Granger Centre
for time series econometrics

Granger Centre Discussion Paper Series

Forecasting in factor augmented regressions under structural change

by

Daniele Massacci and George Kapetaniosy

Granger Centre Discussion Paper No. 21/02

University of Nottingham
UK | CHINA | MALAYSIA

Forecasting in Factor Augmented Regressions under Structural Change

Daniele Massacci* George Kapetanios†
King's College London King's College London

October 28, 2021

Abstract

Factor augmented regressions are widely used to produce out-of-sample forecasts of macroeconomic and financial time series. However, these series are subject to occasional breaks. We study the effect of neglected structural instability on the forecasts produced by factor augmented regressions when the latent factors are estimated by cross-sectional averages from a large panel of variables. Our results show that neglecting structural instability can be very costly in terms of forecasting performance. We derive analytical results to show that both instability in the factor model *and* in the forecasting equation have an impact on the produced forecasts. We further provide numerical results showing that conditioning upon the most recent break tends to produce more accurate forecasts than unconditional estimation methods based on expanding or rolling windows, although the actual gain depends on the location and the magnitude of the breaks.

JEL classification: C13, C32, C38, C53.

Keywords: Factor Augmented Regression, Structural Instability, Out-of-Sample Forecasts, Estimation Window, Cross-Sectional Averages.

*E-mail: daniele.massacci@kcl.ac.uk .
†E-mail: george.kapetanios@kcl.ac.uk .

1 Introduction

Factor augmented regressions are widely used to produce out-of-sample forecasts of macroeconomic and financial time series. For a given target variable, they consist of a forecasting equation in which one or more predictor is latent and it is estimated from a large panel of observable variables. Stock and Watson (2002a), and Bai and Ng (2006), provide seminal methodological contributions when the latent factors are estimated by asymptotic principal components, as studied in Bai and Ng (2002), and Bai (2003). On the empirical side, Stock and Watson (2002b), and Forni *et al.* (2018), employ factor augmented regressions to forecast macroeconomic variables; Ludvigson and Ng (2007), and Giovannelli *et al.* (2021), consider stock returns; Ludvigson and Ng (2009) look at bond risk premia.

The aforementioned contributions work under the maintained assumption that both the factor model and the forecasting equation are stable over time. However, the assumption of structural stability may not be realistic in practice. Stock and Watson (1996) find instabilities in macroeconomic time series. Pástor and Stambaugh (2001), and Timmermann (2001), show similar results for stock returns. Paye and Timmermann (2006), and Rapach and Wohar (2006), employ the procedure developed in Bai and Perron (1998), and document the presence of structural breaks in return prediction models. Timmermann (2008) argues that structural breaks in the data generating process of stock returns generate "pockets" of predictability, which are further analyzed in Farmer *et al.* (2021).

Given the existing evidence of breaks in the data generating process of macroeconomic and financial time series, a large body of literature has addressed the problem of forecasting under structural breaks: see Rossi (2021) for a general overview of the literature, and Timmermann (2018) for a specific focus on financial asset returns. However, to the very best of our knowledge, all existing contributions assume that the predictors of the target variable are either observable or, if latent, they are estimated from a large panel of variables exhibiting a factor structure assumed to be stable over time. For example, in the case of stock returns, observable predictors are provided in Welch and Goyal (2008), whereas Neely *et al.* (2014), Baetje and Menkhoff (2016), Çakmakli and van Dijk (2016), and Gonçalves *et al.* (2017), study predictions based on latent factors under the maintained assumption that the underlying factor model is not subject to structural instability. Within an *in-sample* framework, Corradi and Swanson (2014), and Massacci (2019), study estimation and inference in factor augmented regressions in which either the factor model or the forecasting equation (or both) are subject to breaks. However, to the very best of our knowledge, the literature is silent regarding the consequence of structural instability on the *out-of-sample* forecasting performance of factor augmented regressions.

This paper fills a gap in the literature by studying the problem of out-of-sample forecasting in factor augmented regressions when either the factor model or the forecasting equation (or both) are subject to structural instability. In particular, we focus on the situation in which the information stemming from the breaks is ignored and a misspecified linear model is used. Following Giacomini and White (2006), we focus on the forecasting method, which includes the model itself, as well as other choices made by the forecaster, such as the estimator for the model unknown parameters, and the related length of the estimation window. In terms of estimation, we follow Pesaran (2006) and consider cross-sectional averages estimation for the latent factors: this is appealing as it uses only the cross-sectional dimension, which is not affected by the break. In terms of estimation window, we ask ourselves whether we should use a conditional approach based on post-break observations only, or an unconditional approach that implements an expanding or a rolling window. The choice of the estimation window is addressed in Pesaran and Timmermann (2004), who however do not include latent factors and only consider observable predictors. Our work also complements Pesaran and Timmermann (2007), who study forecast combinations across estimation windows as a tool to mitigate the effect of structural instability on out-of-sample forecasts.

We study the simple yet informative set up of a factor augmented regression with a single latent factor, and one break in the factor model and in the forecasting equation. We obtain two results. First, we derive a closed form expression for the covariance between the realization and the forecast of the variable of interest, and we show how this depends on the choice of the estimation window in relation to the location of the breaks: in particular, when the estimation window begins after the break in the factor model, this does not affect the forecast; conversely, when the reverse occurs, the break in the factor model has an impact on the produced forecast due to rotational indeterminacy typical of latent factor models. Second, through a set of numerical results, we show that when the break in the factor model does not have an impact on the produced forecast, the post-break estimation window is likely to outperform the forecasts obtained by using the exponential and the rolling estimation windows. However, when the break in the factor model affects the produced forecast, the post-break estimation window may still have an edge, but this may depend on the sign and the magnitude of the breaks.

