046**	1.3483	0.7109	0.7269	0.7815
2007:1	**0.4715**	1.3268	0.6649	0.6755	0.7484
2008:1	**0.5270**	1.3105	0.6798	0.6987	0.7791
2009:1	**0.8147**	1.4761	0.9288	0.9544	1.0226
2010:1	**0.7737**	1.4006	0.8815	0.9056	0.9704
2011:1	**0.7388**	1.3357	0.8406	0.8636	0.9253
2012:1	**0.7075**	1.2803	0.8057	0.8278	0.8871
2013:1	**0.6799**	1.2302	0.7750	0.7959	0.8526
2014:1	**0.6556**	1.1860	0.7493	0.7690	0.8231
2015:1	**0.6353**	1.1506	0.7407	0.7570	0.8059

4 Application: Real US GDP Growth Predictions

In the second application, we consider point combinations by the decision maker for quarterly real US GDP growth for the target dates 1992:4 to 2017:1, based on the predictions of two experts.

4.1 Forecasts from the Experts

Motivated by earlier research by Adrian, Boyarchenko and Giannone (2018) on vulnerable growth, the experts use variants of the following model:

$$YG_t \;=\; \gamma_{i,1} \;+\; \gamma_{i,FC}FC_{t-1} \;+\; \gamma_{i,YG}YG_{t-1} \;+\eta_{i,t} \tag{7}$$

where YG_t and FC_t denote real output growth (annualised, quarter on quarter) and an indicator of financial conditions in time t, respectively. The first of the two experts, indexed by $i = 1$, uses credit spreads to capture financial conditions and ignores the autoregressive term, so that $\gamma_{1,YG} = 0$. We refer to this expert as Expert FC. The second expert, indexed by $i = 2$, ignores financial conditions but allows for persistence in output growth, with $\gamma_{2,FC} = 0$. We refer to this as Expert AR. The unrestricted parameters for each expert were calibrated using a copula modelling approach. In both cases, the parameters were based on sample data from 1973:1 to 2017:1, with (kernel-smoothed) ECDF marginals for the target and the explanatory variable. For Expert AR, the dependence between the target variable and the explanatory variable was fitted with a Gaussian copula, giving an estimated (linear, Spearman's) dependence parameter of 0.3575. For Expert FC, dependence was fitted using a rotated Clayton copula, with parameter 0.3358. This Archimedean copula gives asymmetric dependence, with stronger (negative) dependence in the left tail. Put differently, Expert FC allows the fitted dependence between output growth and financial conditions to be stronger during instances of low output growth. The forecasts from both experts were recalibrated using the (full sample) marginal distribution for the target variable. In contrast, Adrian, Boyarchenko and Giannone (2018) use a quantile regression framework to capture asymmetry in the forecast densities for output growth conditional on financial conditions and do not consider forecast combinations of experts who separate dependence from marginal distributions.

We emphasise that, unlike our previous application but like the analysis by Adrian, Boyarchenko and Giannone (2018), this is not a real-time forecasting exercise. Nevertheless, our forecast evaluation uses recursively computed RMSFE. Data considerations are discussed in the data appendix.

Our two experts make quarterly predictions for forecast target dates from 1992:4 to 2017:1. All forecasts are one step ahead—predictions made for the t observation based on information sets containing observations dated time $t-1$ and earlier. (The decision maker estimates combination weights based on evaluations to $t-1$.) Figure 5 plots the recursive predictions, together with the target variable, real US output growth (black line). Both experts track the conditional mean of the series, but with considerable volatility in the case of Expert AR (blue line). Expert FC (red line) captures the low frequency movements in the target variable particularly well and responds to the onset of the recent financial crisis in a timely manner. Expert AR captures the depth of the trough but the forecasts (naturally) lag the target variable.

4.2 Forecast Combination and Estimation

The decision maker recursively combines the forecasts from the two experts using an expanding window for combination weight estimation from 1992:4 to 2016:4, with weight training from 1973:2 to 1992:3. The heterogeneous marginal distributions for the target variable and the individual forecasts from the two experts are fitted using the respective kernel-smoothed ECDFs using an expanding window.[4] For illustrative purposes, Figure 6 plots the corresponding Probability Density Function (PDF, black line) based on the output growth data to 2016:4. Figure 6 also plots the PDFs for Expert FC (red line) and Expert AR (blue line). The marginal for output growth displays a number of characteristics identified by Adrian, Boyarchenko and Giannone (2018). In particular, there is evidence of non-normality in the upper (right) tail and the lower (left) tail, with the latter being relatively high mass and revealing the impact of sharp contractions—vulnerability. The upper tail is somewhat sparse, with a noticeable bulge slightly to the right of the mode—reflecting the pronounced upswing of the business cycle in the sample data. The marginal for Expert AR matches the tails for the target variable, but with less mass near the peak, and little sign of non-normality. The marginal for Expert FC has left skew but less mass in both tails than for output growth. For the target variable and the predictions from Expert FC, the null hypothesis of normality is rejected at the 1% significance level for the Shapiro-Wilk test. The predictions from Expert AR are approximately normally distributed.

Recall that for each forecast origin, the decision maker constructs the (univariate) Probability Integral Transforms (PITS) and converts each PITS series to be a Gaussian-distributed pseudo variable via the inverse CDF. Then, the dependence parameters (of the Gaussian copula) are estimated from the pseudo regression, Equation (5). The copula point combination is mapped onto the unit interval using the normal CDF and then onto the observable scale for output growth using the inverse of the kernel-smoothed ECDF.

[4] The not for publication appendix contains results for the homogeneous marginals case.

4.3 Results

The forecast evaluations use target data for US real quarterly output growth from 1992:4 to 2017:1. Figure 7 displays the predictions from the copula point combination (red line), together with the GR regression benchmark (blue line). The copula point combination has a zero intercept, bounded weights (estimated recursively, on the unit interval) and weights that sum to one across experts. Eyeballing the forecasts, the copula methodology provides fairly accurate point predictions of the low frequency shifts in the conditional mean of US output growth (black line). During the financial crisis, the magnitude of the slump in the observed data is just over twice the size of the copula point combination forecast. The GR regression benchmark performs slightly less well throughout the evaluation sample, with a tendency towards greater volatility for moderate sized fluctuations than from the copula combination. Nevertheless, the two combination forecasts are fairly closely matched throughout the evaluation. In contrast, the two approaches gave very different forecasts for the Federal Funds rate application considered previously, where the marginal distributions exhibited stronger departures from symmetry.

Figure 8 displays the recursively estimated weights from the copula point combination (solid lines) and the benchmark GR regression (dotted lines). The GR regression weights for Expert FC (red dotted line) display greater time variation than for the same expert with the copula point combination (solid red line). There is a pronounced step up in the benchmark combination weight on this expert during the financial crisis, reflecting the importance of financial conditions during that era. In general, the copula point combination weights exhibit greater stability, with greater weight on Expert FC, and with the weight on this expert increasing non-monontonically throughout the sample from around 60% to 67%. Regardless of the point combination methodology, the weight on Expert FC typically exceeds 50%.

Table 4: Relative RMSFE, Output Growth

Expert/Combination		
Granger-Ramanathan (GR)	1.0000	
GR Equal Weight (EW)	1.0105	
Copula Combination	0.9584	$*$
Copula Combination EW	0.9697	
Expert FC	0.9751	
Expert AR	1.2885	

As a rough guide, using the Harvey et al. (1997) small-sample adjustment of the Diebold and Mariano (1995) test, $*$ denotes 5% and $**$ 1% significance in the two-sided test against the GR benchmark, respectively.

Table 4 displays end-evaluation relative RMSFE results for the combinations and experts, with the GR regression as a benchmark (with unrestricted recursively estimated parameters). The equal weight GR regression, row 3, has a 1% higher RMSFE than the benchmark. The copula point combination, row 4, provides the best RMSFE, with a performance gain of just over 4% relative to the benchmark. Recall that this version of the copula point combination has restricted coefficients. Relaxing these restrictions lifts the relative RMSFE by around 5% relative to the benchmark. The equal weight copula combination, row 5, is slightly less effective than the recursive weight equivalent, with a performance gap of around 1%. Expert FC, row 6, outperforms the GR regression by around 2.5%. Expert AR, row 7, is inferior to the benchmark, with a performance difference of approximately 29%. In this forecasting output growth application, the recursive copula combination is the best of this set of experts and combinations, but the performance gain is modest.

Table 5 displays the recursively computed RMSFEs (not normalised) for expanding window forecast evaluations, starting in 1992:4 and ending in the dates shown in the first column. The remaining four columns refer to the benchmark GR Regression, the

GR Regression with Equal Weights (GR EW), the Copula Combination (CC) and the Copula Combination with Equal Weights (CC EW), respectively. The best performing combination for each row is marked in bold. (Figure A4 in the not for publication appendix plots the four combinations shown in Table 5.) There are three main findings. First, the recursive weight copula point combination typically provides the lowest RMSFE but the performance gain is often modest—around 1% for much of the pre-crisis sample, rising to around 4% by the end of the evaluation. Second, the equal weight copula combination performs less well than the recursive weight version for most of the evaluation. Third, by the end of the sample, recursive weights dominate their equal weight counterparts for both the copula point combinations and the GR regressions.

Summarising the results based on recursive RMSFE for this forecasting output growth application, the copula point combinations perform somewhat better than the benchmark, with recursive weights typically giving better predictive performance than equal weights. But the performance gain over the benchmark GR regression is much smaller than for the Federal Funds rate application.

5 Conclusions

Our methodology for copula point combinations builds directly on the GR regression approach by allowing for asymmetrically distributed variables. More general dependence structures between the experts' forecasts and the target variable are feasible and should be explored in subsequent research with copula combinations. One such appealing extension would be to allow the combination dependence to vary in the tails of the distributions. Hierarchical Archimedean copulas offer a feasible route in that direction, extending the framework proposed in this paper.

Table 5: Recursive RMSFE, Economic Growth, by Combination

	GR	GR EW	CC	CC EW
1993:1	2.5852	2.6180	2.4998	**2.4740**
1994:1	1.8850	1.9829	**1.8358**	1.8393
1995:1	1.9997	2.1682	**1.9560**	1.9898
1996:1	1.7704	1.9086	**1.7582**	1.7656
1997:1	1.8565	2.0061	**1.8310**	1.8617
1998:1	1.8036	1.9482	**1.7769**	1.8126
1999:1	1.7791	1.8975	**1.7604**	1.7843
2000:1	1.8737	1.9991	**1.8531**	1.8894
2001:1	2.2337	2.4029	**2.1989**	2.2658
2002:1	2.2441	2.4080	**2.2149**	2.2681
2003:1	2.1638	2.3086	**2.1327**	2.1837
2004:1	2.1615	2.3100	**2.1394**	2.1818
2005:1	2.0759	2.2183	**2.0547**	2.0956
2006:1	2.0330	2.1790	**2.0104**	2.0555
2007:1	2.0534	2.1951	**2.0250**	2.0723
2008:1	2.0877	2.1974	**2.0560**	2.0987
2009:1	2.3618	2.3363	**2.2465**	2.2624
2010:1	2.3122	2.3024	**2.1998**	2.2209
2011:1	2.3212	2.3122	**2.2124**	2.2312
2012:1	2.3008	2.3100	**2.1905**	2.2164
2013:1	2.2599	2.2661	**2.1573**	2.1804
2014:1	2.2911	2.3054	**2.1917**	2.2152
2015:1	2.2556	2.2802	**2.1556**	2.1851
2016:1	2.2095	2.2321	**2.1149**	2.1429
2017:1	2.1805	2.2034	**2.0898**	2.1145

References

Adrian,T., N. Boyarchenko and D. Giannone (2018) "Vulnerable Growth", *American Economic Review*, forthcoming.

Barnard, G.A. (1963) "New Methods of Quality Control", *Journal of the Royal Statistical Society A*, 126, 255-258.

Bassetti, F., R. Casarin and F. Ravazzolo (2018) "Bayesian Nonparametric Calibration and Combination of Predictive Distributions", *Journal of the American Statistical Association*, forthcoming.

Bates, J.M. and C.W.J. Granger (1969) "The Combination of Forecasts", *Operational Research Quarterly*, 20, 4, 451-468.

Berkowitz, J. (2001) "Testing Density Forecasts, With Applications to Risk Management", *Journal of Business and Economic Statistics*, 19, 4, 465-474.

Bernanke, B. (2015) "The Taylor Rule: A benchmark for monetary policy", Brookings, blog, April 28.

Billio, M., R. Casarin, F. Ravazzolo and H. van Dijk (2013) "Time-varying Combinations of Predictive Densities using Nonlinear Filtering", *Journal of Econometrics*, 177, 2, 213-232.

Bjørnland, H., K. Gerdrup, A. Jore, C. Smith and L.A. Thorsrud (2011) "Does Forecast Combination Improve Norges Bank Inflation Forecasts?", *Oxford Bulletin of Economics and Statistics*, 74, 163-179.

Capistrian, C. and A. Timmermann (2009) "Forecast Combination with Entry and Exit of Experts", *Journal of Business and Economic Statistics*, 27, 428-440.

Carlstrom, C.T. and M. Jacobson (2015) "Do Forecasters Agree on a Taylor Rule?", *Economic Commentary*, 2015-10, Federal Reserve Bank of Cleveland.

Clemen, R.T. and Winkler, R.L. (1986) "Combining Economic Forecasts", *Journal of*

Business and Economic Statistics 4, 39-46.

Conflitti, C., C. De Mol and D. Giannone (2015) "Optimal Combination of Survey Forecasts", *International Journal of Forecasting*, 31, 1096-1103.

Diebold, F.X., T.A. Gunther and A.S. Tay (1998) "Evaluating Density Forecasts with Applications to Financial Risk Management", *International Economic Review*, 39, 4, 863-883.

Del Negro, M., R.B. Hasegawa and F. Schorfheide (2016) "Dynamic Prediction Pools: An investigation of financial frictions and forecasting performance", *Journal of Econometrics*, 192 (2), 391-405.

Diebold, F.X. and R.S. Mariano (1995) "Comparing Predictive Accuracy", *Journal of Business and Economic Statistics*, 13, 253-263.

Diebold, F.X. and M. Shin (2018) "Machine Learning for Regularised Survey Forecast Combination: Partially-Egalitarian Lasso and its Derivatives", *International Journal of Forecasting*, in press, NBER Working Paper 24967.

Galbraith, J.W. and S. van Norden (2012) "Assessing Gross Domestic Product and Inflation Probability Forecasts Derived from Bank of England Fan Charts", *Journal of the Royal Statistical Society*, Series A (Statistics in Society), 175, 3, 713-727.

Geweke, J. and G. Amisano (2011) "Optimal Prediction Pools", *Journal of Econometrics*, 164, 130-141.

Genest, C. and J.V. Zidek (1986) "Combining Probability Distributions: A critique and an annotated bibliography", *Statistical Science*, 1, 1, 114-135.

Giglio, S., B. Kelly and S. Pruitt (2016) "Systemic Risk and the Macroeconomy: An empirical evaluation", *Journal of Financial Economics*, 119, 3, 457-471.

Gneiting, T., A.E. Raftery, A.H. Westveld III and T. Goldman (2005) "Calibrated Probabilistic Forecasting Using Ensemble Model Output Statistics and Minimum CRPS

Estimation", *Monthly Weather Review*, 133, 1098-1118.

Granger, C. and R. Ramanathan (1984) "Improved Methods of Combining Forecasts", *Journal of Forecasting*, 3, 197-204.

Harvey, D., S. Leybourne and P. Newbold (1997) "Testing the Equality of Prediction Mean Squared Errors", *International Journal of Forecasting*, 13, 281-291.

Henderson, D.W. and W.J. McKibbin (1993) "A Comparison of Some Basic Monetary Policy Regimes for Open Economies: Implications of different degrees of instrument adjustment and wage persistence", *Carnegie-Rochester Conference Series on Public Policy*, 39, 221-317.

Joe, H. (2006) "Generating Random Correlation Matrices Based on Partial Correlations", *Journal of Multivariate Analysis*, 97, 10, 2177-2189.

Johanssen, B. K. and E. Mertens (2016) "A Time Series Model of Interest Rates with the Effective Lower Bound", FEDS Working Paper, 2016-033.

Jouini, M.N. and R.T. Clemen (1996) "Copula Models for Aggregating Expert Opinion", *Operations Research*, 44, 3, 444-457.

Jore, A.S., J. Mitchell and S.P. Vahey (2010) "Combining Forecast Densities from VARs with Uncertain Instabilities", *Journal of Applied Econometrics*, 25, 621-634.

Lima, L.R. and F. Meng (2017) "Out of Sample Return Predictability: A Quantile Combination Approach", *Journal of Applied Econometrics*, 32, 4, 877-895.

Loaiza-Maya, R. and M.S. Smith (2018) "Real-Time Macroeconomic Forecasting with a Heteroskedastic Inversion Copula", *Journal of Business and Economic Statistics*, forthcoming.

Madansky, A. (1964) "Externally Bayesian Groups", Santa Monica, CA: RAND Corporation.

Meucci, A. (2005) "Beyond Black-Litterman in Practice: A five-step recipe to input views on non-normal markets", *Risk*, 19, 114-119.

Nelsen, R. (2006) *An Introduction to Copulas*, 2nd ed., New York: NY: Springer.

Patton, A.J. (2006) "Modelling Asymmetric Exchange Rate Dependence", *International Economic Review*, 47, 2, 527-556.

Patton, A.J. (2012) "A Review of Copula Models for Economic Time Series", *Journal of Multivariate Analysis*, 110, 4-18.

Ranjan, R. and T. Gneiting (2010) "Combining Probability Forecasts", *Journal of the Royal Statistical Society Series B*, 72, 71-91.

Rosenblatt, M. (1952) "Remarks on a Multivariate Transformation", *The Annals of Mathematical Statistics* 23, 3, 470-472.

Rossi, B. and T. Sekhposyan (2018) "Alternative Tests for Correct Specification of Conditional Predictive Densities", *Journal of Econometrics*, forthcoming.

Salgado, S., F. Guvenen and N. Bloom (2016) "Skewed Business Cycles", unpublished manuscript.

Schefzik R., T. Gneiting and T.L. Thorarinsdottir (2013) "Uncertainty Quantification in Complex Simulation Models Using Ensemble Copula Coupling", *Statistical Science*, 28 (4), 616-640.

Scotti, C. (2011) "A Bivariate Model of Federal Reserve and ECB Main Policy Rates", *International Journal of Central Banking*, 7, 3, 37-78.

Smith, M.S. and W. Maneesoonthorn (2018) "Inversion Copulas from Nonlinear State Space Models", *International Journal of Forecasting*, 34, 3, 389-407.

Smith, M.S. and S.P. Vahey (2016) "Asymmetric Forecast Densities for U.S. Macroeconomic Variables from a Gaussian Copula Model of Cross-Sectional and Serial Dependence", *Journal of Business and Economic Statistics*, 34, 3, 416-434.

Timmermann, A. (2006) "Forecast Combinations", Handbook of Economic Forecasting, 135-196.

Taylor, J.B. (1993) "Discretion versus Policy Rules in Practice", *Carnegie-Rochester Conference Series on Public Policy*, 39, 195-214.

Taylor, J.B. (1999) "A Historical Analysis of Monetary Policy Rules", in J.B. Taylor, ed., *Monetary Policy Rules*, Chicago: University of Chicago Press, 319-341.

Waggoner, D.F. and T. Zha (2012) "Confronting Model Misspecification in Macroeconomics", *Journal of Econometrics*, 171, 2, 167-184.

Yellen, J.L. (2012) "Perspectives on Monetary Policy", Speech to the Boston Economic Club Dinner, Boston, Massachusetts, June 6.

Data Appendix

A. Federal Funds Rate Application

We use the same data as Bernanke (2015). The data can be downloaded from Brookings. The quarterly seasonally adjusted real GDP, nominal GDP, PCE and core PCE variables come from the Federal Reserve Bank of Philadelphia's Real-Time Data Center for Macroeconomists, which maintains the database containing historical vintages of these series. Inflation is the year-on-year change between $t-5$ and $t-1$ corresponding to the latest available vintage of data.

