Efficient Prediction under Model Instabilities [*]

Tae-Hwy Lee[†] Shahnaz Parsaeian [‡] Aman Ullah[§]

Abstract
This paper aims to improve the prediction under model instabilities. Model instability is defined as a permanent change in the parameters of the model. We introduce a combined estimator of the post-break data and full-sample data and show that this combined estimator has a lower MSFE compared to the post-break estimator, which is a standard solution under model instabilities. The combination weight lies between zero and one. Monte Carlo experiment demonstrates our theoretical findings.

Key Words: Model Instabilities, Combined Estimation.

1. Introduction

Many macroeconomic and financial time series are subject to structural breaks. Since the 1960s, a voluminous literature on structural changes has been developed. Structural change in linear regressions was considered early on by Chow (1960), and Quandt (1960). Seminal works were mostly designed for the specific case of a single change. Bai (1997a) and Bai (1997b) study the least squares estimation of a regression model with a single break and with multiple breaks, respectively. Bai and Perron (1998) extend the sup-type test to models with multiple changes and propose a double maximum test against the alternative under which only the maximum number of breaks is prescribed. They also consider a sequential test for the null hypothesis of l breaks against the alternative of $l + 1$ breaks. The literature on detecting the structural break is massive and there are some cost efficient program to detect the breaks. For a comprehensive survey on structural changes, see Perron (2006) and Casini and Perron (2018).

Since the seminal work by Bates and Granger (1969), forecast combination has been proved to be an effective way to improve forecasting performance. Especially under model instability, the performance of the forecast can be boosted by forecast combinations method, see Diebold and Pauly (1987), Clements and Hendry (1998, 2006), Stock and Watson (2004), Pesaran and Timmermann (2005, 2007), Timmermann (2006), Pesaran and Pick (2011), Pesaran and Timmermann (2007), Rossi (2013), and Pesaran et al. (2013) inter alia.

The goal of this paper is to introduce a combined estimator which lowers the forecast error under model instabilities. The standard solution for forecasting under model instabilities is to use the post-break estimator, but this estimator can not improve the forecast when there is not enough observations in the post-sample. Our proposed combined estimator is the combination of the post-break estimator and the full-sample estimator with combination weight $w \in [0, 1]$. Using the pre-break data biases the forecast, but reduces the forecast error variance.

We undertake a Monte Carlo experiment with different set up for the break points, different numbers of the regressors, and various break sizes in both coefficients and error variances to compare the forecasting performance of our proposed combined estimator

[*]We thank the seminar participants at JSM (Denver) July 2019.
[†]Department of Economics, University of California, Riverside, CA 92521. E-mail: taelee@ucr.edu
[‡]Department of Economics, University of California, Riverside, CA 92521 .E-mail: spars002@ucr.edu
[§]Department of Economics, University of California, Riverside, CA 92521. E-mail: aman.ullah@ucr.edu

with the post-break estimator. We calculate the mean squared forecast error (MSFE) for the proposed combined estimator and the post-break estimator and evaluate their performance under various set up. This experiment confirms the out-performance of our proposed combined estimator relative to the post-break estimator in forecasting under structural breaks in the sense of MSFE.

The outline of the paper is as follows. Section 2 sets up the model under structural break and introduces the proposed combined estimator and its asymptotic risk. For simplicity, we discuss the problem under a single break, but generalization of the method to the multiple breaks is straightforward. Section 3 reports Monte Carlo simulation. Finally, Section 4 concludes.

2. The Structural Break Model

Consider the linear structural break model as $y_t = x_t'\beta_t + \sigma_t\varepsilon_t$, in which the $k \times 1$ vector of coefficients, β_t, and the error variance, σ_t, are subject to a break at time T_1. So we can write the model as:

$$y_t = \begin{cases} x_t'\beta_{(1)} + \sigma_{(1)}\varepsilon_t & \text{if} \quad 1 \ < t \leq T_1 \\ x_t'\beta_{(2)} + \sigma_{(2)}\varepsilon_t & \text{if} \quad T_1 < t < T, \end{cases} \tag{1}$$

where x_t is $k \times 1$, $\varepsilon_t \sim$ i.i.d.$(0, 1)$, $t \in \{1, \ldots, T\}$ and T_1 is the break point with $1 < T_1 < T$. In this set up we have only one break (two regimes). Assume that we know the break point.