The Center also publishes the Federal Reserve Greenbook estimates of the output gap for 1987 to 2006. For 2007 to 2009, Bernanke uses the Greenbook output gap estimates available from Federal Reserve Board of Governors. Bernanke uses the output gap estimates prepared for the first FOMC meeting of each quarter. For the post-2009 data, Bernanke calculates a real-time output gap using the CBO's estimates of potential output and the first estimates of real GDP from the Bureau of Economic Analysis. For the 1996:1 vintage, the observation for 1995:1 is missing, due to a delayed data release caused by a US federal government shutdown. In addition, the revised estimate is used rather than the real-time estimates for 2013:3 to 2013:4, due to comprehensive revisions by the BEA. The Federal Reserve Bank of St Louis' database, ALFRED, provides different vintages of estimates of potential output from the Congressional Budget Office (CBO). These vintages are compiled from the historical releases of the CBO's Budget and Economic Outlook.

The quarterly Federal Funds rate can be downloaded from FRED and is the average of the daily effective rate (not seasonally adjusted) in percent.

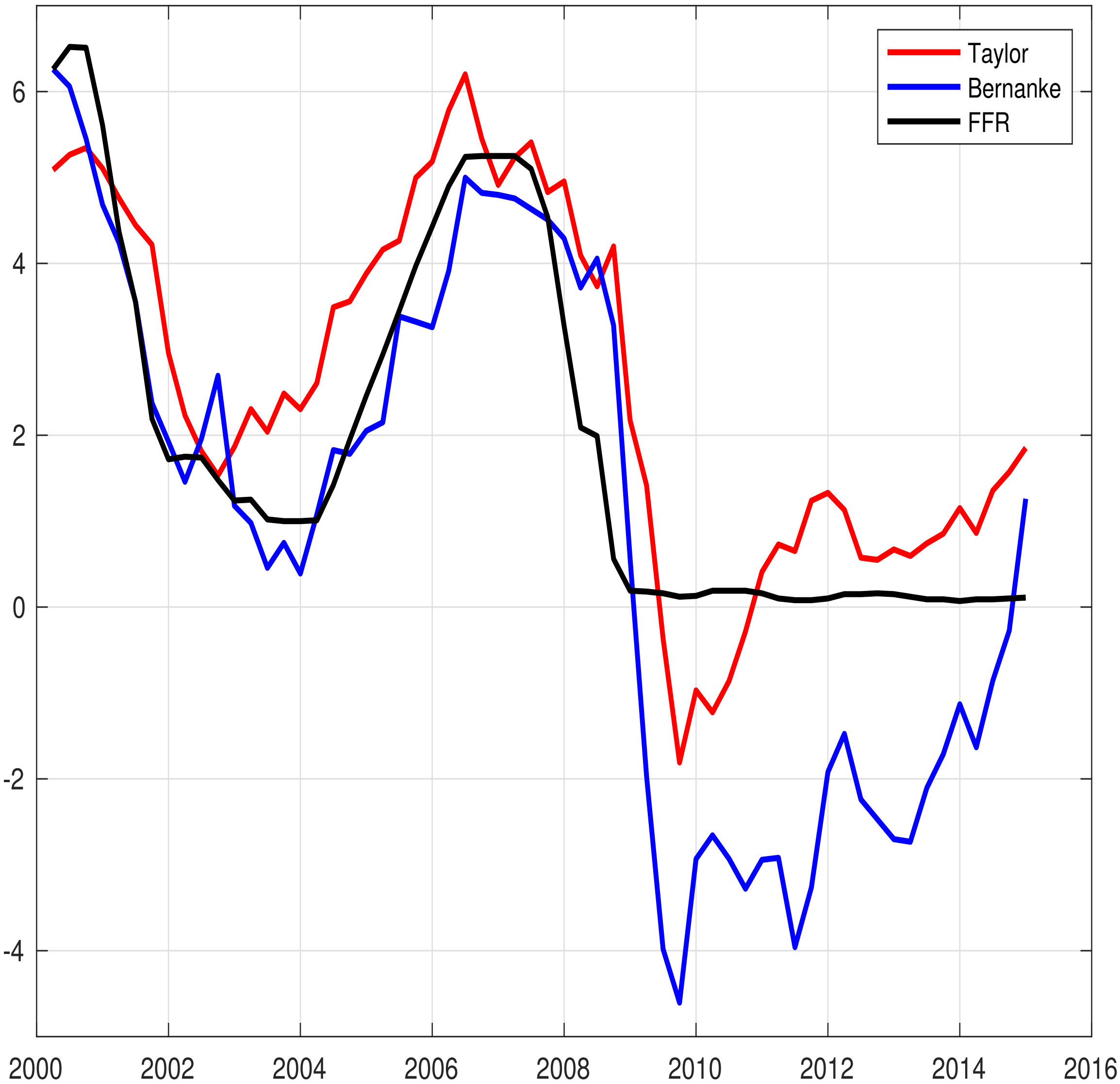
The Taylor expert uses the GDP deflator as the measure of inflation. The Bernanke expert, however, uses the FOMC's preferred real-time measure of inflation which is based on core personal consumption expenditure (PCE), excluding volatile food and energy prices. In addition, the Bernanke expert also uses a higher weight on the output gap, a coefficient of 1 rather than 0.5 used by the Taylor expert.

B. GDP Growth Application

The (revised) measure of economic activity for the target variable, following Adrian et al (2018), is the quarterly seasonally adjusted Real Gross Domestic Product as a percentage change from preceding period (annualised).

Our two experts use different explanatory variables to predict real output growth. Expert FC uses the spread between the Baa and Aaa corporate bonds, available from the Federal Reserve Bank of St Louis' FRED database (the observation from the third month of each quarter). Expert AR uses lagged real output growth as the explanatory variable.

FIGURE 1: FORECASTS FROM EXPERTS AND FFR



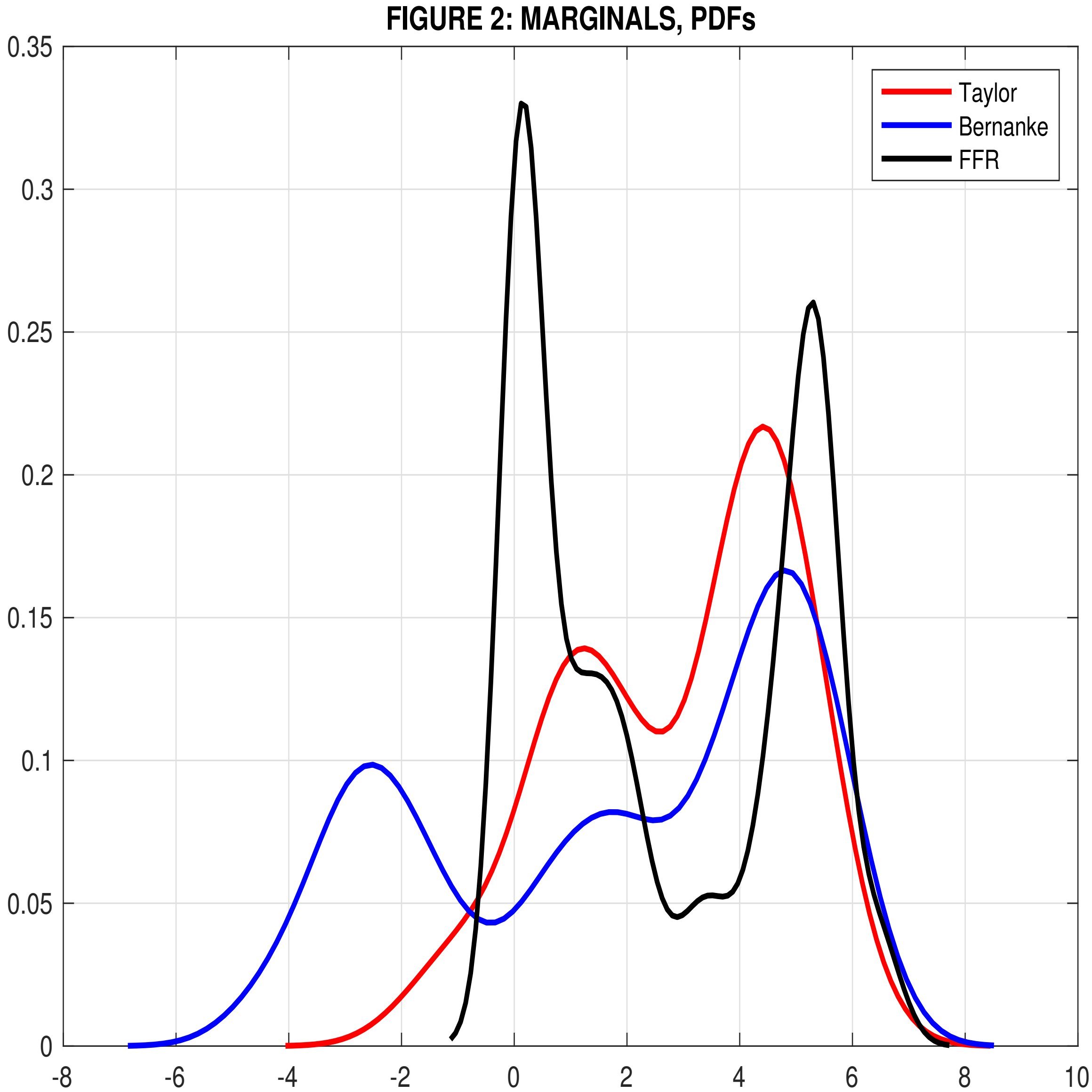
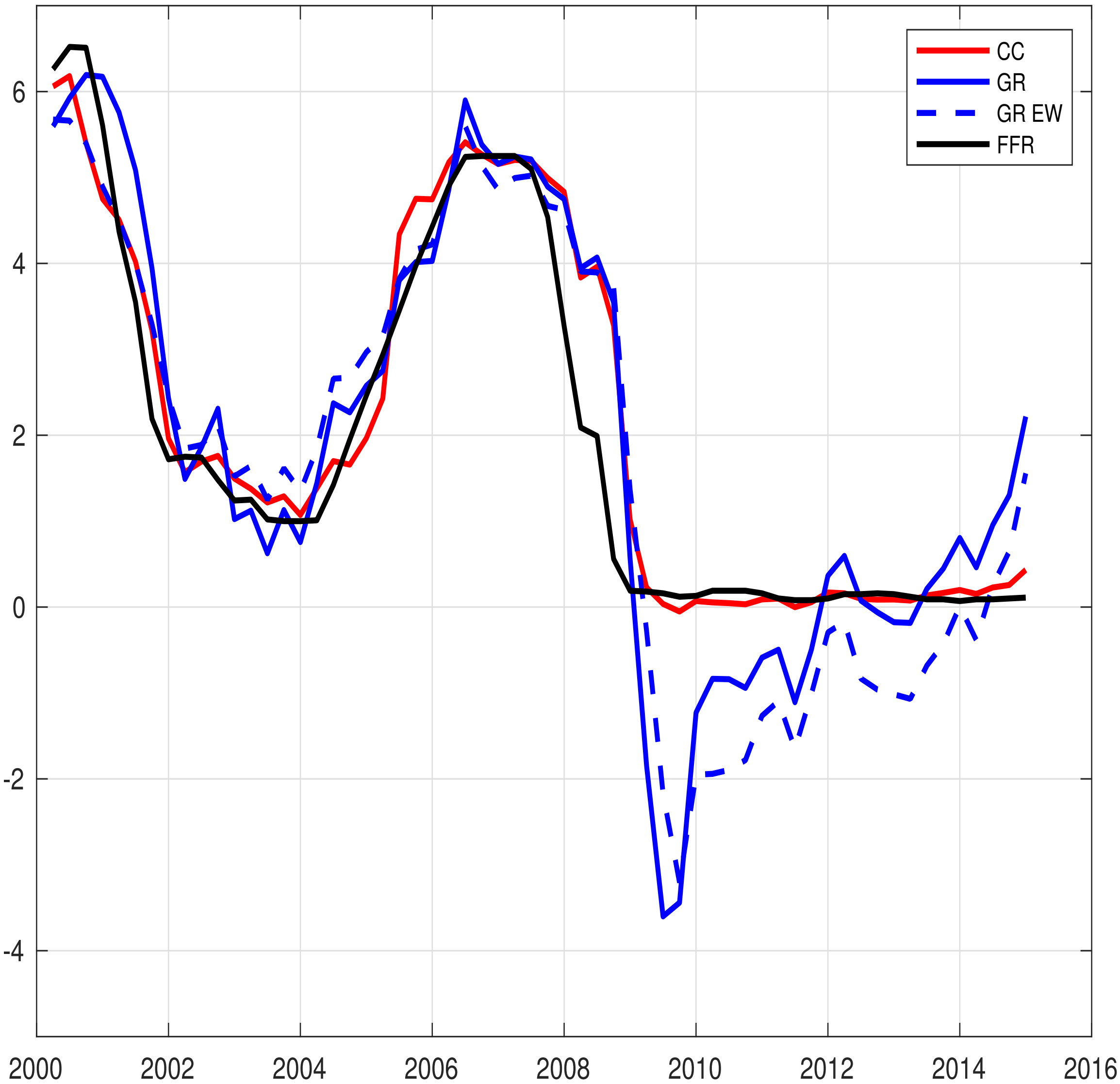


FIGURE 2: MARGINALS, PDFs
Taylor
Bernanke
FFR
0.35
0.3
0.25
0.2
0.15
0.1
0.05
0
-8
-6
-4
-2
0
2
4
6
8
10

FIGURE 3: COMBINATIONS AND FFR



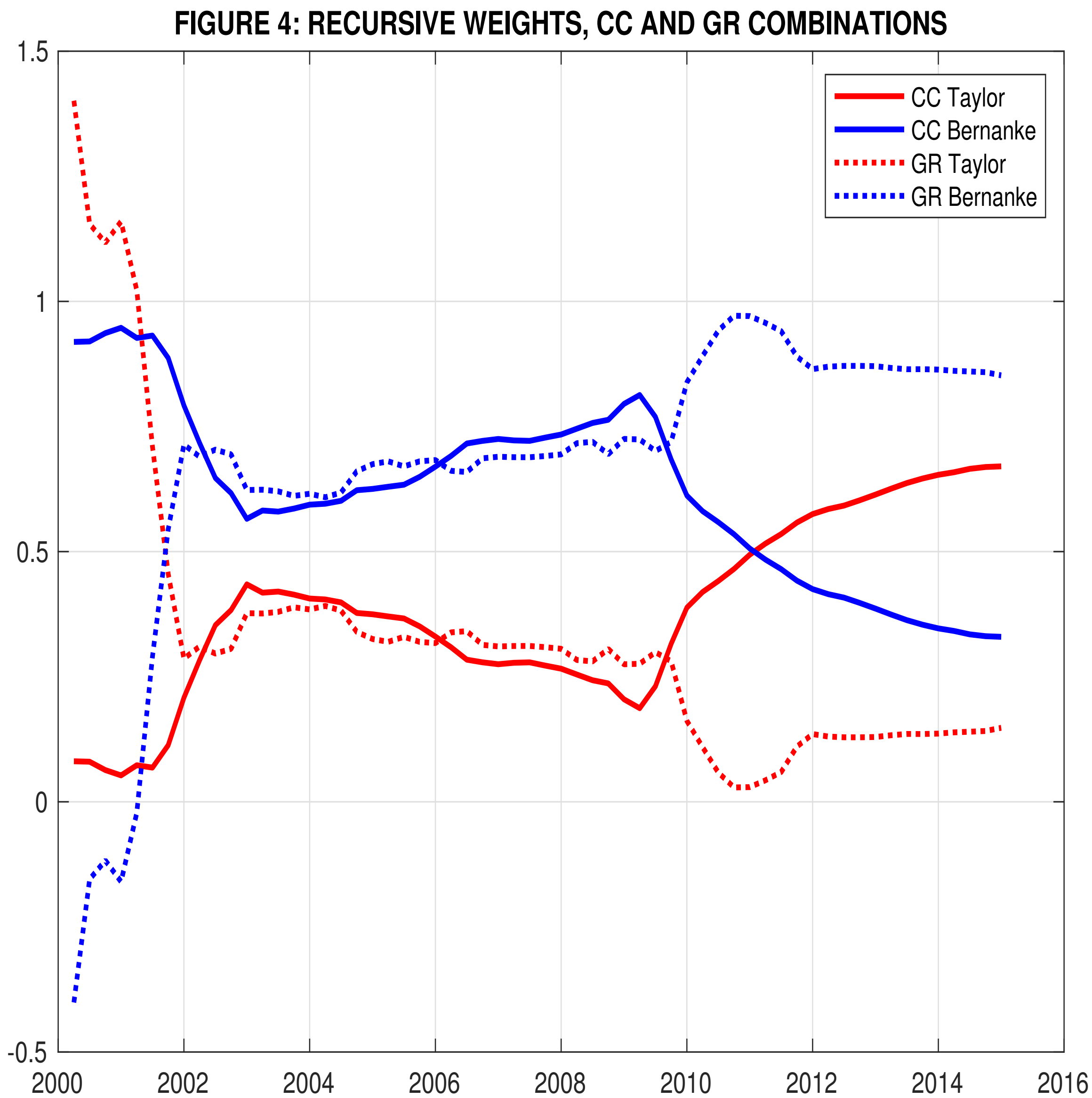
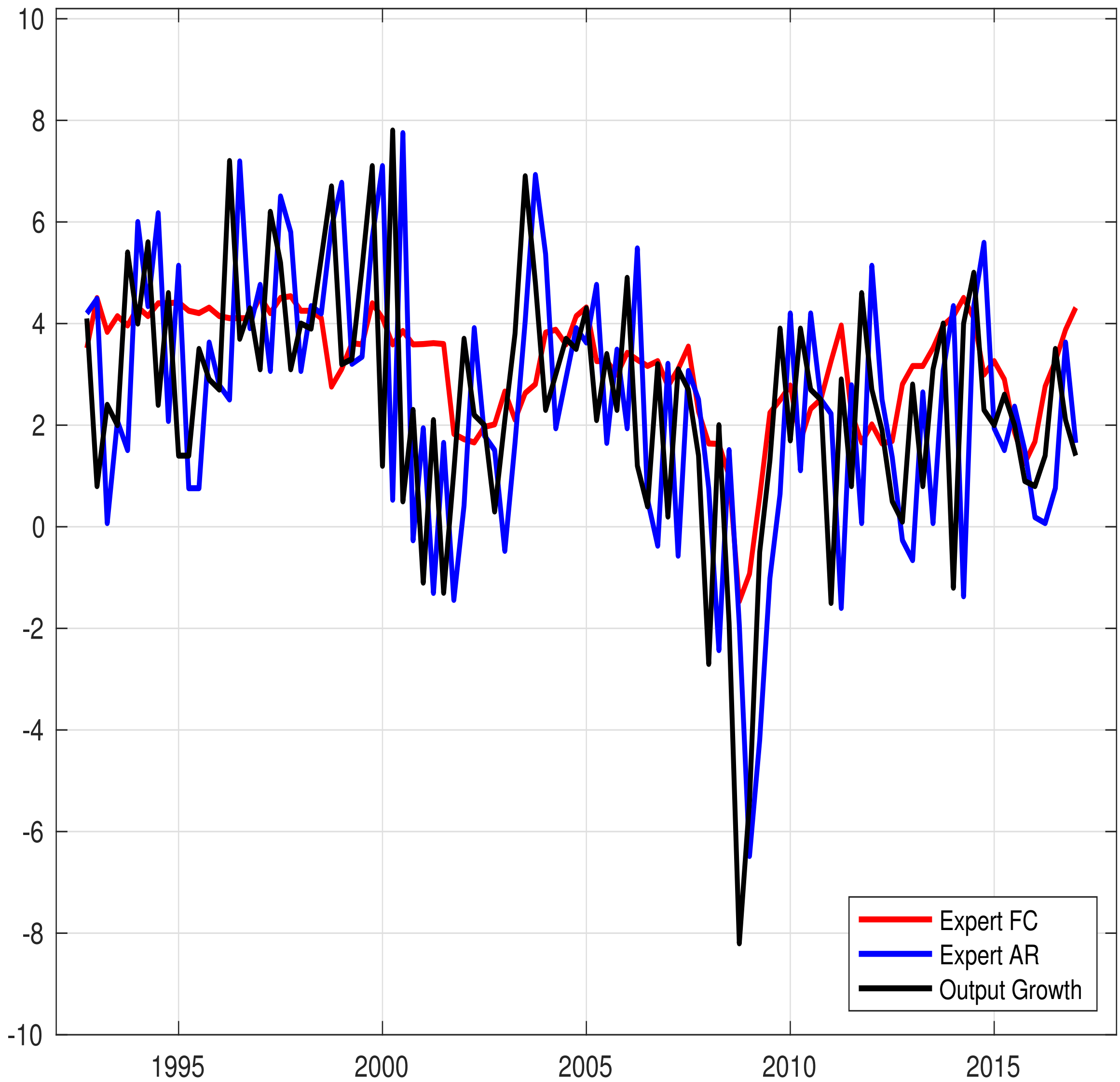


FIGURE 4: RECURSIVE WEIGHTS, CC AND GR COMBINATIONS
CC Taylor
CC Bernanke
GR Taylor
GR Bernanke
1.5
1
0.5
0
-0.5
2000
2002
2004
2006
2008
2010
2012
2014
2016

FIGURE 5: FORECASTS FROM EXPERTS AND OUTPUT GROWTH



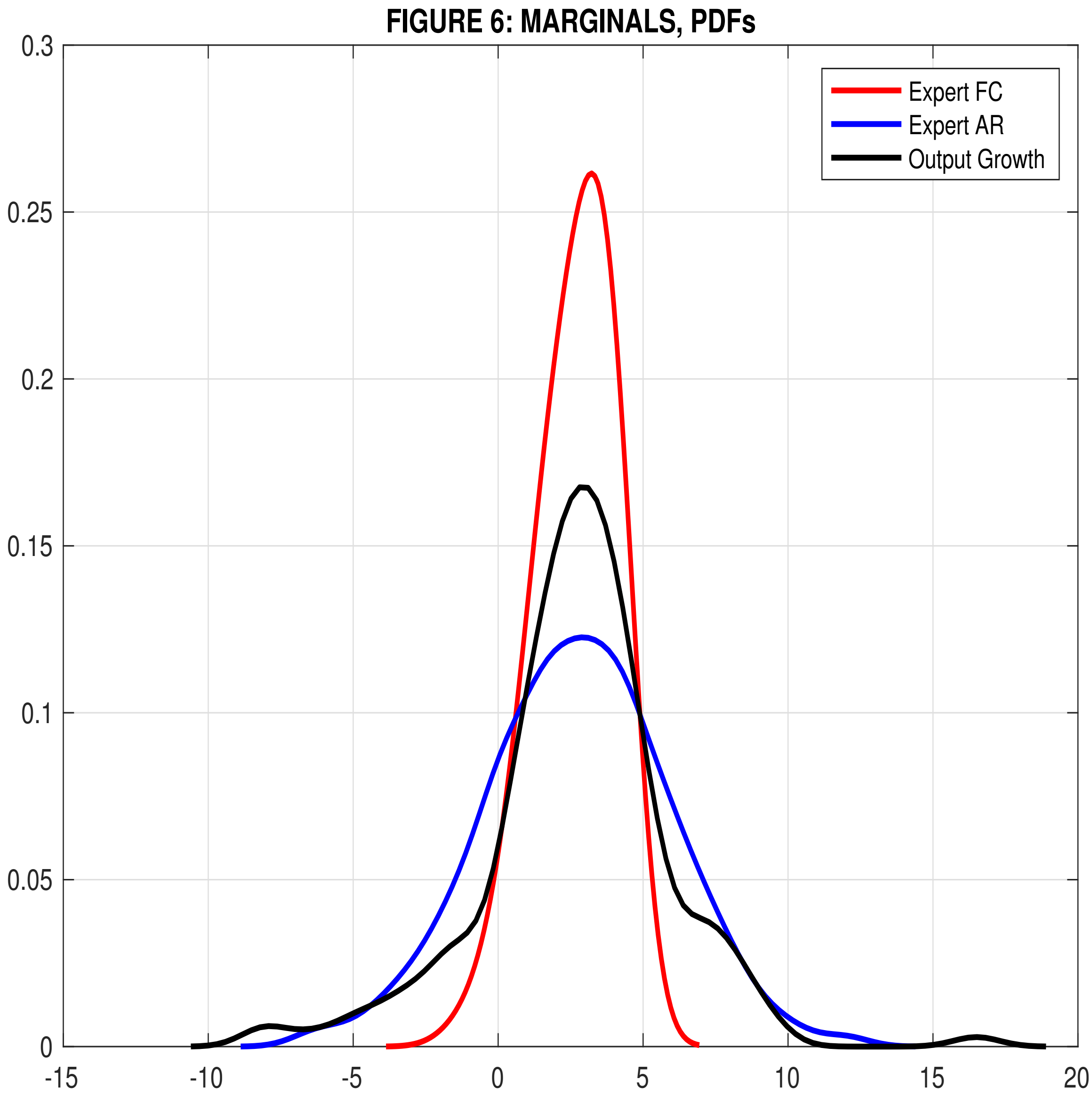
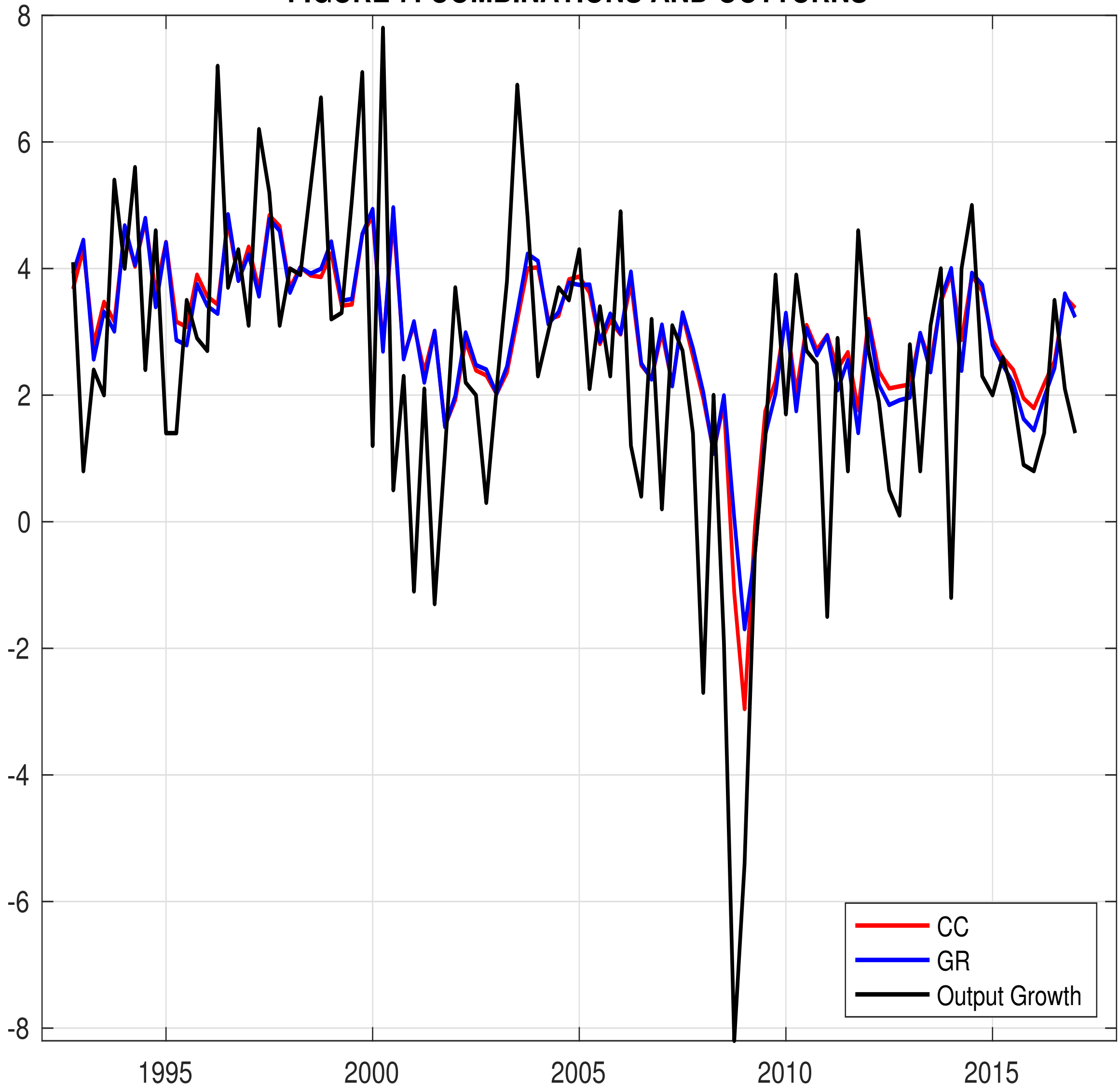


FIGURE 6: MARGINALS, PDFs
0.3
0.25
0.2
0.15
0.1
0.05
0
-15
-10
-5
0
5
10
15
20
Expert FC
Expert AR
Output Growth

FIGURE 7: COMBINATIONS AND OUTTURNS



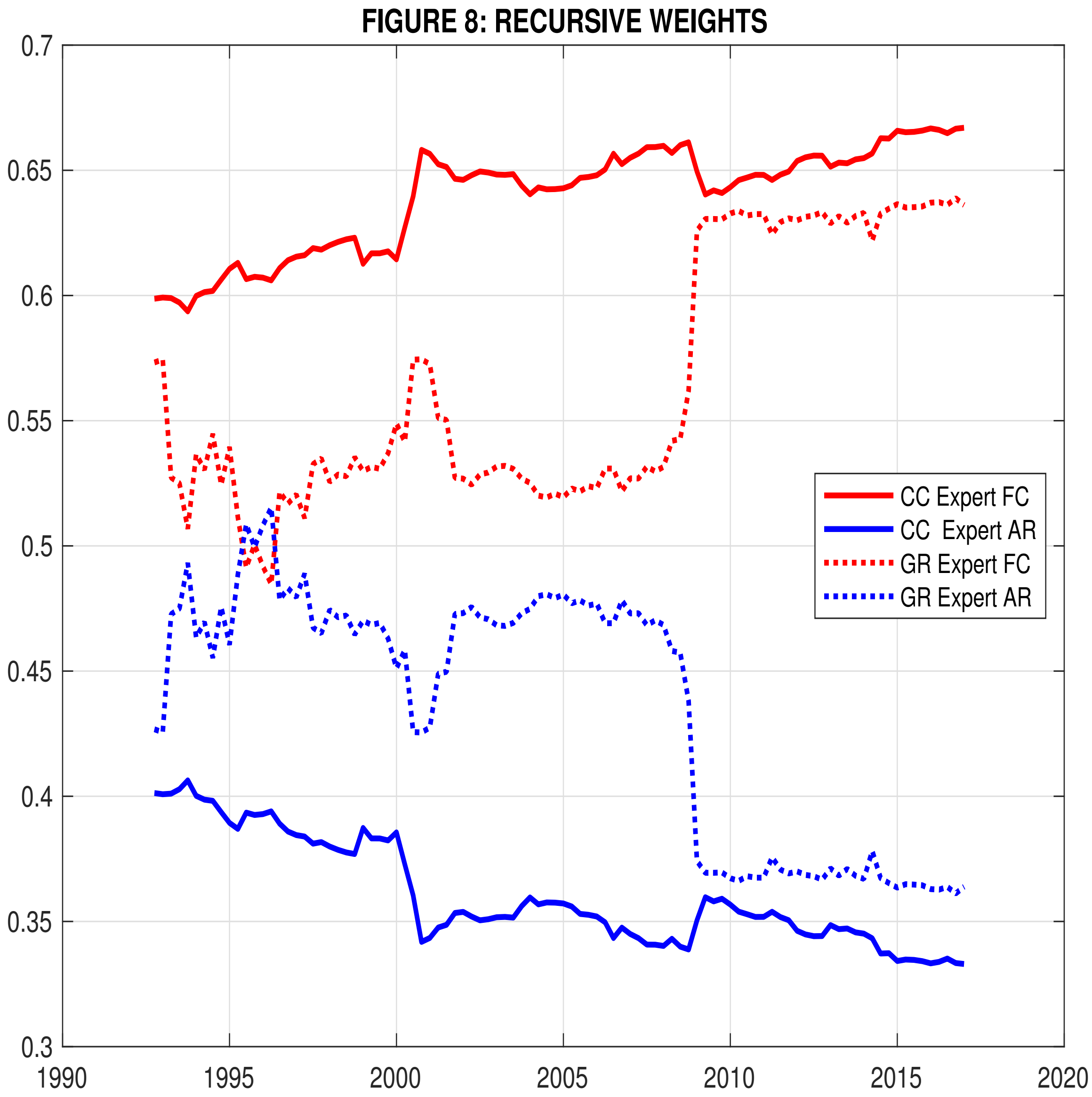


FIGURE 8: RECURSIVE WEIGHTS
0.7
0.65
0.6
0.55
0.5
0.45
0.4
0.35
0.3
1990
1995
2000
2005
2010
2015
2020
CC Expert FC
CC Expert AR
GR Expert FC
GR Expert AR

Not for publication appendix

"Improved Methods for Combining Point Forecasts for an Asymmetrically Distributed Variable"

Karagedikli, Vahey and Wakerly

June 22, 2019

A. Additional Figures for Applications

Figure A1 plots the predictions of recalibrated experts with recursively estimated weights (red solid line), together with the target variable, FFR_t (blue solid line). For each recalibrated expert, the fit to FFR_t is generally improved relative to the corresponding naive expert (shown in Figure 1) in the aftermath of the most recent financial crisis, when the recalibrated experts predict low but positive interest rates. Nevertheless, the forecast errors are typically larger (in absolute value) than for the copula point combination, especially for the Taylor expert.

Figure A2 displays the recursive RMSFE for various combinations and experts for the Federal Funds rate application. It plots the GR regression combination with recursive weights (blue line), the GR regression with equal weights (blue broken line), the copula point combination with recursive weights (red line) and the copula point combination with equal weights (red broken line). The copula combination dominates, although the gap to the GR benchmark narrows during 2008.

Figure A3 displays the recursive RMSFE for various combinations and recalibrated experts for the Federal Funds rate application. It plots the recalibrated Taylor and Bernanke experts (cyan lines, solid and broken, respectively), the GR regression combinations with recalibrated experts with recursive weights (magenta line) and with equal weights (magenta broken line). The copula point combination is displayed again (red line), together with the GR regression benchmark (blue line). The recalibrated combinations perform worse than the copula combination throughout the evaluation, and typically worse than the recalibrated Bernanke expert. For much of the sample performance of the GR combination with recalibrated experts is similar to the benchmark, but after the slump in 2008, the relative performance of the benchmark deteriorates.

Figure A4 displays the recursive RMSFEs for the vulnerable growth application. The copula point combination (red solid line) and its equally weighted counterpart (red broken line), together with the GR regression benchmark (blue line) and its equal weight counterpart (blue broken line). The recursive copula combination ranks best from 1995 onwards, although the gap between this and the GR benchmark is typically around 1% until the onset of the financial crisis. The equal weight copula combination and benchmark are fairly close for most of the sample, with the copula variant preferred after the GFC. The performance gap in favour of the recursive weight specification over the equal weight specification is slightly larger for the copula combination than for the GR regression.

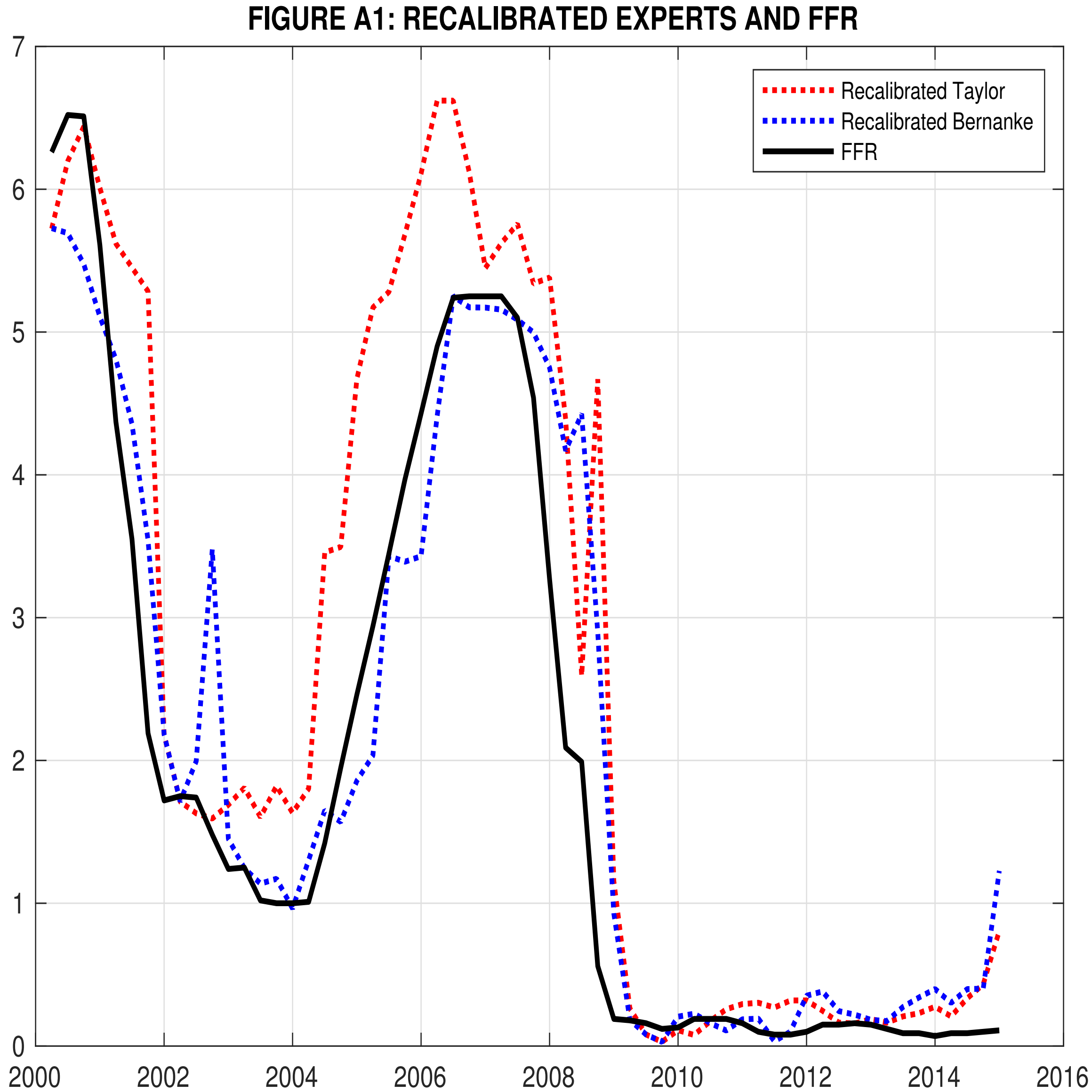
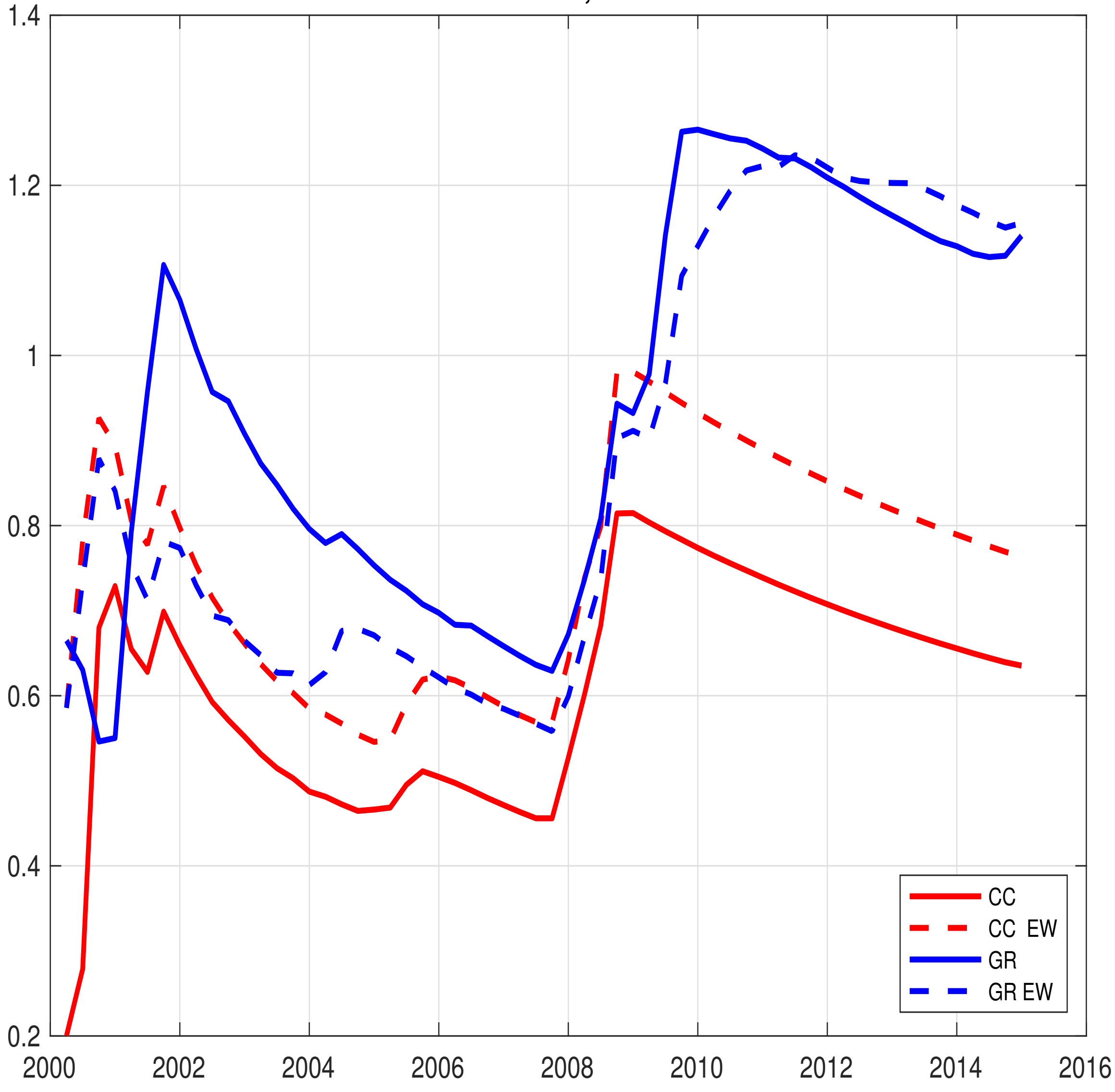