2.1 Combined Estimator with Weight $w \in [0, 1]$

This section introduces the combined estimator of the post-break data estimator and full-sample data estimator with a combination weight $w \in [0, 1]$. The proposed estimator is:

$$\widehat{\beta}_w = w\widehat{\beta}_{Full} + (1 - w)\widehat{\beta}_{(2)}, \tag{2}$$

where $\widehat{\beta}_{Full}$ is the estimator for the coefficient by using full-sample data, $t = \{1, \ldots, T\}$, and $\widehat{\beta}_{(2)}$ is the estimator for the coefficient using the data after the breakpoint. Before deriving the asymptotic risk for this combined estimator, at first we want to derive the distribution for each estimator.

2.2 Full-Sample Estimator

The full-sample estimator is constructed under the null hypothesis that there is no break in the model, $\beta_{(1)} = \beta_{(2)}$, so it uses all of the observations to estimate β. Under the alternative hypothesis there is a break such that $\beta_1 = \beta_2 + \frac{\delta_1}{\sqrt{T}}$. We denote this full-sample estimator as $\widehat{\beta}_{Full}$, and estimate the coefficient by the Generalized Least Square (GLS) as:

$$\begin{aligned}
\widehat{\beta}_{Full} &= \Big(X'\Omega^{-1}X\Big)^{-1} X'\Omega^{-1}Y \\
&= \left(\sum_{t=1}^{T} x_t x_t' \frac{1}{\sigma_t^2}\right)^{-1} \sum_{t=1}^{T} x_t y_t \frac{1}{\sigma_t^2} \\
&= \left(\sum_{t=1}^{T} x_t x_t' \frac{1}{\sigma_t^2}\right)^{-1} \left[\sum_{t=1}^{T_1} x_t x_t' \frac{\beta_{(1)}}{\sigma_{(1)}^2} + \sum_{t=T_1+1}^{T} x_t x_t' \frac{\beta_{(2)}}{\sigma_{(2)}^2} + \sum_{t=1}^{T} \frac{x_t \sigma_t \varepsilon_t}{\sigma_t^2}\right] \\
&= \left(\sum_{t=1}^{T} x_t x_t' \frac{1}{\sigma_t^2}\right)^{-1} \left[\sum_{t=1}^{T_1} x_t x_t' \frac{\beta_{(1)} - \beta_{(2)}}{\sigma_{(1)}^2} + \sum_{t=1}^{T} x_t x_t' \frac{\beta_{(2)}}{\sigma_t^2} + \sum_{t=1}^{T} \frac{x_t \sigma_t \varepsilon_t}{\sigma_t^2}\right],
\end{aligned} \tag{3}$$

where $\Omega = \text{diag}(\sigma_{(1)}^2, \ldots, \sigma_{(1)}^2, \sigma_{(2)}^2, \ldots, \sigma_{(2)}^2)$ is a $T \times T$ matrix and $X = (x_1, x_2, \ldots, x_T)'$ $= (X_1' \; X_2')'$ is $T \times k$ matrix of regressors. So X_1 is $T_1 \times k$ matrix of observations before the break point, and X_2 is $(T - T_1) \times k$ matrix of observations after the break point. Assume that $T - T_1 \geq k + 1$, so we will have enough observations in the post-break sample. This is the assumption to have well defined estimator otherwise the number of unknown parameters would be larger than number of observations. By simplification:

$$\begin{aligned}
\sqrt{T}\Big(\widehat{\beta}_{Full} - \beta_{(2)}\Big) &= \Big(\sum_{t=1}^{T} x_t x_t' \frac{1}{T\sigma_t^2}\Big)^{-1} \left[\sum_{t=1}^{T_1} x_t x_t' \frac{b_1\sqrt{T}\Big(\beta_{(1)} - \beta_{(2)}\Big)}{Tb_1 \; \sigma_{(1)}^2} + \sum_{t=1}^{T} \frac{x_t \sigma_t \varepsilon_t}{\sqrt{T}\sigma_t^2}\right] \\
&= \Big(\frac{X'\Omega^{-1}X}{T}\Big)^{-1} \Big(\frac{X_1'\Omega_1^{-1}X_1}{Tb_1}\Big) \, b_1\delta_1 + \Big(\frac{X'\Omega^{-1}X}{T}\Big)^{-1} \Big(\sum_{t=1}^{T} \frac{x_t \sigma_t \varepsilon_t}{\sqrt{T}\sigma_t^2}\Big) \\
&\xrightarrow{d} N\Big(Q^{-1}Q_1 b_1 \delta_1, Q^{-1}\Big),
\end{aligned} \tag{4}$$

where $V_F = \Big(\frac{X'\Omega^{-1}X}{T}\Big)^{-1} \xrightarrow{p} Q^{-1}$ is the variance of the full-sample estimator, $b_1 = \frac{T_1}{T}$ shows the proportion of pre-break observations, $\Big(\frac{X_1'\Omega_1^{-1}X_1}{T_1}\Big) \xrightarrow{p} Q_1$ and Ω_1 is the variance of the pre-break data which is equal to $\Omega_1 = \text{diag}(\sigma_{(1)}^2, \ldots, \sigma_{(1)}^2)$. This is the two step GLS method in which in the first step we need to estimate the $\widehat{\sigma_{(1)}^2}$ and $\widehat{\sigma_{(2)}^2}$, and the second step estimates the coefficients.

2.3 Post-break Estimator

The post-break estimator focuses on the observations after the break point. So:

$$\widehat{\beta}_{(2)} = (X_2' \, \Omega_2^{-1} \, X_2)^{-1} X_2' \, \Omega_2^{-1} \, Y_2, \tag{5}$$

where $\Omega_2 = \text{diag}(\sigma_{(2)}^2, \ldots, \sigma_{(2)}^2)$ is a $(T - T_1) \times (T - T_1)$ matrix. By simplification:

$$\begin{aligned}
\sqrt{T}\Big(\widehat{\beta}_{(2)} - \beta_{(2)}\Big) &= \left(\frac{X_2' \, \Omega_2^{-1} \, X_2}{T - Tb_1}\right)^{-1} \left(\frac{X_2' \, \Omega_2^{-1} \, \sigma_{(2)}\varepsilon_2}{\sqrt{1 - b_1}\sqrt{T - Tb_1}}\right) \\
&\xrightarrow{d} N\Big(0, \frac{1}{1 - b_1} Q_2^{-1}\Big),
\end{aligned} \tag{6}$$

where $V_{(2)} = \frac{1}{1-b_1}\Big(\frac{X_2' \, \Omega_2^{-1} \, X_2}{T - T_1}\Big)^{-1} \xrightarrow{p} \frac{1}{1-b_1} Q_2^{-1}$ is the variance of the post-break estimator.

Theorem 1. *The joint distribution of the full-sample estimator and post-break estimator is:*

$$\sqrt{T}\begin{bmatrix}\widehat{\beta}_{Full}-\beta_{(2)}\\ \widehat{\beta}_{(2)}-\beta_{(2)}\end{bmatrix}\xrightarrow{d} V^{1/2}Z, \tag{7}$$

where $Z \sim \mathrm{N}(\theta, I_{2k}), \theta = V^{-1/2}\begin{bmatrix}Q^{-1}Q_1 b_1\delta_1\\ 0\end{bmatrix}, V = \begin{bmatrix}V_F & V_F\\ V_F & V_{(2)}\end{bmatrix}, V_F = \underset{T\to\infty}{\operatorname{plim}}\left(\frac{X'\Omega^{-1}X}{T}\right)^{-1}$ $= Q^{-1}$ *and* $V_{(2)} = \frac{1}{1-b_1}\underset{T\to\infty}{\operatorname{plim}}\left(\frac{X_2'\Omega_2^{-1}X_2}{T-T_1}\right)^{-1} = \frac{1}{1-b_1}Q_2^{-1}$.

Given that,

$$\sqrt{T}\big(\widehat{\beta}_{(2)}-\widehat{\beta}_{Full}\big)\xrightarrow{d} G'\,V^{1/2}Z \tag{8}$$

and

$$\sqrt{T}\big(\widehat{\beta}_{(2)}-\beta_{(2)}\big)\xrightarrow{d} G_2'V^{1/2}Z, \tag{9}$$

where $G = (-I\ \ I)'$ and $G_2 = (0\ \ I)'$.