FIGURE A1: RECALIBRATED EXPERTS AND FFR
Recalibrated Taylor
Recalibrated Bernanke
FFR
0
1
2
3
4
5
6
7
2000
2002
2004
2006
2008
2010
2012
2014
2016

FIGURE A2: RECURSIVE RMSFE, CC AND GR COMBINATIONS



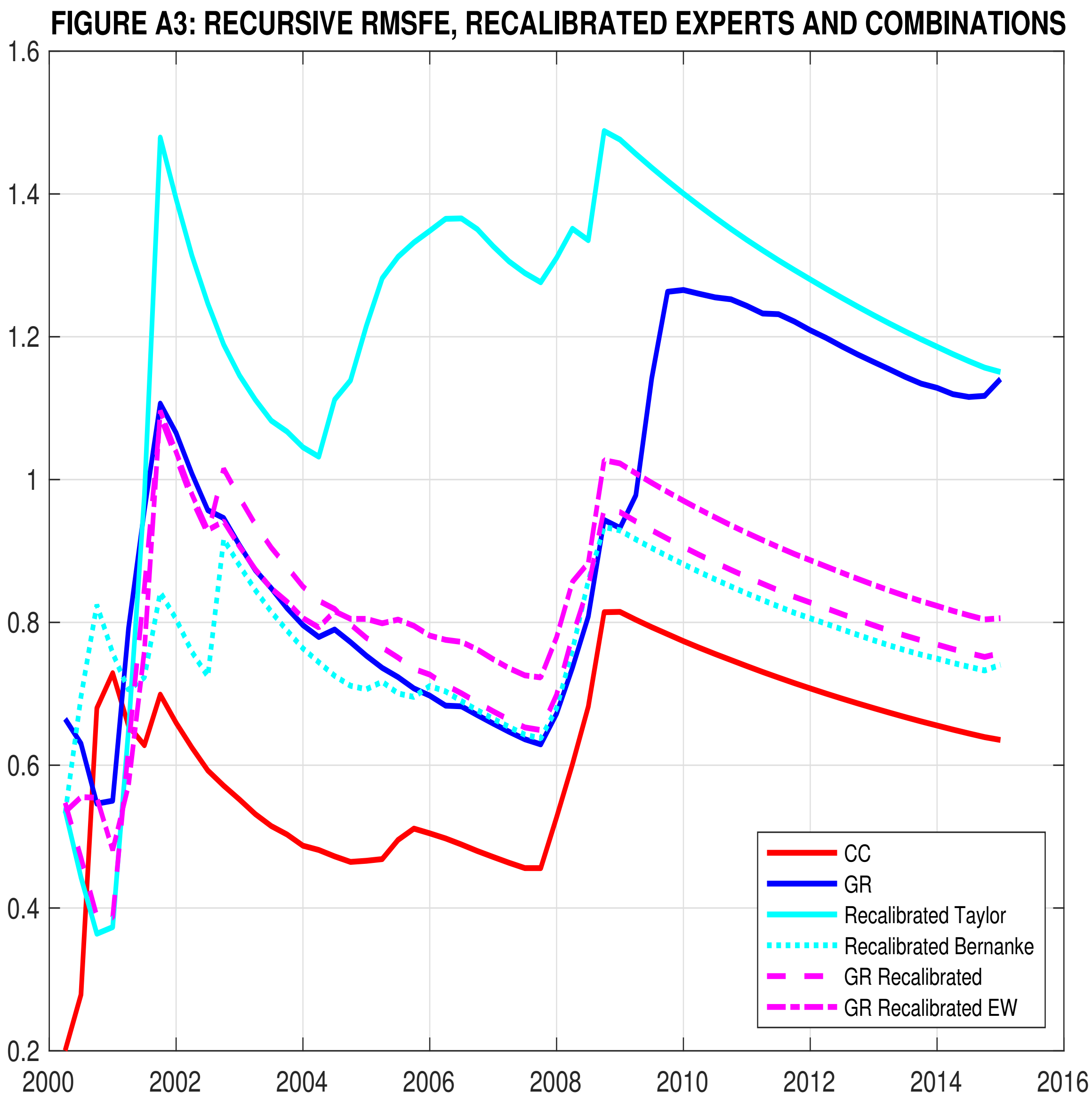
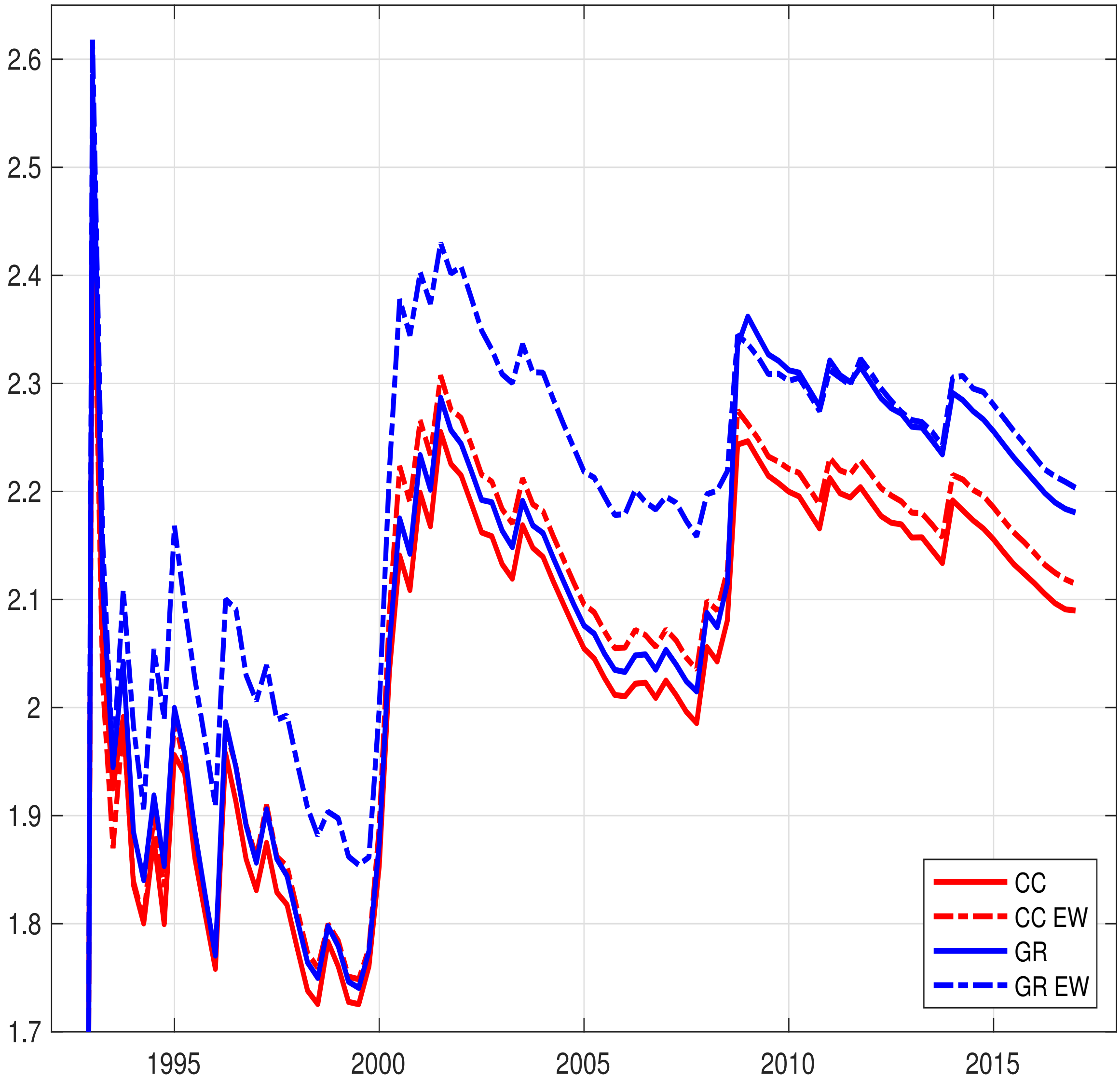


FIGURE A3: RECURSIVE RMSFE, RECALIBRATED EXPERTS AND COMBINATIONS
1.6
1.4
1.2
1
0.8
0.6
0.4
0.2
2000
2002
2004
2006
2008
2010
2012
2014
2016
CC
GR
Recalibrated Taylor
Recalibrated Bernanke
GR Recalibrated
GR Recalibrated EW

FIGURE A4: RECURSIVE RMSFE, COMBINATIONS



B. Robustness to Kernel Density Smoothing

In our two applications, we use non-parametric kernel-smoothed ECDFs for the marginal distributions, fitted recursively through the expanding window. Throughout the analysis reported in the main text of the paper, non-parametric kernel fitting uses the MATLAB function ksdensity, with a calibrated bandwidth. Higher bandwidths increase the smoothing. The calibrated parameters used in the forecast evaluations reported in the main text are: Federal Funds rate 0.15, Taylor expert 0.15, Bernanke expert 0.15, output growth 0.4, Expert CS 0.4, and Expert AR 0.4.

In figures B1 through B20, we report results based on varying the bandwidth parameter, kw, through the range 0.1 to 0.5 in each application. The general pattern of results is that the qualitative nature of the results is fairly robust to the choice of bandwidth parameter. Figures B1 through B10 are for the predicting Federal Funds rate application. Figures B11 through B20 are the corresponding plots for the forecasting output growth application.

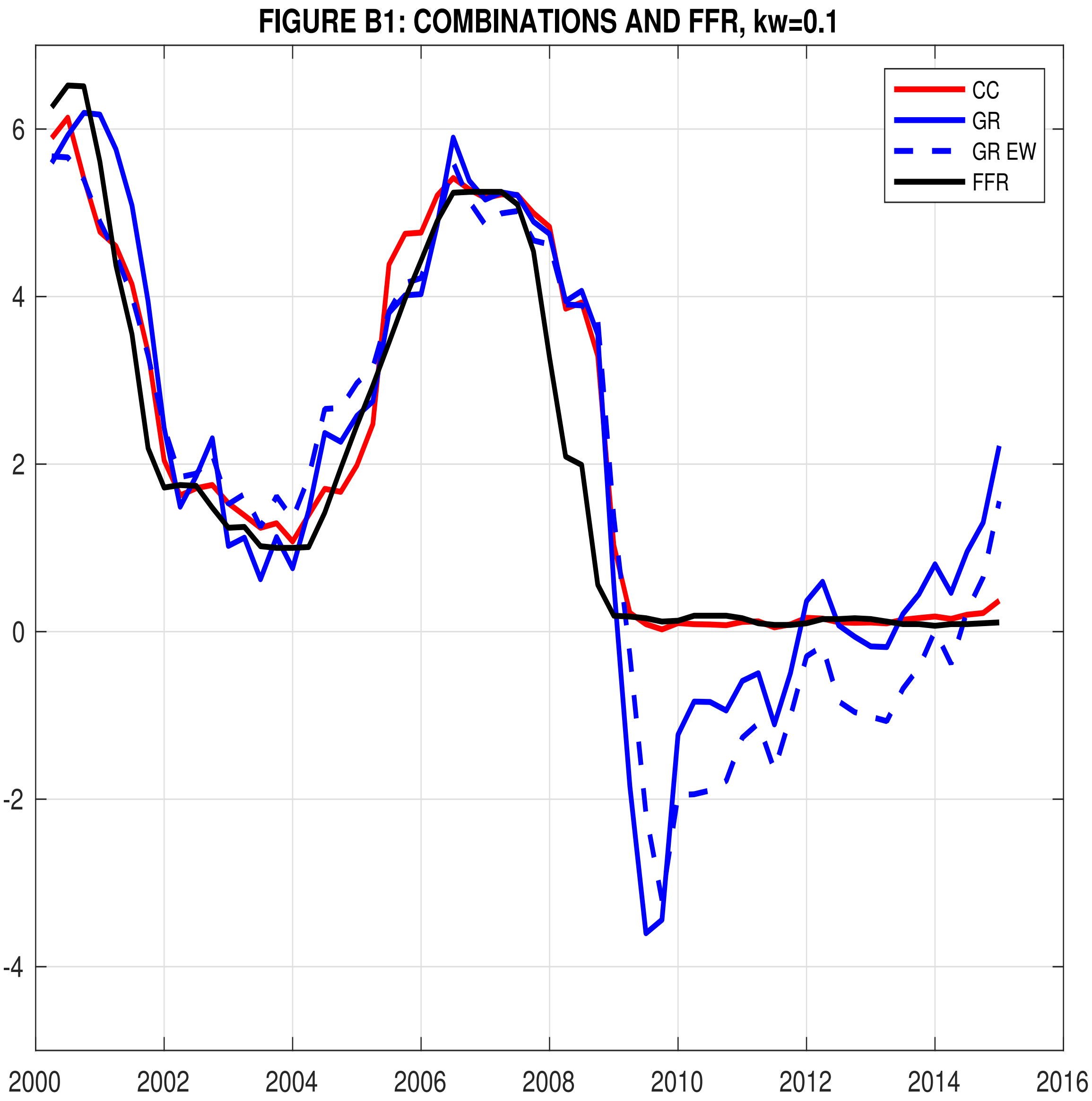


FIGURE B1: COMBINATIONS AND FFR, kw=0.1
CC
GR
GR EW
FFR
6
4
2
0
-2
-4
2000
2002
2004
2006
2008
2010
2012
2014
2016

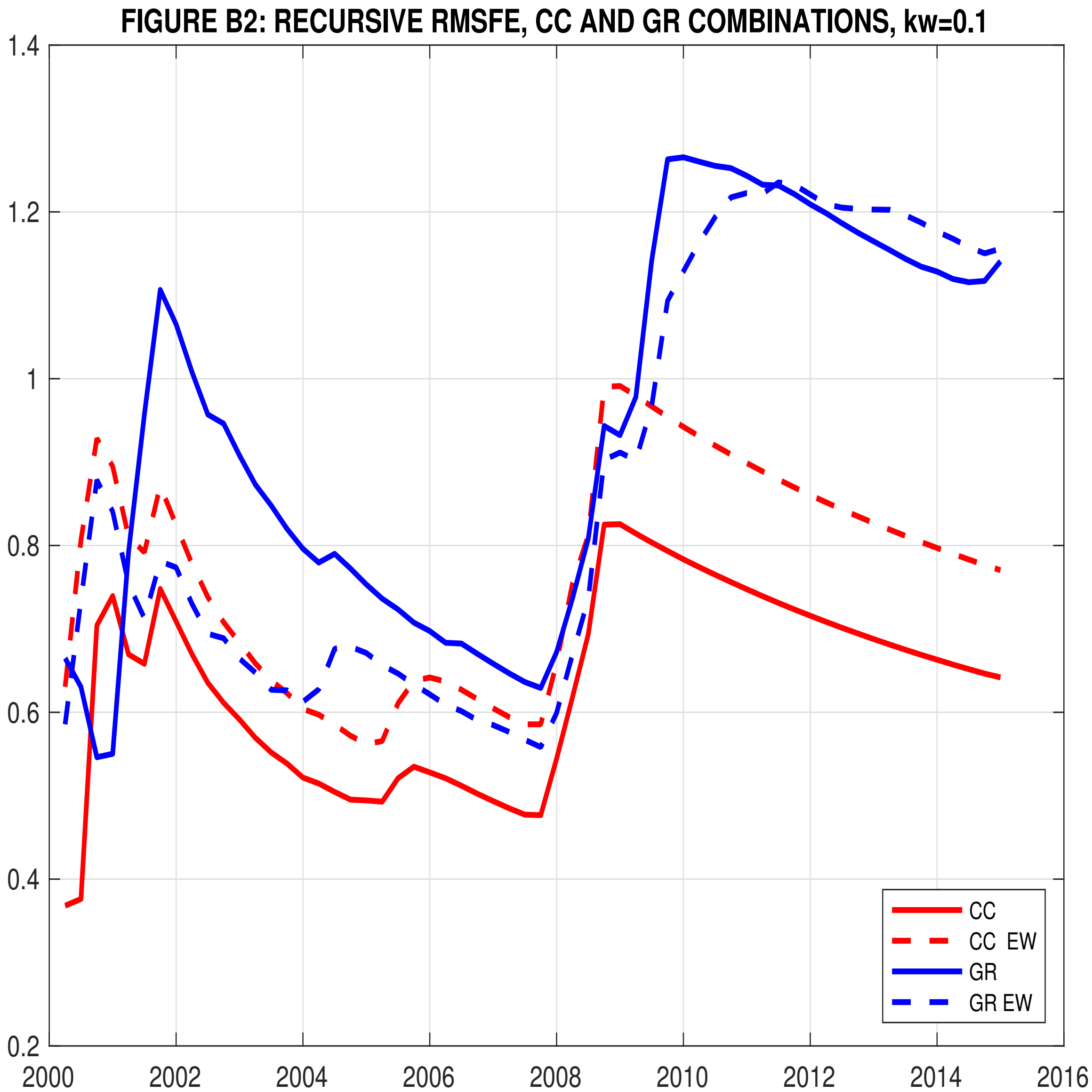


FIGURE B2: RECURSIVE RMSFE, CC AND GR COMBINATIONS, kw=0.1
1.4
1.2
1
0.8
0.6
0.4
0.2
2000
2002
2004
2006
2008
2010
2012
2014
2016
CC
CC EW
GR
GR EW

FIGURE B3: COMBINATIONS AND FFR, kw =0.2

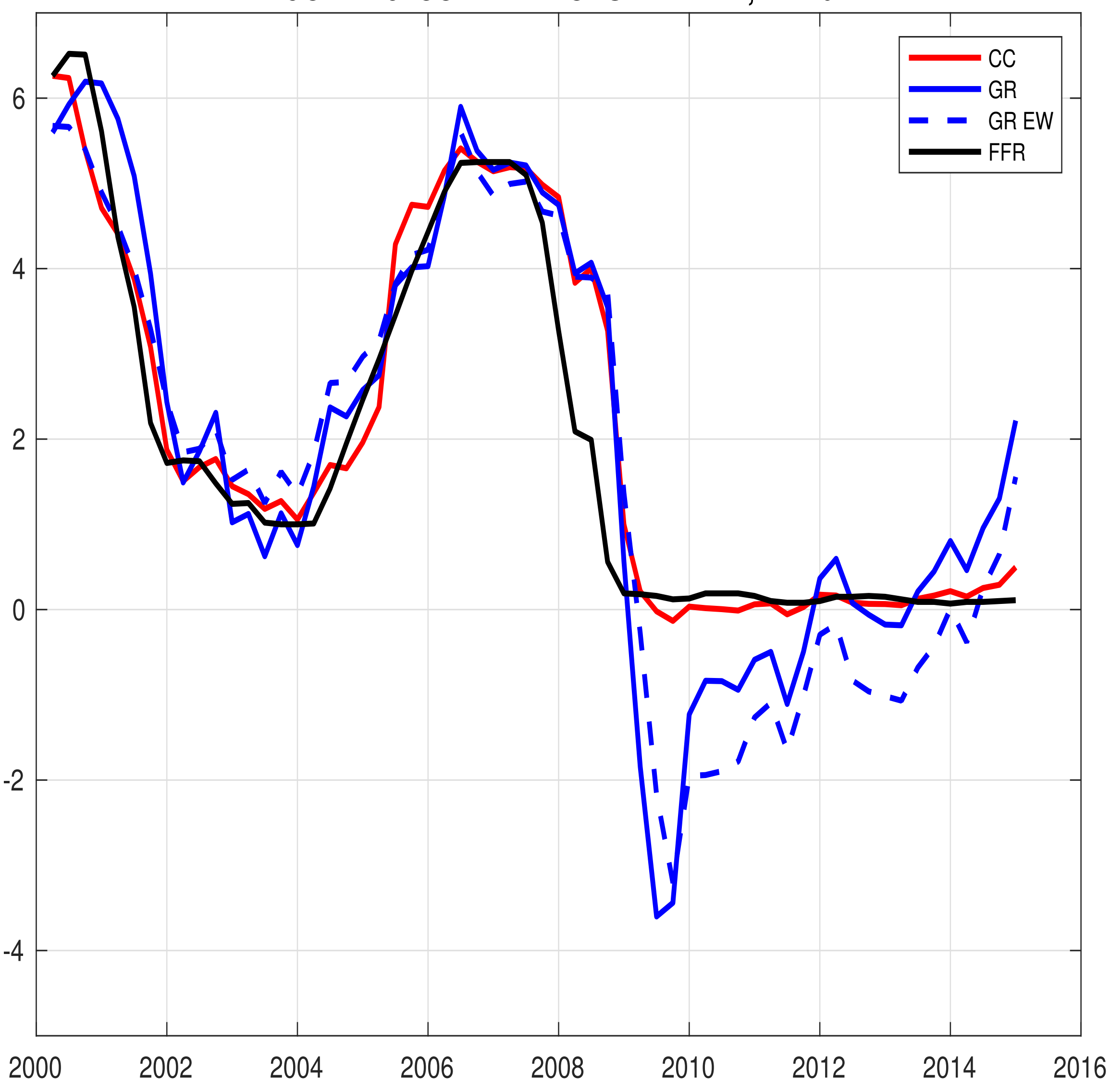
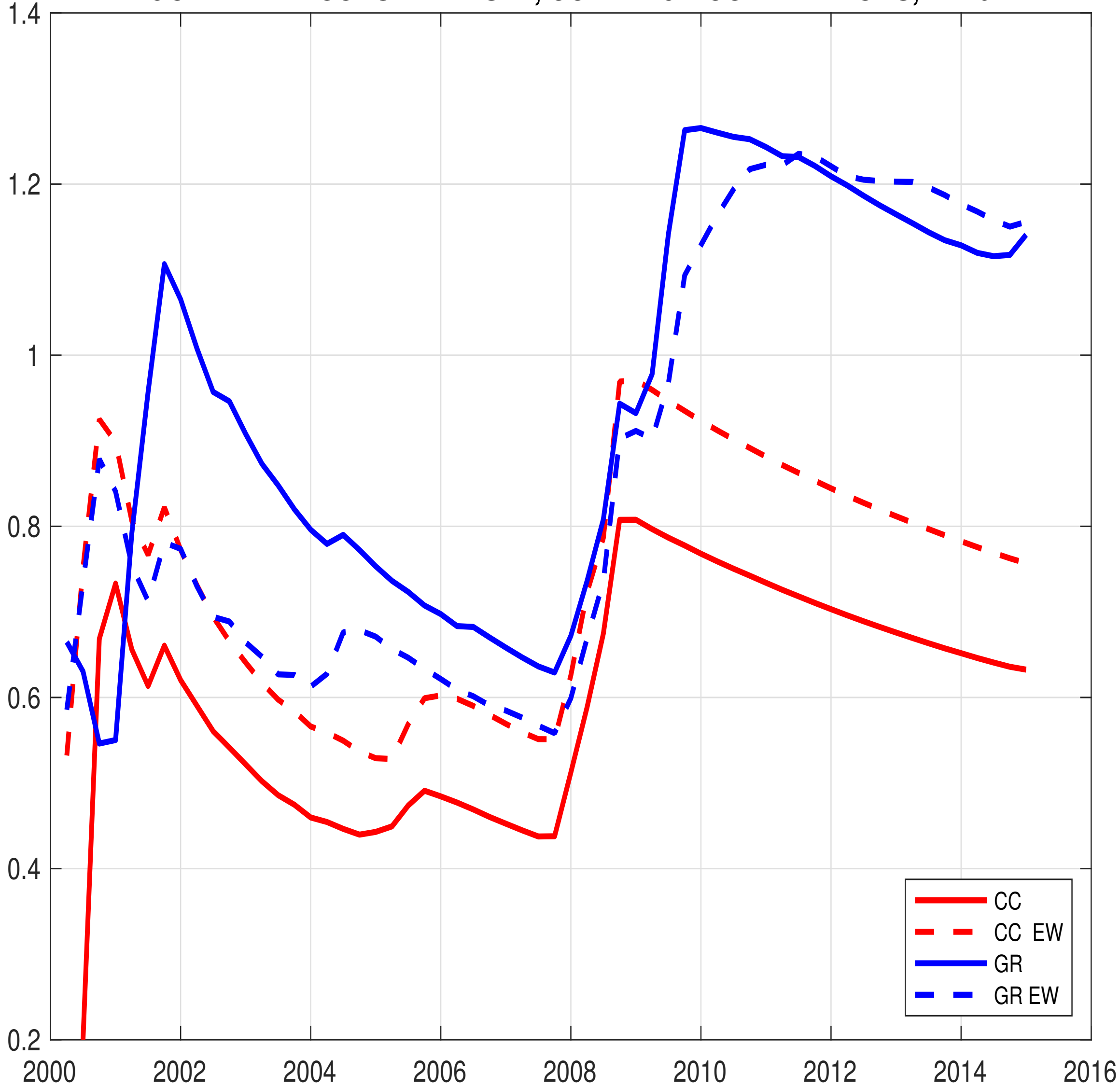


FIGURE B4: RECURSIVE RMSFE, CC AND GR COMBINATIONS, kw=0.2



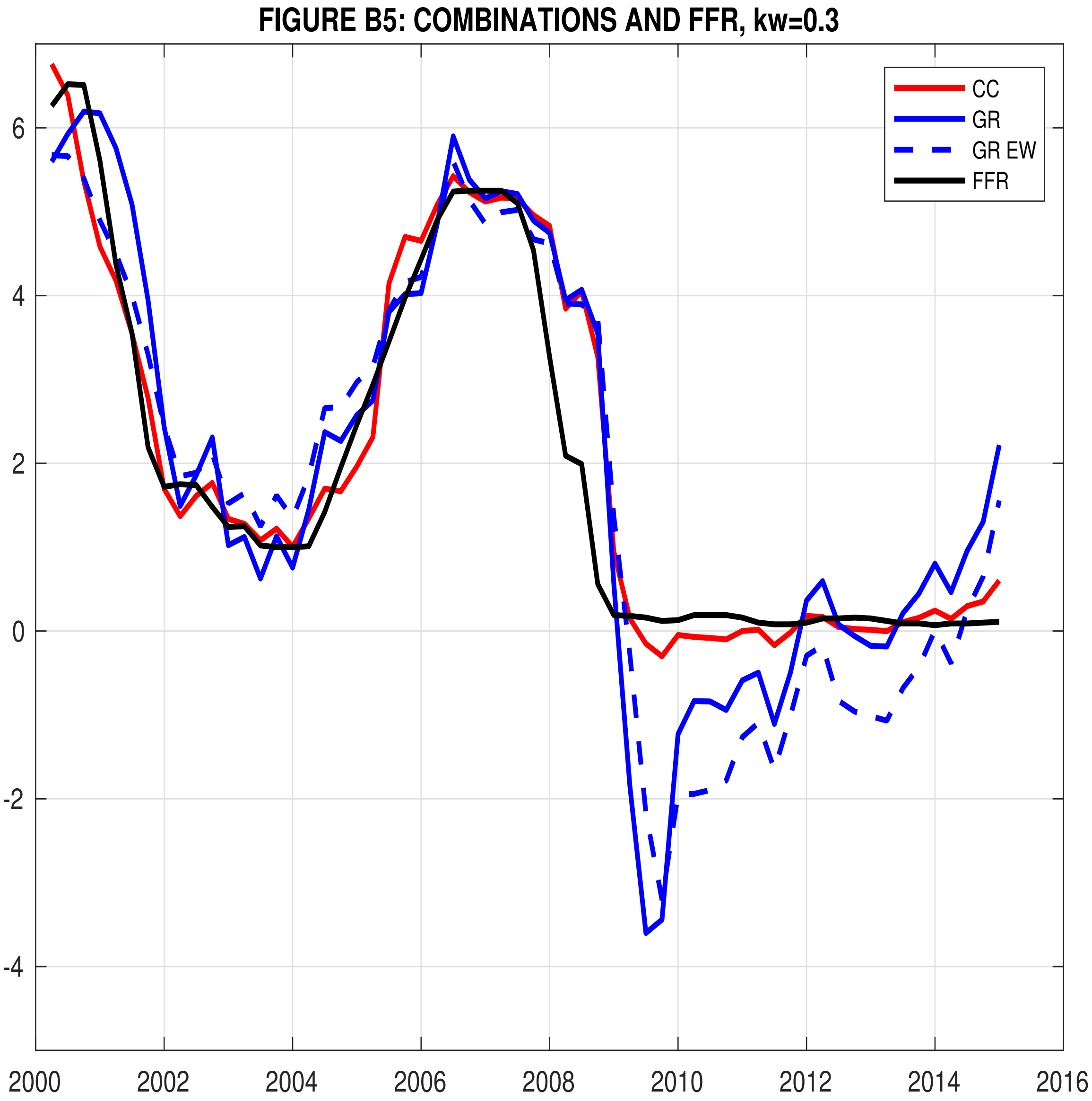


FIGURE B5: COMBINATIONS AND FFR, kw=0.3
CC
GR
GR EW
FFR
6
4
2
0
-2
-4
2000
2002
2004
2006
2008
2010
2012
2014
2016

FIGURE B6: RECURSIVE RMSFE, CC AND GR COMBINATIONS, kw=0.3

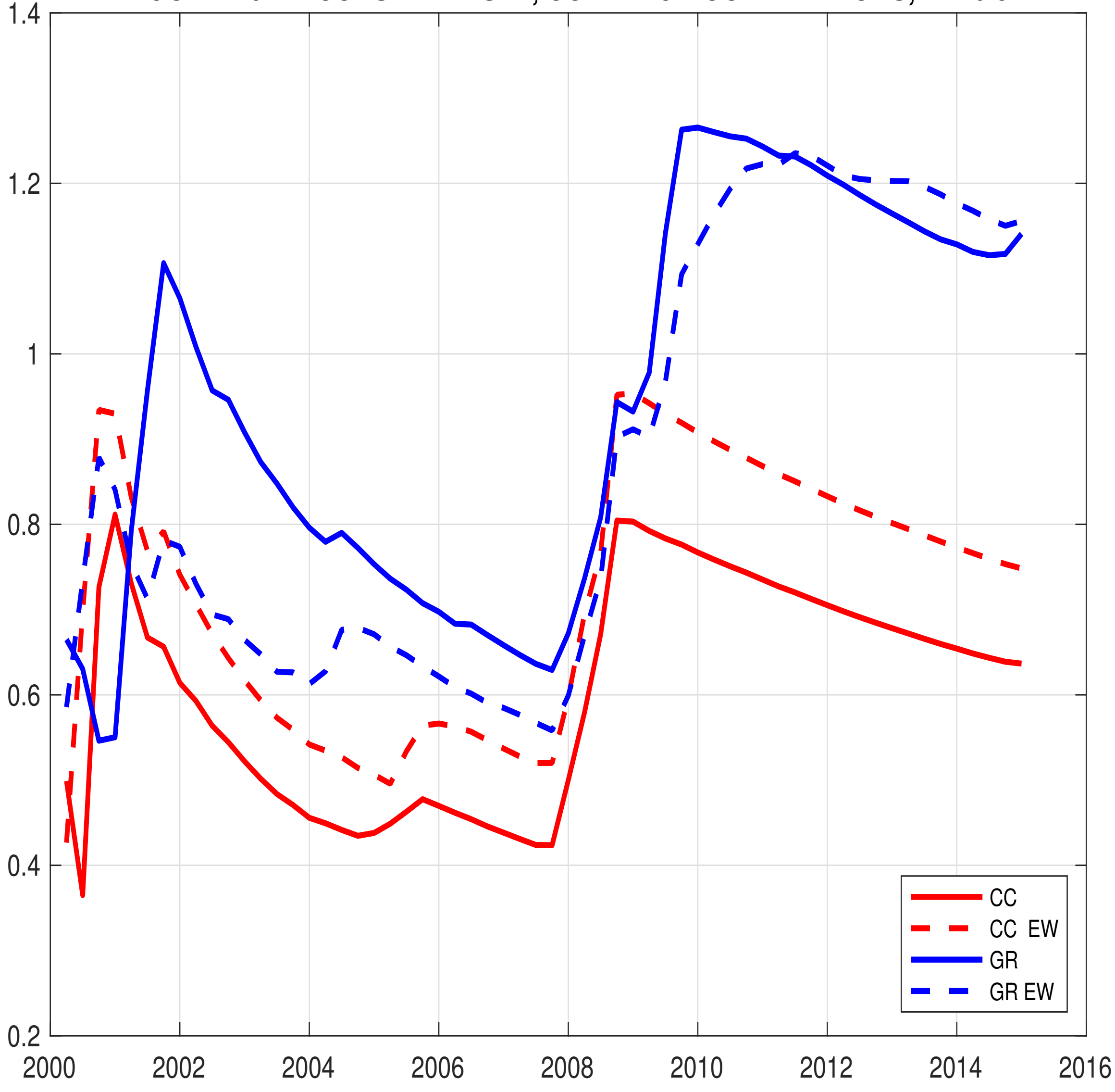
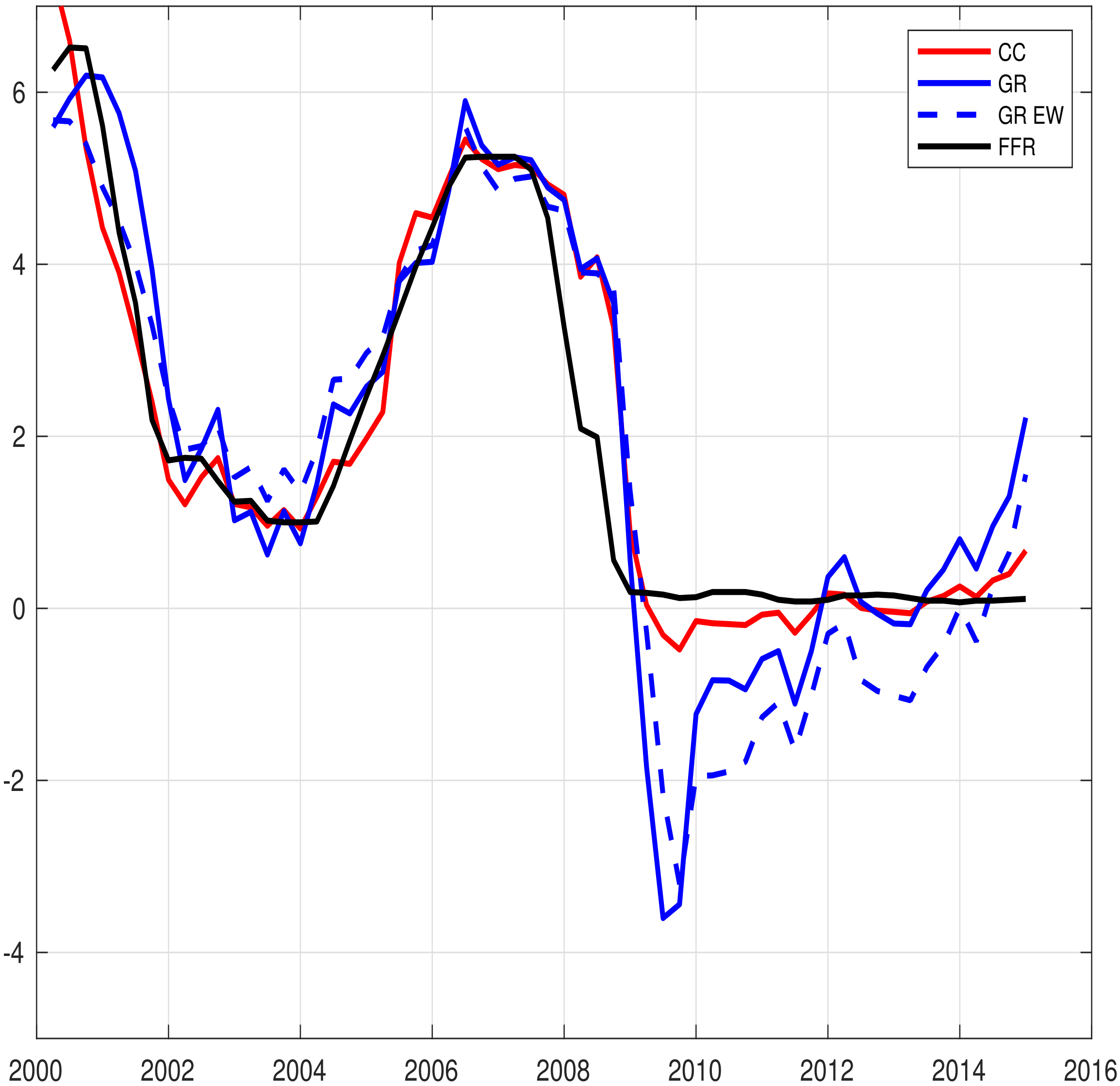