□

Theorem 1 presents the joint asymptotic distribution of the full-sample estimator and the post-break estimator, and the combined estimator with weight $w \in [0,1]$. The joint asymptotic distribution of the full-sample estimator and the post-break estimator is normally distributed. Based on the derived distribution, the full-sample estimator is biased but more efficient. On the other side, the post-break estimator is unbiased but it is less efficient compared to the full- sample estimator because it is not using all available information in the sample. Thus, the proposed combined estimator exploits the trade-off between bias and variance of the two estimators.

2.4 Asymptotic Risk for the Combined Estimator

In this section we want to find the asymptotic risk for the proposed combined estimator for the case that we have only one break. Since our focus is on forecasting, it seems reasonable to consider $\beta_{(2)}$ as the true parameter vector in the definition of the risk.

Lemma 2. When an estimator has asymptotic distribution, $\sqrt{T}(\widehat{\beta}-\beta)\xrightarrow{d}\varpi$, then we can derive risk for this estimator as $\rho(\widehat{\beta},\mathbb{W}) = \mathbb{E}(\varpi'\,\mathbb{W}\,\varpi)$. See Lehmann and Casella (1998).

□

Based on Lemma 2, we can write the asymptotic risk for our estimator as:

$$\begin{aligned}
\rho(\widehat{\beta}_w, \mathbb{W}) &= \mathbb{E}\left[T(\widehat{\beta}_w - \beta_{(2)})' \, \mathbb{W} \, (\widehat{\beta}_w - \beta_{(2)})\right] \\
&= T\,\mathbb{E}\left[\left(\widehat{\beta}_{(2)} - \beta_{(2)}\right) - w\left(\widehat{\beta}_{(2)} - \widehat{\beta}_{Full}\right)\right]' \mathbb{W} \left[\left(\widehat{\beta}_{(2)} - \beta_{(2)}\right) - w\left(\widehat{\beta}_{(2)} - \widehat{\beta}_{Full}\right)\right] \\
&= \rho(\widehat{\beta}_{(2)}, \mathbb{W}) + w^2\,\mathbb{E}\left[Z'V^{1/2}G\,\mathbb{W}\,G'V^{1/2}Z\right] - 2w\,\mathbb{E}\left[Z'V^{1/2}G\,\mathbb{W}\,G_2'V^{1/2}Z\right] \\
&= \rho(\widehat{\beta}_{(2)}, \mathbb{W}) + w^2 \operatorname{tr}\left[V^{1/2}\,G\,\mathbb{W}\,G'\,V^{1/2}\;\mathbb{E}(ZZ')\right] \\
&- 2w \operatorname{tr}\left[V^{1/2}\,G\,\mathbb{W}\,G_2'\,V^{1/2}\;\mathbb{E}(ZZ')\right] \\
&= \rho(\widehat{\beta}_{(2)}, \mathbb{W}) + w^2 \operatorname{tr}\left[V^{1/2}\,G\,\mathbb{W}\,G'\,V^{1/2}\,(I_{2k} + \theta\theta')\right] \\
&- 2w \operatorname{tr}\left[V^{1/2}\,G\,\mathbb{W}\,G_2'\,V^{1/2}\,(I_{2k} + \theta\theta')\right] \\
&= \rho(\widehat{\beta}_{(2)}, \mathbb{W}) + w^2\,\left[\operatorname{tr}\Big(\mathbb{W}\left(V_{(2)} - V_F\right)\Big) + \theta'V^{1/2}G\,\mathbb{W}\,G'V^{1/2}\theta\right] \\
&- 2w \operatorname{tr}\Big(\mathbb{W}\left(V_{(2)} - V_F\right)\Big).
\end{aligned} \tag{10}$$

Theorem 3. *The asymptotic risk for the combined estimator with weight w can be written as:*

$$\begin{aligned}
\rho(\widehat{\beta}_w, \mathbb{W}) = \rho(\widehat{\beta}_{(2)}, \mathbb{W}) - w\,\Bigg(&2\operatorname{tr}\big(\mathbb{W}\left(V_{(2)} - V_F\right)\big) - w\Big[\operatorname{tr}\big(\mathbb{W}\left(V_{(2)} - V_F\right)\big) \\
&+ \theta'V^{1/2}G\,\mathbb{W}\,G'V^{1/2}\theta\Big]\Bigg).
\end{aligned} \tag{11}$$

□

Based on Theorem 3, the combined estimator with weight w has a smaller risk than the post-break estimator if the term inside the parenthesis be positive.