FIGURE B7: COMBINATIONS AND FFR, kw =0.4



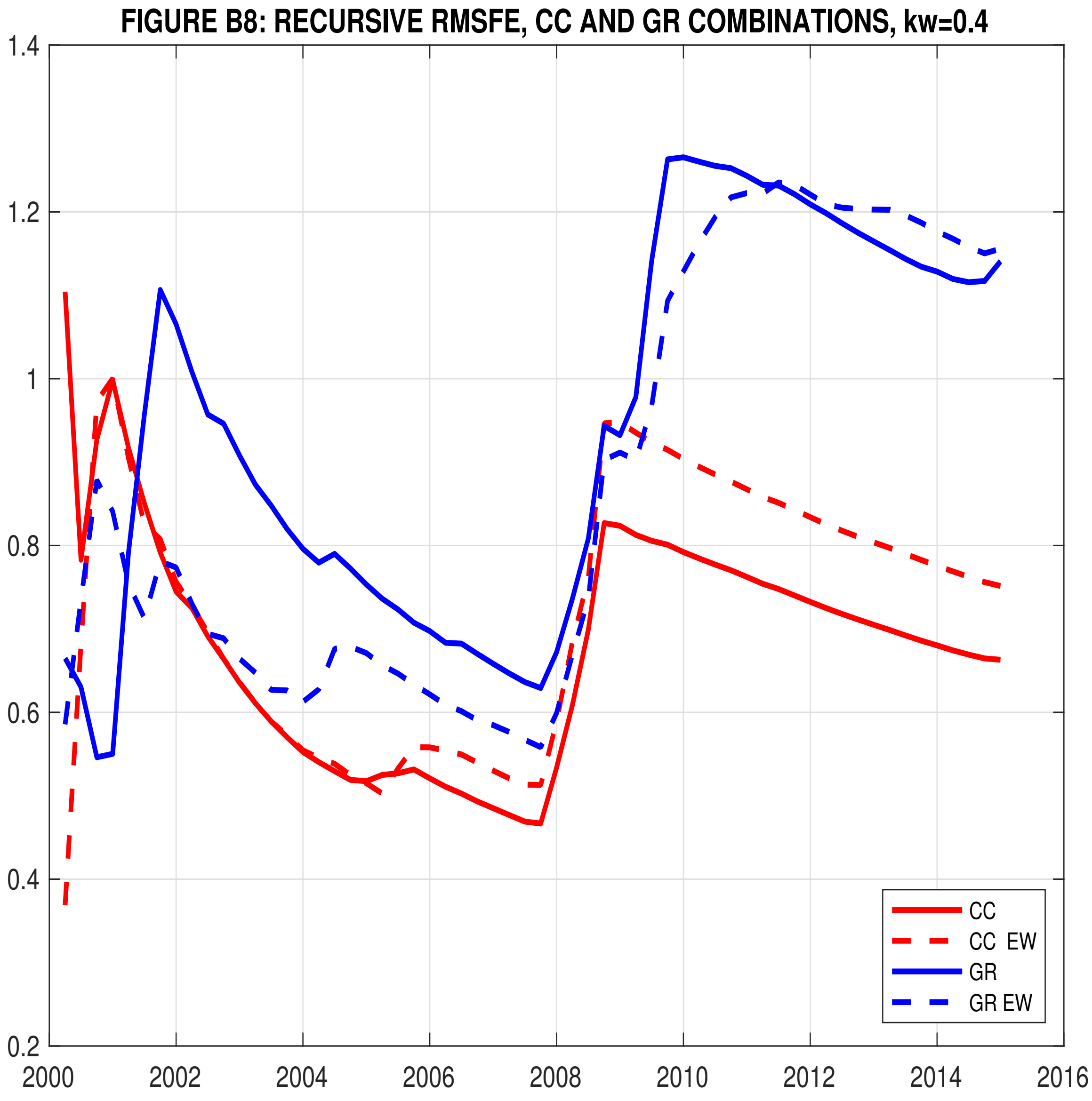
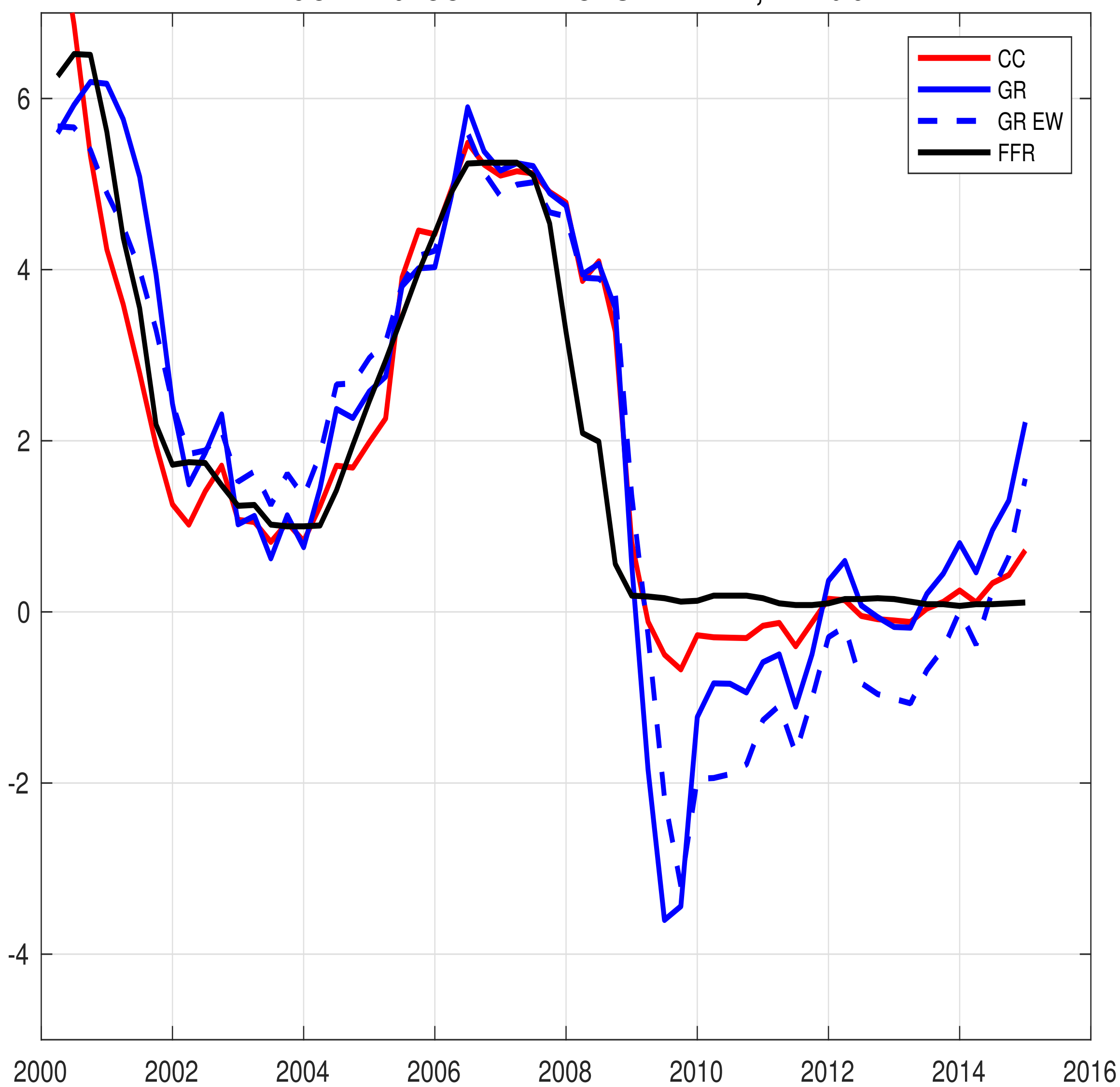


FIGURE B8: RECURSIVE RMSFE, CC AND GR COMBINATIONS, kw=0.4
1.4
1.2
1
0.8
0.6
0.4
0.2
2000
2002
2004
2006
2008
2010
2012
2014
2016
CC
CC EW
GR
GR EW

FIGURE B9: COMBINATIONS AND FFR, kw =0.5



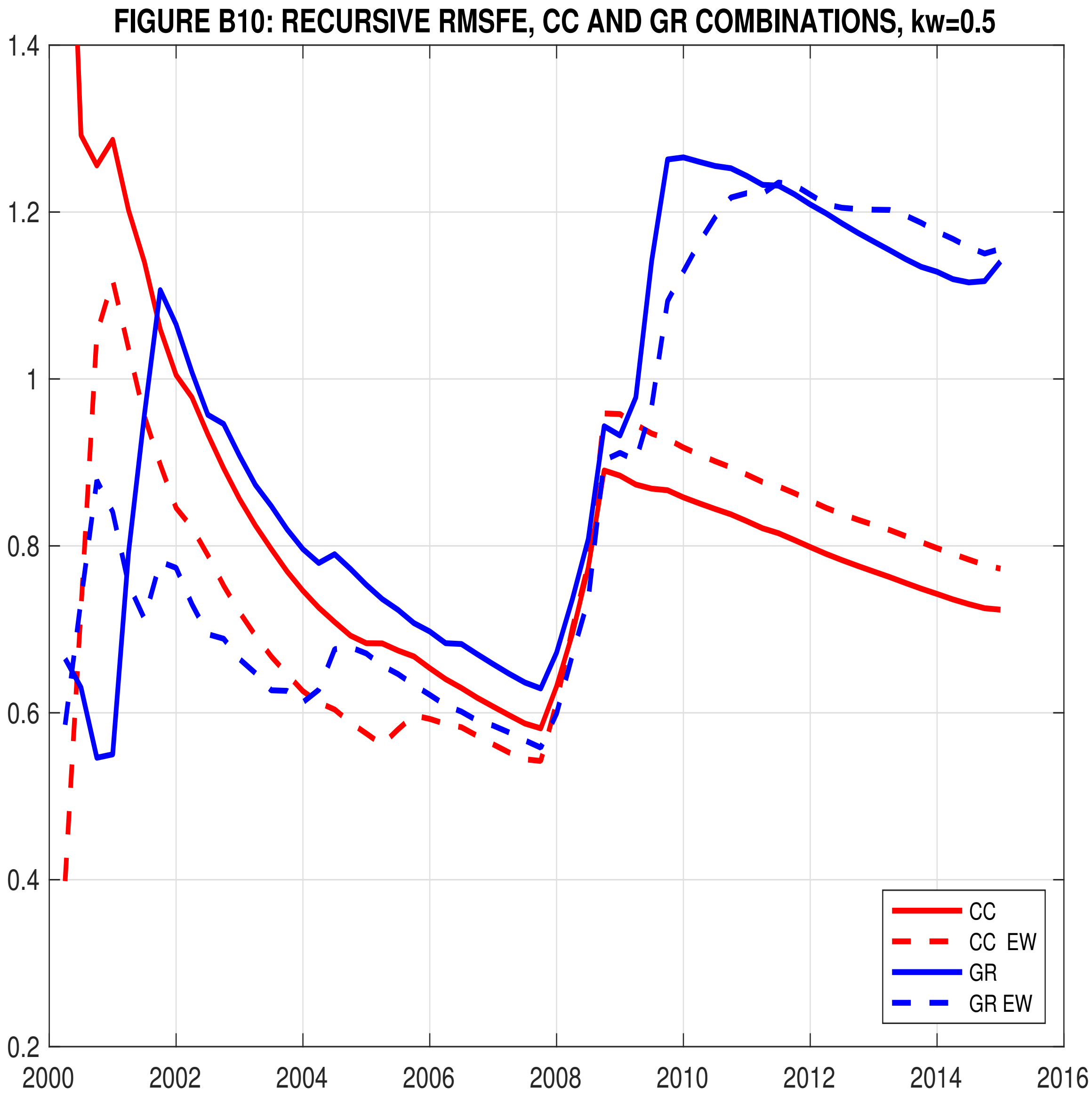


FIGURE B10: RECURSIVE RMSFE, CC AND GR COMBINATIONS, kw=0.5
1.4
1.2
1
0.8
0.6
0.4
0.2
2000
2002
2004
2006
2008
2010
2012
2014
2016
CC
CC EW
GR
GR EW

FIGURE B11: COMBINATIONS AND OUTTURNS, kw=0.1

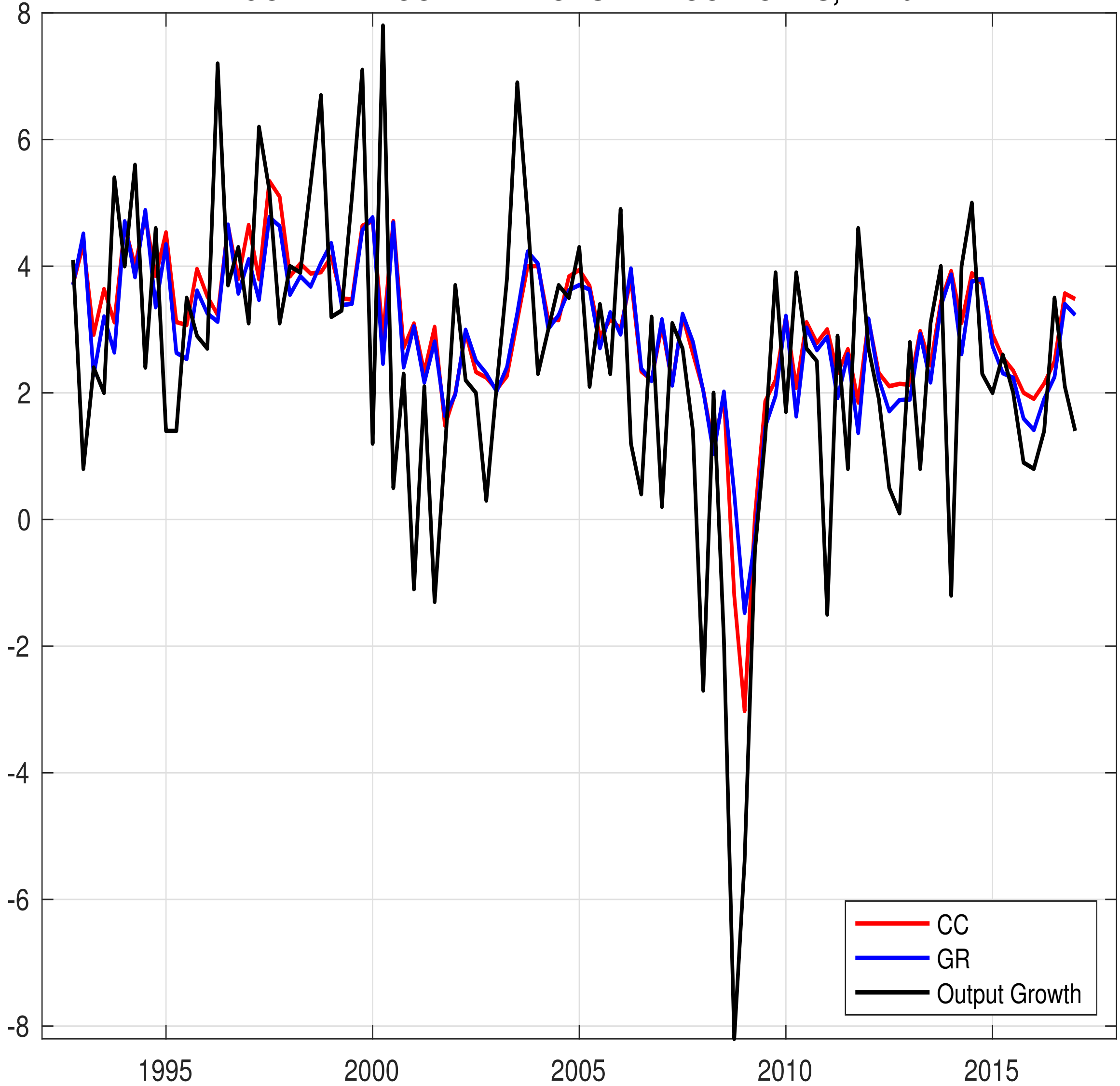


FIGURE B12: RECURSIVE RMSFE, COMBINATIONS, kw=0.1

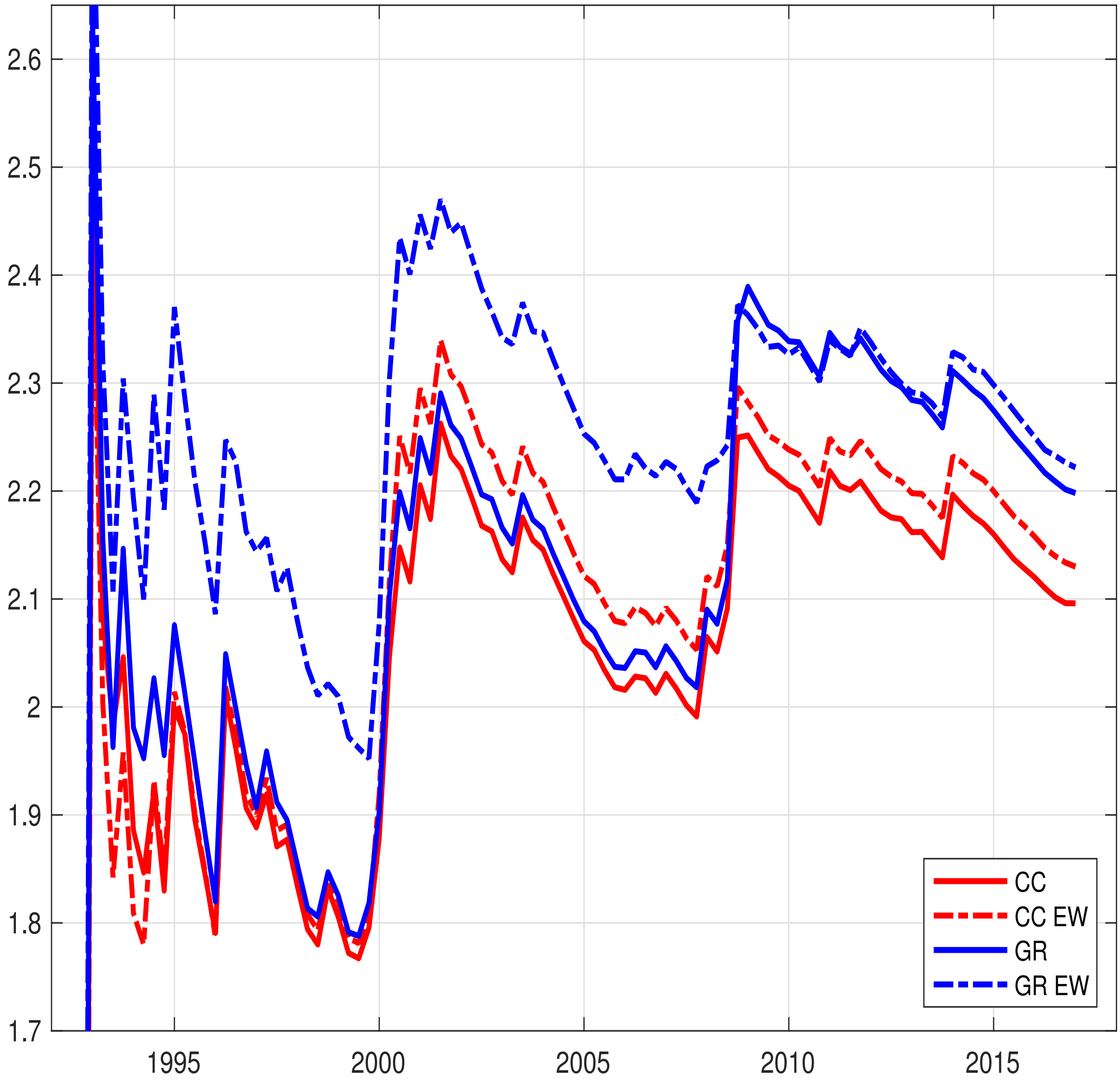


FIGURE B13: COMBINATIONS AND OUTTURNS, kw=0.2

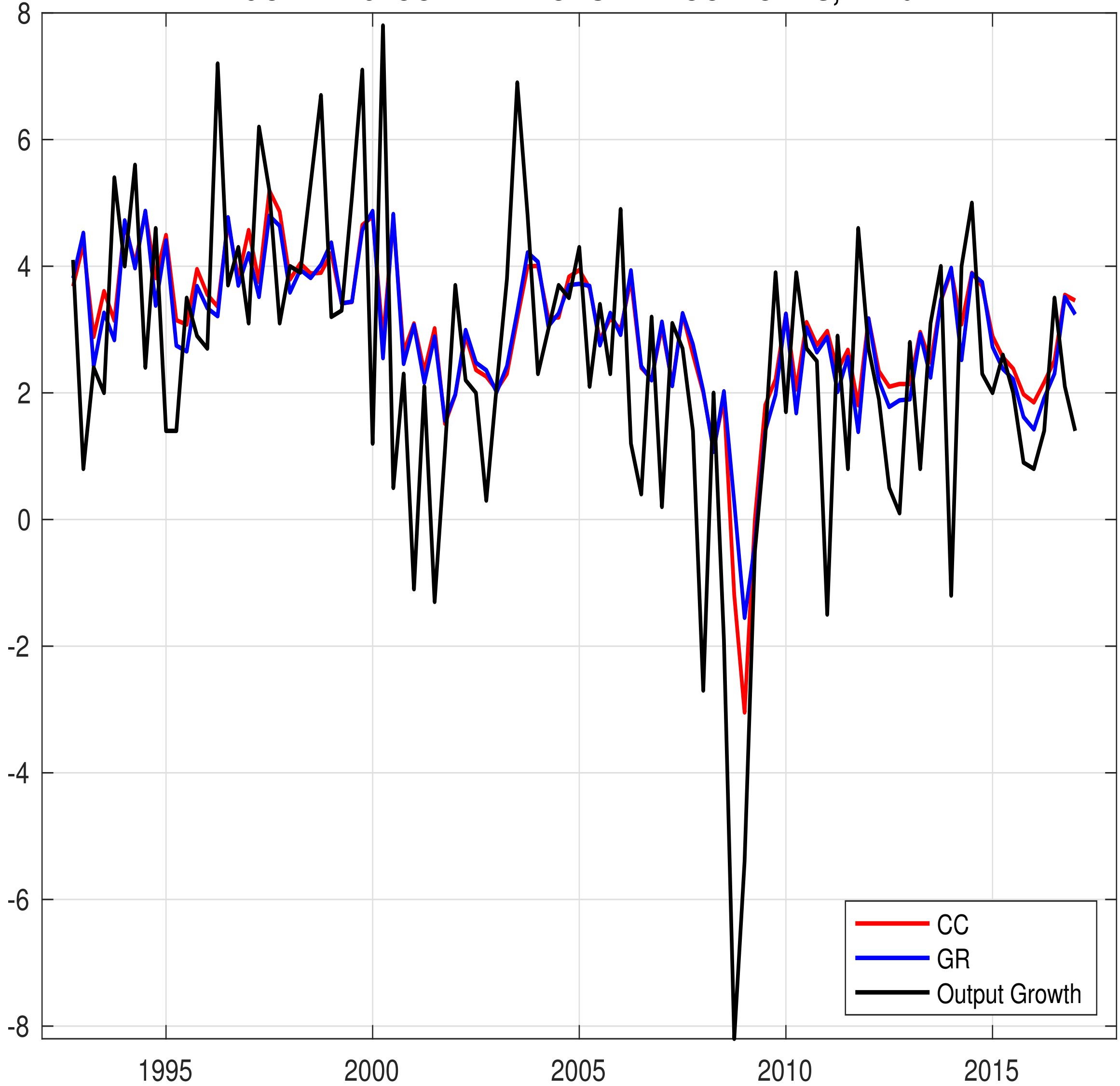
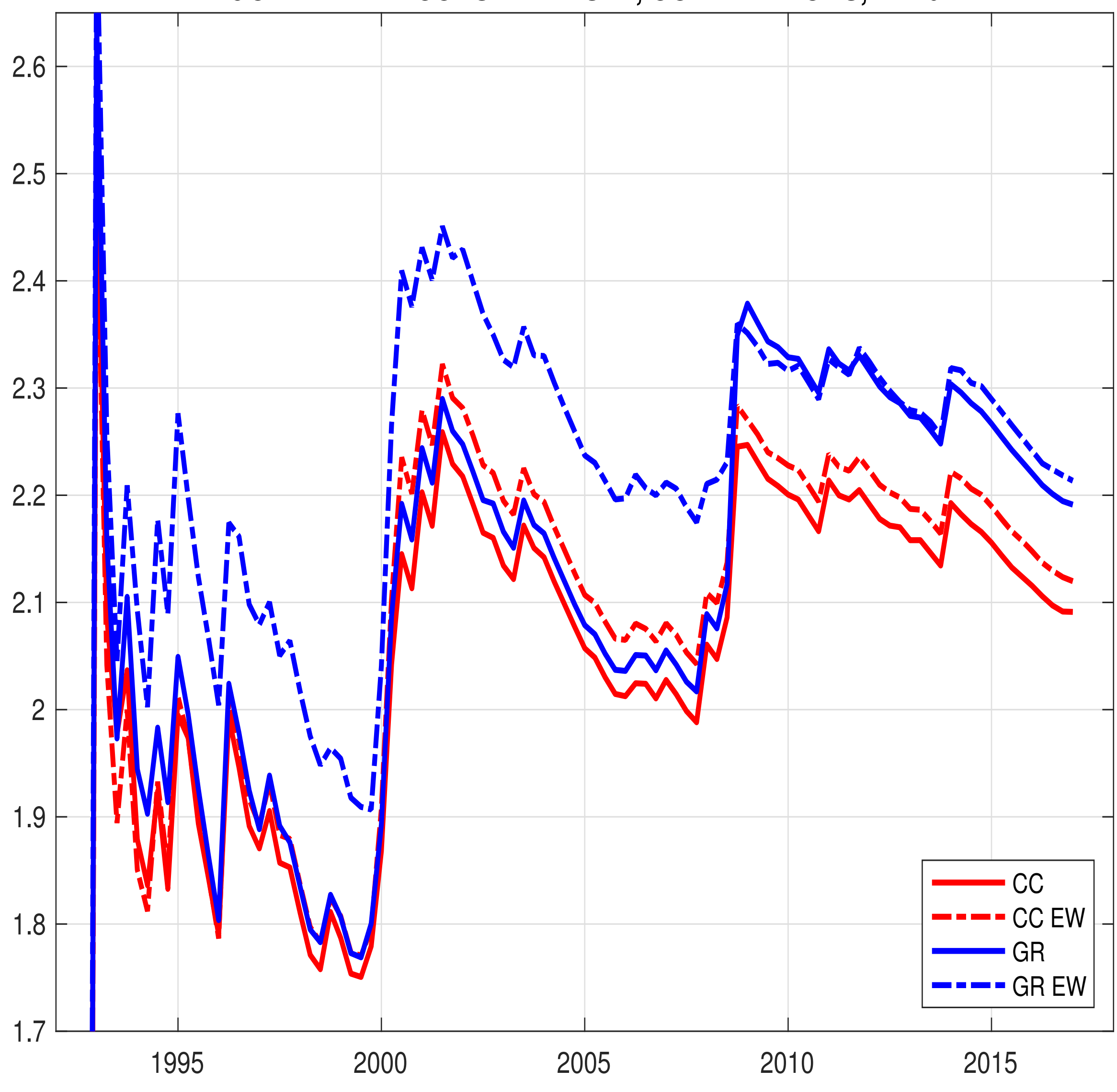


FIGURE B14: RECURSIVE RMSFE, COMBINATIONS, kw=0.2



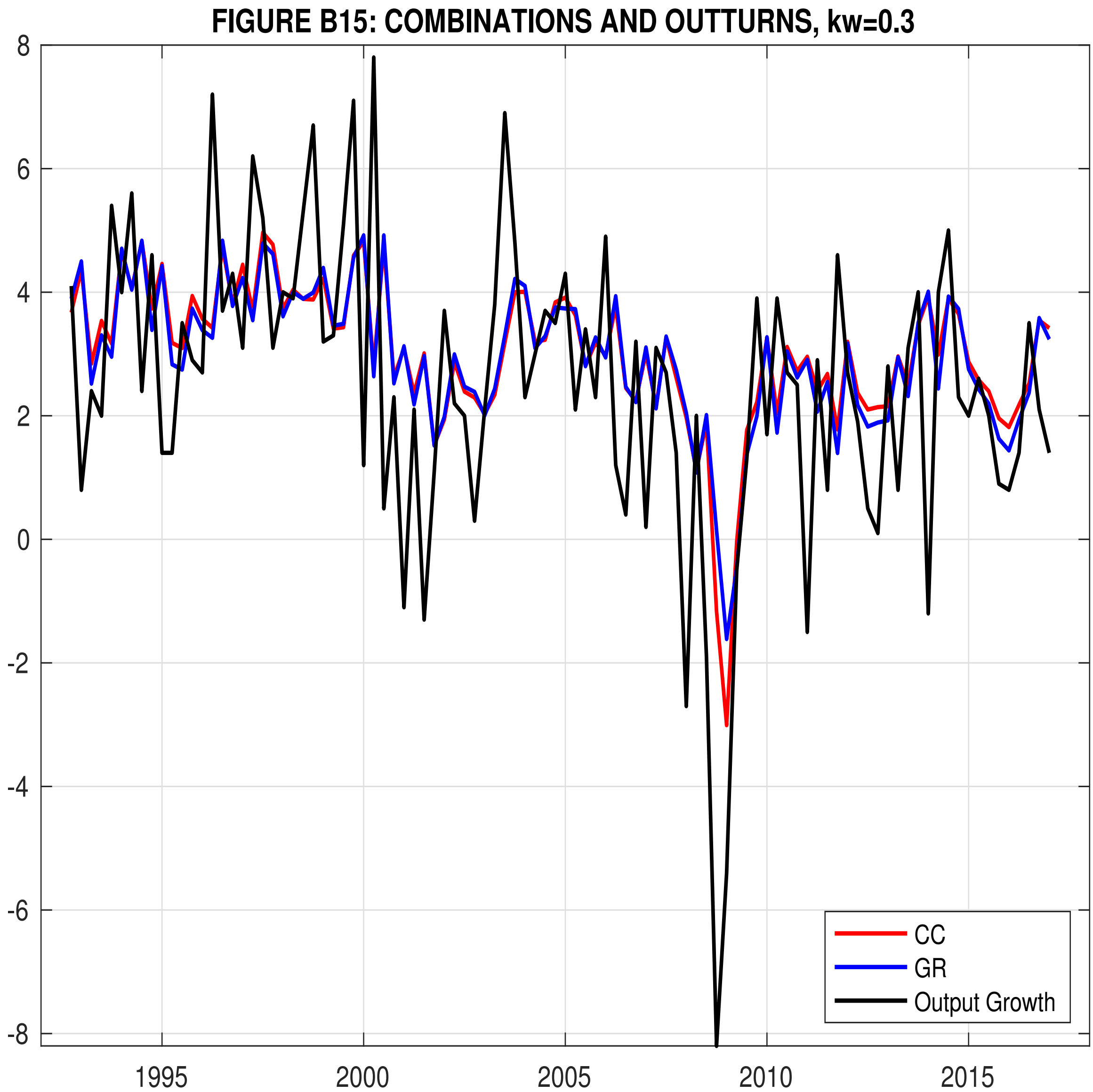


FIGURE B15: COMBINATIONS AND OUTTURNS, kw=0.3
8
6
4
2
0
-2
-4
-6
-8
1995
2000
2005
2010
2015
CC
GR
Output Growth