If $0 \leq w \leq \frac{2\operatorname{tr}\left(\mathbb{W}\left(V_{(2)} - V_F\right)\right)}{\operatorname{tr}\left(\mathbb{W}\left(V_{(2)} - V_F\right)\right) + \theta'V^{1/2}G\,\mathbb{W}\,G'V^{1/2}\theta}$, then this condition is satisfied and the optimal w^* can be derived as:

$$w^*(\mathbb{W}) = \frac{\operatorname{tr}(\mathbb{W}\left(V_{(2)} - V_F\right))}{\operatorname{tr}(\mathbb{W}\left(V_{(2)} - V_F\right)) + \theta'V^{1/2}G\,\mathbb{W}\,G'V^{1/2}\theta}. \tag{12}$$

Finally by plugging the optimal $w^*(\mathbb{W})$ in equation (12) into the risk function, we have the following theorem.

Theorem 4. *The asymptotic risk for the combined estimator with weight w can be derived by plugging back the optimal $w^*(\mathbb{W})$ in equation (12) into the risk function, and is equal to:*

$$\rho(\widehat{\beta}_w, \mathbb{W}) = \rho(\widehat{\beta}_{(2)}, \mathbb{W}) - \frac{\left[\operatorname{tr}(\mathbb{W}\left(V_{(2)} - V_F\right))\right]^2}{\operatorname{tr}(\mathbb{W}\left(V_{(2)} - V_F\right)) + \theta'V^{1/2}G\,\mathbb{W}\,G'V^{1/2}\theta}. \tag{13}$$

□

Theorem 4 shows that under the condition that $0 \leq w \leq \frac{2\operatorname{tr}\left(\mathbb{W}\left(V_{(2)} - V_F\right)\right)}{\operatorname{tr}\left(\mathbb{W}\left(V_{(2)} - V_F\right)\right) + \theta'V^{1/2}G\,\mathbb{W}\,G'V^{1/2}\theta}$, the risk of the combined estimator is less than the post-break estimator which means that there is a gain using the pre-break data under model instabilities.

Table 1: Simulation results with $T = 100$

		$q = 0.5$		$q = 1$	
	λ	$k = 3$	$k = 8$	$k = 3$	$k = 8$
	0	0.960	0.884	0.972	0.911
$b_1 = 0.6$	0.5	0.999	0.982	0.999	0.981
	1	0.999	0.997	0.999	0.996

3. Monte Carlo Simulations

This section provides some Monto Carlo results based on the combined estimator with weight $w \in [0, 1]$ and the post-break estimator. The goal is to compare the $MSFEs$ for these estimators. To do this, Let $t = \{1, ..., T\}$ with $T = 100$, $q = \frac{\sigma_{(1)}}{\sigma_{(2)}} \in \{0.5, 1\}$ shows the magnitude of the break in the error variances and $k = \{3, 8\}$ represents different number of regressors. We do the Monte Carlo for the case that the proportion of the pre-break sample observations, $b_1 = \frac{T_1}{T}$, is equal to 0.6. Suppose $x_t \sim N(0, 1)$, and $\varepsilon_t \sim i.i.d.\ N(0, 1)$. The data generating process for the single break case is:

$$y_t = \begin{cases} x_t'\beta_{(1)} + q\varepsilon_t & \text{if} \quad 1 \quad \leq t \leq T_1 \\ x_t'\beta_{(2)} + \varepsilon_t & \text{if} \quad T_1 < t \leq T. \end{cases} \tag{14}$$

Let $\beta_{(2)}$ be a vector of ones, and $\beta_{(1)} = \beta_{(2)} + \frac{\delta_1}{\sqrt{T}}$ under the local alternative. Assume that $\lambda \equiv \frac{\boldsymbol{\delta_1}}{\sqrt{T}} \in \{0, 0.5, 1\}$ represents different break sizes in the coefficient. The number of replications is 1000.

Table 1 shows the results of the Monte Carlo. We report the results based on Relative MSFE with respect to the post-break estimator which is the benchmark estimator, i.e.,

$$RMSFE_w = \frac{MSFE(\widehat{\beta}_w)}{MSFE(\widehat{\beta}_{(2)})}. \tag{15}$$

Table 1 shows the relative MSFE for different numbers of regressors, k, different break ratios in the error variance, q, and different break sizes in the coefficient, λ. Based on the results of Table 1, the relative MSFE are all less than one which means that the combined estimator has a lower MSFE than the post-break estimator. If the break happens towards the end of the sample (large b_1), the gain from using the combine estimator increases because the post-break estimator cannot perform well when there is not enough observations in the post-break sample. For the small break sizes, there is huge advantage in using the combined estimator and this advantage increases as the number of regressors (k) increases. For the large break size ($\lambda = 1$), the relative MSFE is almost one which means that the post-break estimator and the combined estimator perform equally under the large break. Actually for the large break size, there is not much gain using the combined estimator because the full-sample estimator adds more bias to the combine estimator.

4. Conclusion

In this paper we introduce the combined estimator of the full-sample estimator (using all observations in the sample), and the post-break estimator which uses the observations after the break point. The standard solution for forecasting under structural break is to use the post-break estimator, but it has been shown that using pre-break observations can decrease the MSFE. As our combined estimator uses the pre-break observations, we are able to reduce the variance of forecast error at the cost of adding some bias. Based on the simulation results, the combined estimator has a lower MSFE compared to the post-break estimator regardless of the breakpoint or break size.

REFERENCES

Bai, J. (1997a). “Estimation of a change point in multiple regression models.” Review of Economics and Statistics 79, 551-563.

Bai, J. (1997b). “Estimating multiple breaks one at a time.” Econometric Theory 13, 315-352.

Bai, J. and Perron, P. (1998). “Estimating and testing linear models with multiple structural changes.” Econometrica 66, 47-78.

Bai, J. and Perron, P. (2003). “Computation and analysis of multiple structural change models.” Journal of Applied Econometrics 18, 1-22.

Bates, J.M. and C.W.J. Granger (1969). “The Combination of Forecasts.” Operations Research Quarterly 20, 451-468.

Casini, A. and Perron, P. (2018). “Structural breaks in time series.” arXiv preprint arXiv:1805.03807.

Clements, M.P. and D.F. Hendry (1998). “Forecasting Economic Time Series.” Cambridge University Press.

Clements, M.P. and D.F. Hendry (2006). “Forecasting with Breaks.” In Elliott, G., C.W.J. Granger and A. Timmermann (Eds.), Handbook of Economic Forecasting vol. 1. North-Holland, 605-658

Chow, Gregory C. (1960). “Tests of Equality Between Sets of Coefficients in Two Linear Regressions.” Econometrica 28:3, 591-605.

Diebold, F.X. and P. Pauly (1987). “Structural Change and the Combination of Forecasts.” Journal of Forecasting 6, 21-40.

Lehmann, E. L., Casella, G. (1998). Theory of Point Estimation. 2nd ed. New York: Springer.

Perron, P. (2006). “Dealing with structural breaks.” In T.C. Mills & K. Patterson (eds.), Palgrave Hand-book of Econometrics, Econometric Theory 1, 278-352. Palgrave Macmillan.

Pesaran, M.H. and A. Timmermann (2005). “Small Sample Properties of Forecasts from Auto-regressive Models under Structural Breaks.” Journal of Econometrics 129, 183-217.

Pesaran, M.H., Andreas Pick and Mikhail Pranovich (2013). “Optimal forecasts in the presence of structural breaks.” Journal of Econometrics, 177(2), 134-152.

Pesaran, M.H., Timmermann, Allan (2007). “Selection of estimation window in the presence of breaks.” Journal of Econometrics 137, 134-161.

Pesaran, M.H., and Andreas Pick (2011). “Forecast Combination Across Estimation Windows.” Journal of Business and Economic Statistics, 29:2, 307-318, DOI: 10.1198/ jbes.2010.09018

Quandt, R. E. (1960). “Tests of thehypothesis that a linear regression obeys two separate regimes.” Journal of the American Statistical Association 55, 324-330.

Rossi, B. (2013). “Advances in forecasting under instability.” In G. Elliott, & A. Timmermann (Eds.), Handbook of Economic Forecasting, Chapter 21, 1203 - 1324. Elsevier volume 2, Part B.

Stock, J. H., and Watson, M. W. (2004). “Combination Forecasts of Output Growth in a Seven-Country Data Set.” Journal of Forecasting 23, 405-430.

Timmerman, Allan (2006). “Forecast Combinations.” In G. Elliott, C. W. Granger, and A. Timmermann (Eds.), Handbook of economic forecasting (pp. 135-196). Elsevier.