FIGURE B16: RECURSIVE RMSFE, COMBINATIONS, kw=0.3

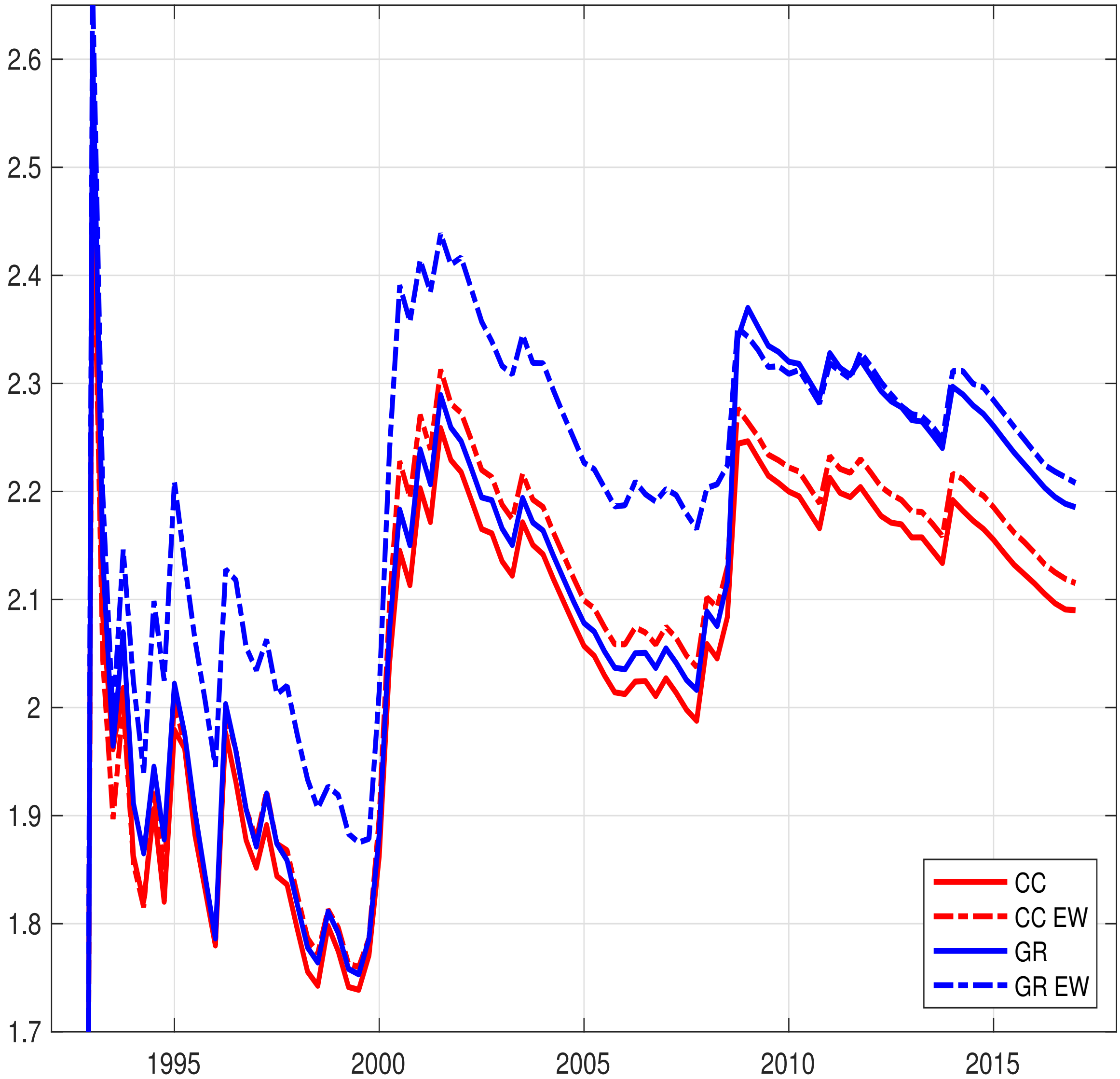


FIGURE B17: COMBINATIONS AND OUTTURNS, kw=0.4

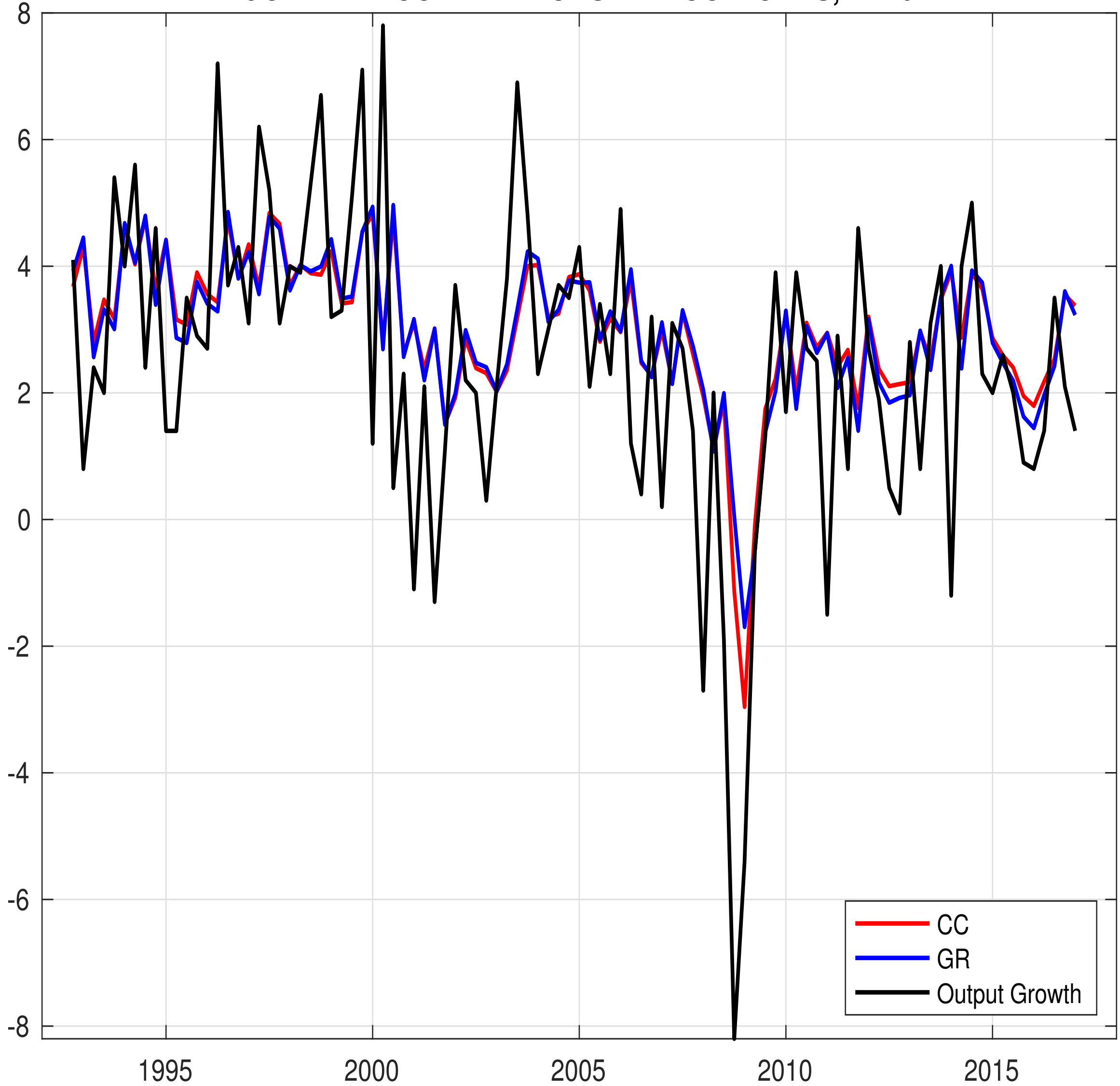


FIGURE B18: RECURSIVE RMSFE, COMBINATIONS, kw=0.4

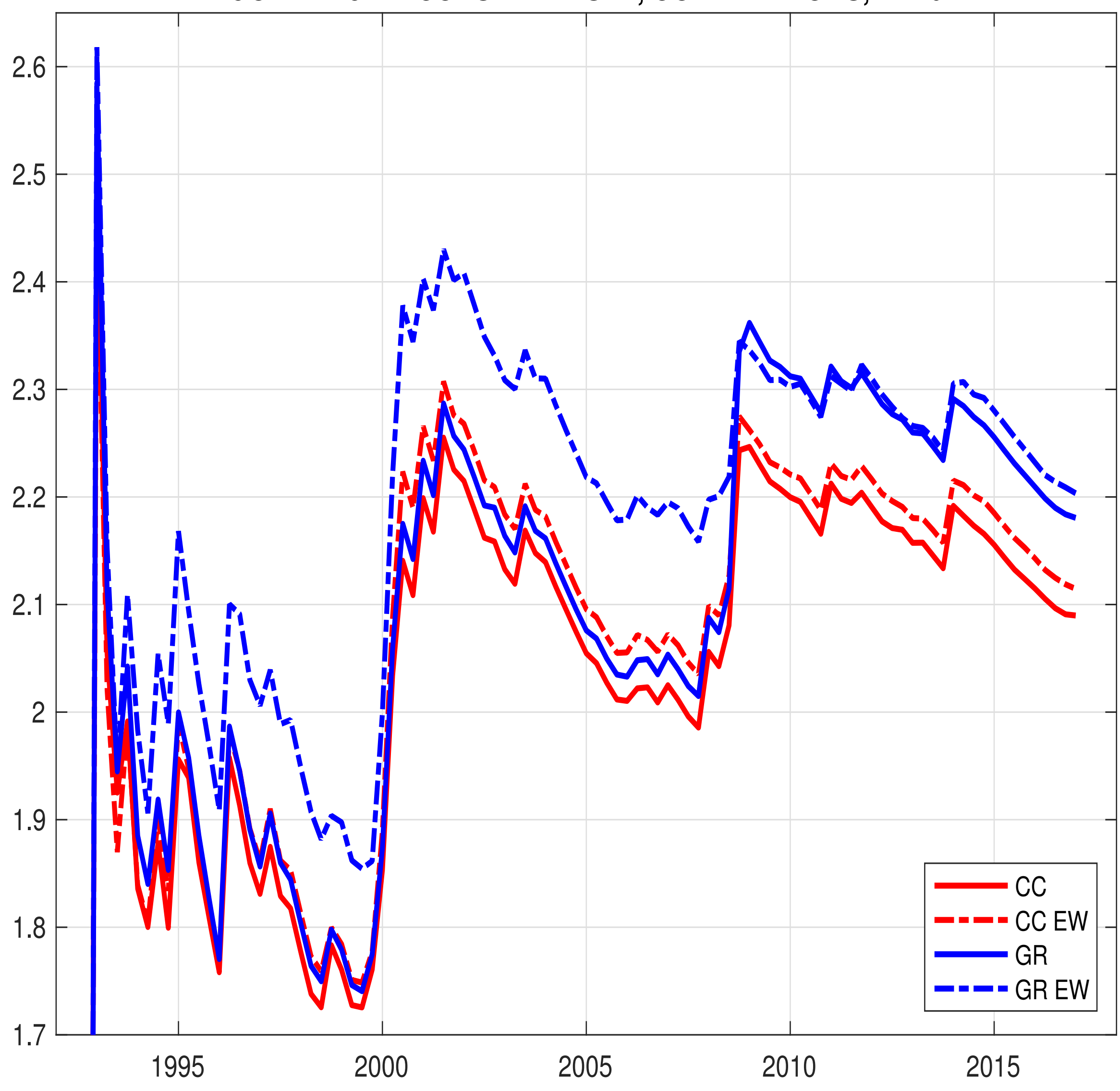


FIGURE B19: COMBINATIONS AND OUTTURNS, kw=0.5

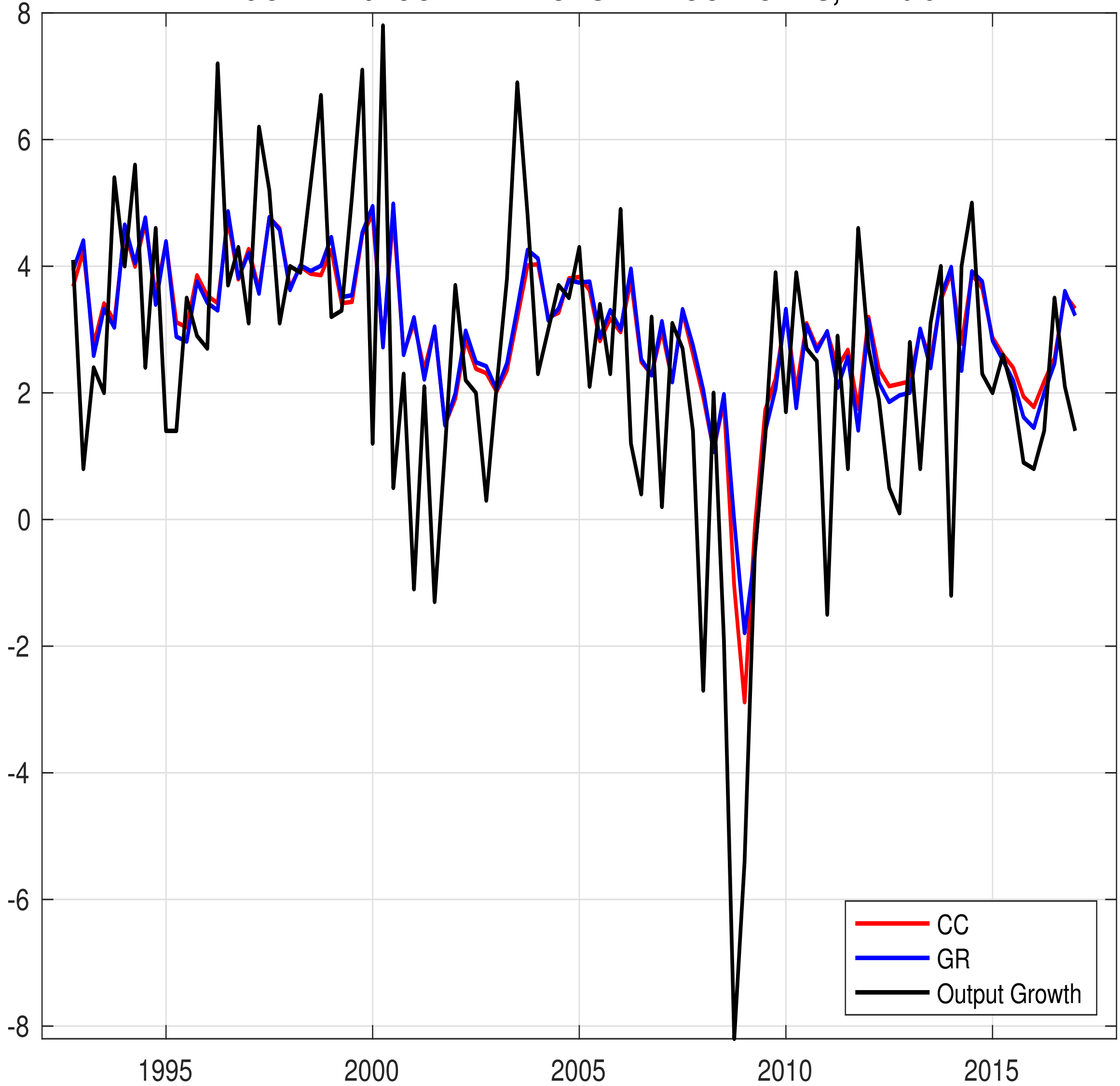
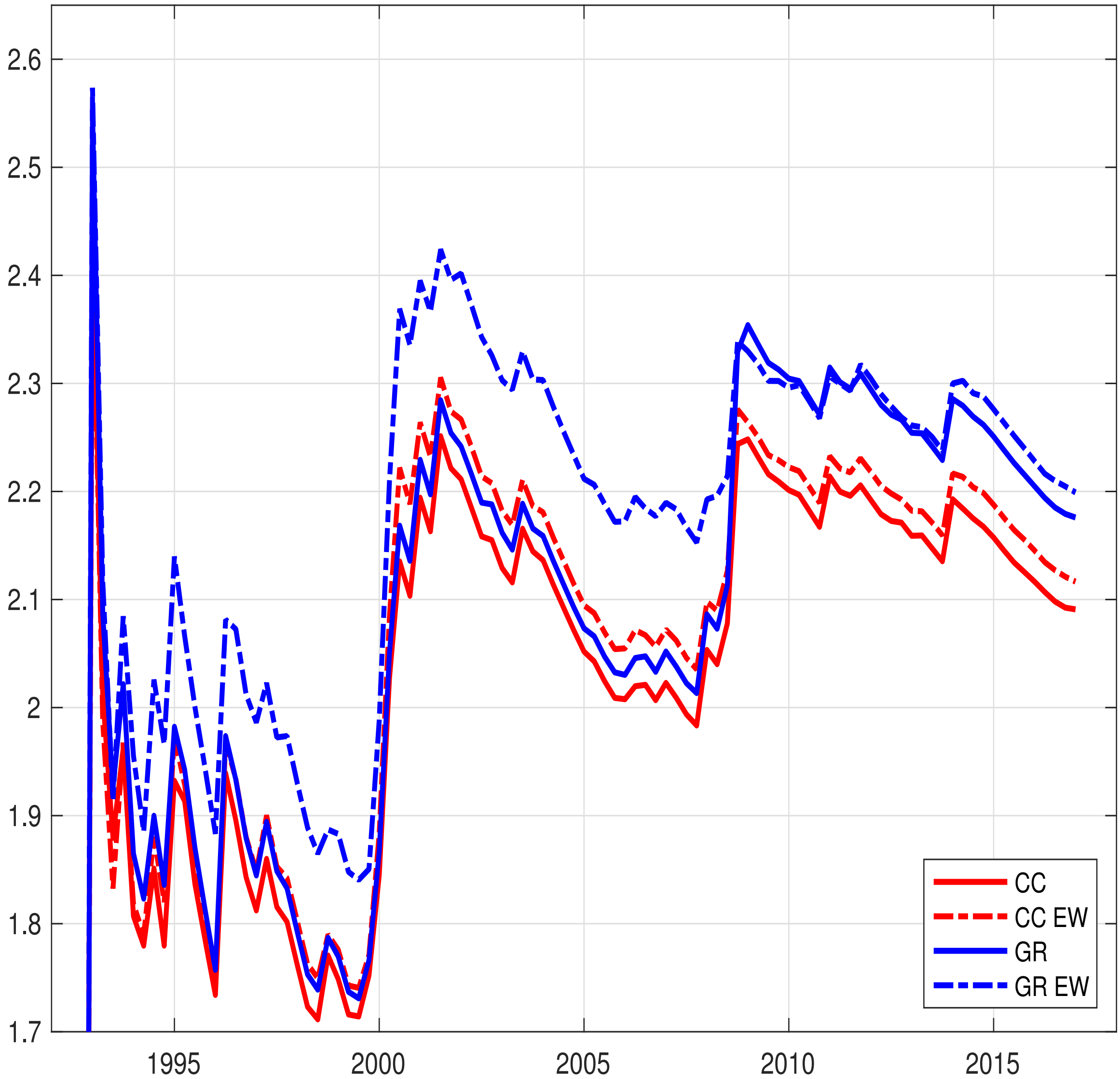


FIGURE B20: RECURSIVE RMSFE, COMBINATIONS, kw=0.5



C. Homogeneous and Heterogeneous Marginals

In the results reported for both applications in the main text the copula point combination uses heterogeneous marginals. We report results for the homogeneous case in figures C1 through C4. The homogeneous marginals results are slightly weaker than the heterogeneous case in terms of RMSFE by the end of the evaluation but slightly better for some samples. Figures C1 and C2 are for the predicting Federal Funds rate application. Figures C3 and C4 are the corresponding plots for the forecasting output growth application.

FIGURE C1: COMBINATIONS AND FFR, homogeneous marginals

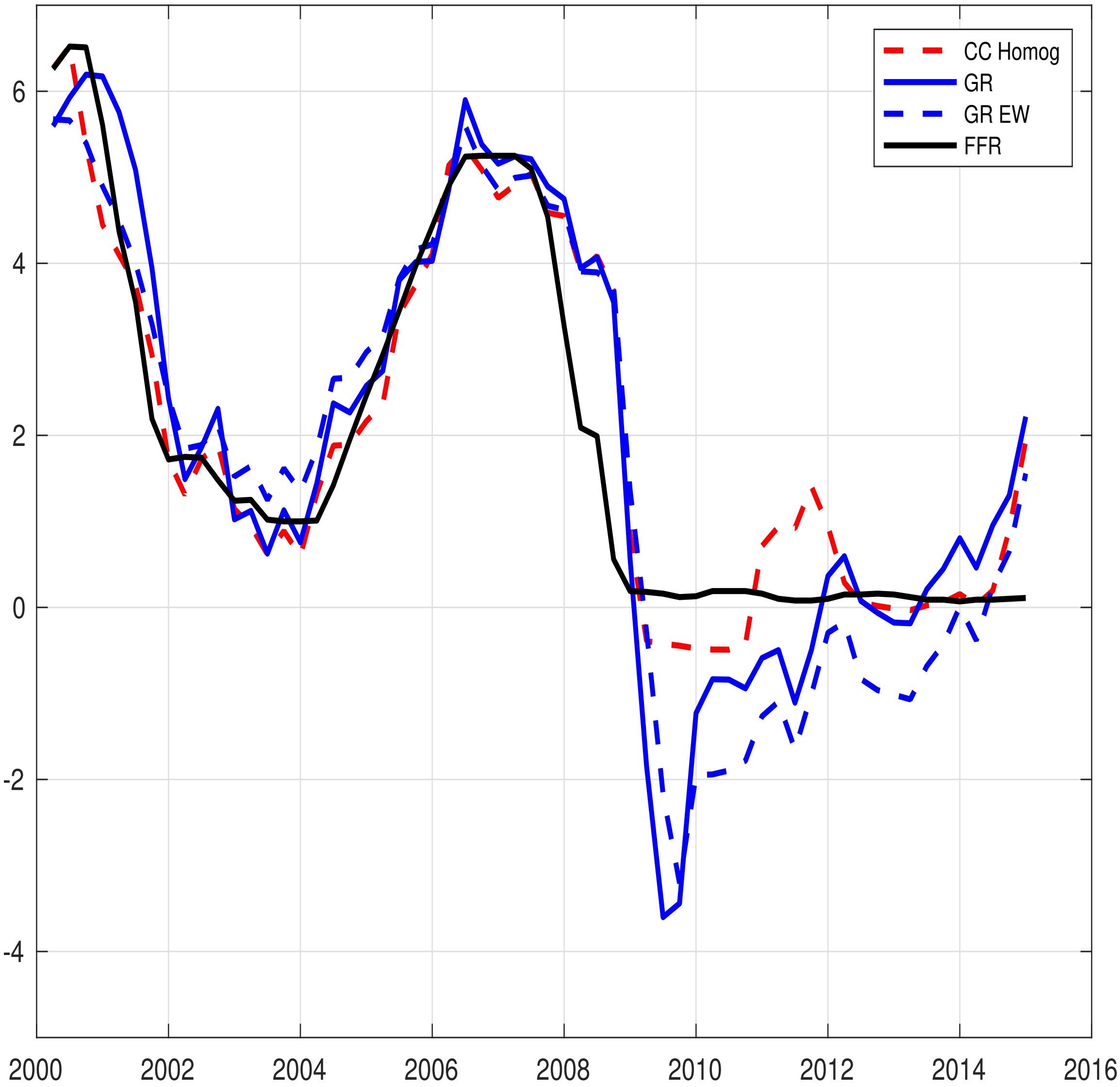
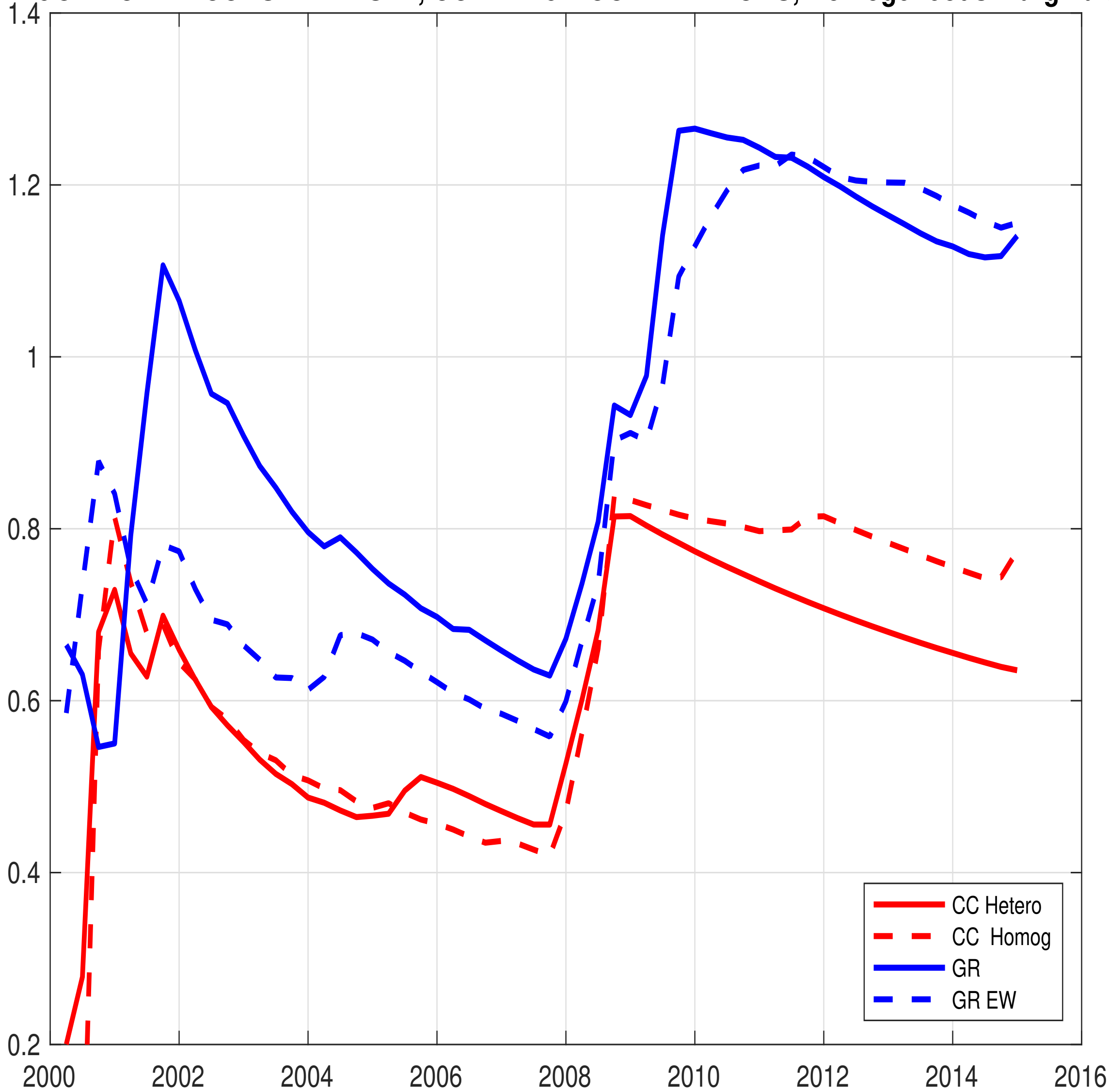


FIGURE C2: RECURSIVE RMSFE, CC AND GR COMBINATIONS, Homogeneous marginals



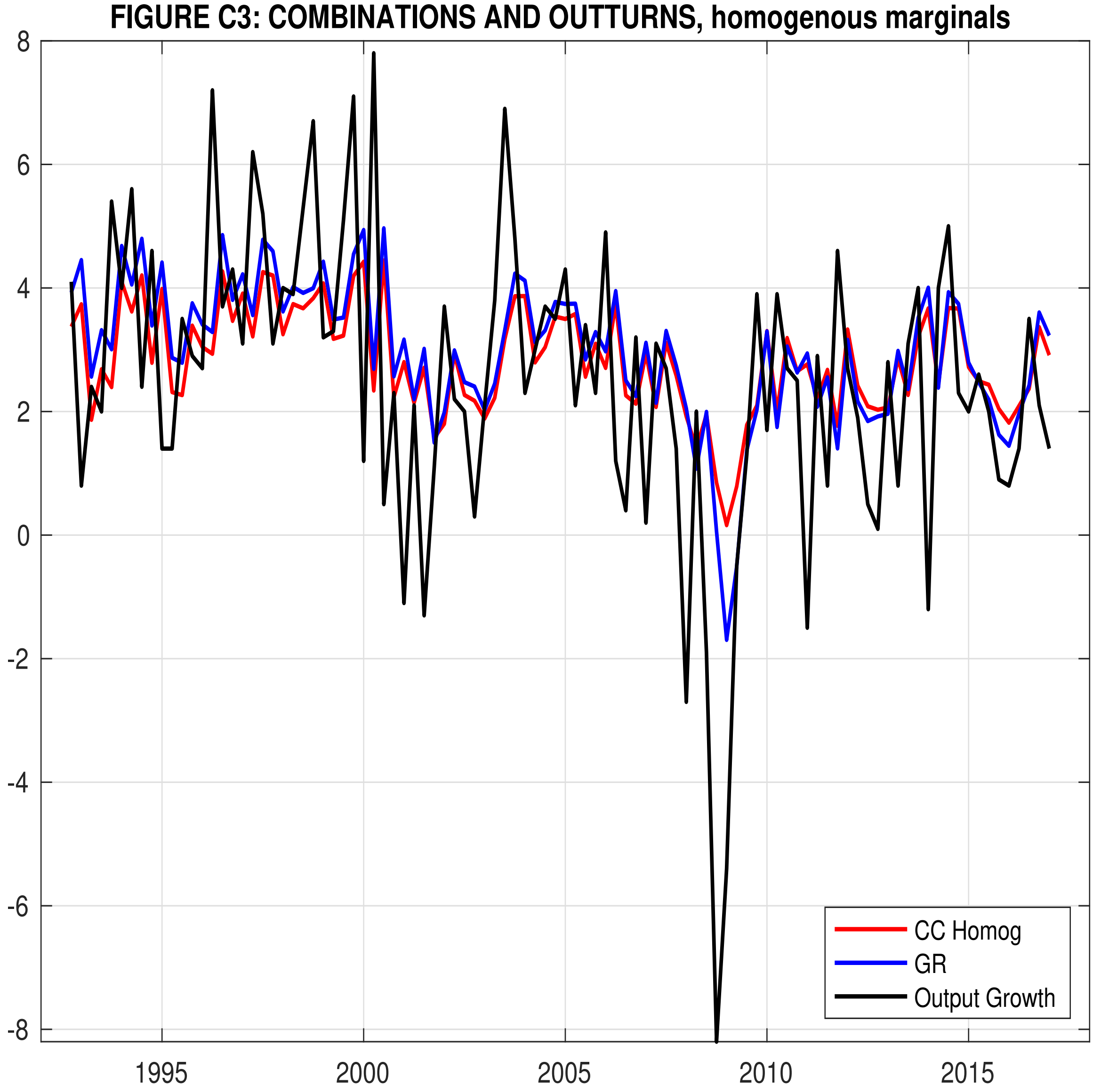


FIGURE C3: COMBINATIONS AND OUTTURNS, homogenous marginals
8
6
4
2
0
-2
-4
-6
-8
1995
2000
2005
2010
2015
CC Homog
GR
Output Growth

FIGURE C4: RECURSIVE RMSFE, COMBINATIONS, homogenous marginals