The rest of the paper is organized as follows. Section 2 sets up the problem. Section 3 derives analytical results that quantify the costs of ignoring breaks when using factor augmented regressions for forecasting purposes. Section 4 provides numerical results. Finally, Section 5 concludes. Mathematical proofs are provided in Appendix A.

2 Set up

We consider the model

$$\mathbf{x}_t = \mathbb{I}\left(t \leq T_{\mathbf{x}}^0\right) \boldsymbol{\lambda}_1 f_t + \mathbb{I}\left(t > T_{\mathbf{x}}^0\right) \boldsymbol{\lambda}_2 f_t + \mathbf{e}_t, \quad t = 1, \ldots, T, \quad 1 < T_{\mathbf{x}}^0 < T, \tag{1}$$

$$y_{t+1} = \mathbb{I}\left(t \leq T_y^0\right) \gamma_1 f_t + \mathbb{I}\left(t > T_y^0\right) \gamma_2 f_t + \varepsilon_{t+1}, \quad t = 1, \ldots, T, \quad 1 < T_y^0 < T, \tag{2}$$

where $\mathbb{I}(\cdot)$ is the indicator function and T denotes the time series dimension. Starting from (1), $\mathbf{x}_t = (x_{1t}, \ldots, x_{Nt})' \in \Re^N$ is the $N \times 1$ vector of observable dependent variables; f_t is the latent factor such that $\mathrm{E}(f_t) = 0$; $\mathbf{e}_t = (e_{1t}, \ldots, e_{Nt})' \in \Re^N$ is the $N \times 1$ vector of idiosyncratic components; $\boldsymbol{\lambda}_{j_\mathbf{x}} = (\lambda_{j_\mathbf{x}1}, \ldots, \lambda_{j_\mathbf{x}N})'$ is the $N \times 1$ vector of factor loadings in state $j_\mathbf{x} = 1, 2$, whose $i - th$ element is $\lambda_{j_\mathbf{x}i}$, for $i = 1, \ldots, N$; $T_\mathbf{x}^0$ is the break date in the data generating process of $\mathbf{x}_t$. Moving to (2), $y_{t+1} \in \Re$ is the dependent variable; f_t is the same factor entering (1); ε_{t+1} is the error term; γ_{j_y} is the slope coefficient associated to f_t in state $j_y = 1, 2$; T_y^0 is the break date, which is not constrained to be the same as $T_\mathbf{x}^0$.

The model in (1) and (2) is a factor augmented regression with structural instability. For ease of tractability, the model has one zero-mean factor and one break both in the factor model in (1) and in the forecasting model in (2). Our aim is to out-of-sample forecast y_{T+1} given the information available at time T when the breaks hit the data generating processes of $\mathbf{x}_t$ and y_t at $T_\mathbf{x}^0$ and T_y^0, respectively, before the forecast is made, so that $T_\mathbf{x}^0 < T$ and $T_y^0 < T$. In particular, we are interested in the situation in which the two breaks occur close to the end of the sample and enough observations after the breaks are not available to consistently estimate the model. Formally, this means that we study the case in which the time dimension T is fixed and does not tend to infinity.

It is customary in the literature to estimate linear factor augmented regressions using a two-step procedure, which first estimates the latent factors from a large panel of variables by asymptotic principal components, and then imputes the estimated factors into the forecasting model: see Bai and Ng (2006). While this procedure is valid in a linear setting, it may encounter problems when the model faces structural instability. In particular, asymptotic principal components estimation requires $N, T \to \infty$ at the same rate to achieve consistency (up to a rotation). Since T is fixed in our setting, the principal components estimator for the factor f_t in (1) would not in general be consistent. In order to overcome this issue, we follow an alternative route and estimate f_t using cross-sectional averages of the elements of $\mathbf{x}_t$, as originally proposed in Pesaran (2006). Cross-sectional averaging is appealing in the kind of problem we are facing since it only employs the cross-sectional dimension and thus require $N \to \infty$ only for consistency (up to a rotation), whereas the time series dimension T can be kept fixed.

3 Analytical results

To keep the analysis simple, we assume the number of latent factors (namely, one) in (1) is known. Following Pesaran and Timmermann (2004), we assume that also T_y^0 in (2) is known. Further, in (2) we let T_y^e be the pre-estimation window, with $1 \leq T_y^e \leq T_y^0$: the number of pre-break and post-break observations is $T_y^0 - T_y^e$ and $T - \left(T_y^0 + 1\right)$, respectively; the total number of observations to estimate the model is $T - \left(T_y^e + 1\right)$. Specification of the pre-estimation window T_y^e is required as the forecasting model in (2) is estimated along the time series dimension. On the other hand, we do not need to specify a pre-estimation window for the factor model in (1) since the latent factor f_t is estimated using the cross-sectional averages estimator, which only requires $N \to \infty$ to achieve consistency (up to a rotation).

In what follows, we assess the cost of ignoring the breaks occurring in $T_{\mathbf{x}}^0$ and T_y^0 along two complementary perspectives, namely their effect on the estimator for γ_2 in (2) and on the point forecast of y_{T+1}: these are covered in Sections 3.1 and 3.2, respectively.

3.1 Cross-sectional average estimation

Following Pesaran (2006), the cross-sectional average estimators $\hat{f}_t$ for f_t, and the least squares estimator $\hat{\gamma}_2\left(T_y^e\right)$ for γ_2, are

$$\hat{f}_t = \bar{x}_{\mathbf{w}t} = \sum_{i=1}^{N} w_i x_{it}, \quad \hat{\gamma}_2\left(T_y^e\right) = \left[\sum_{t=1}^{T-1} \mathbb{I}\left(t > T_y^e\right) \hat{f}_t \hat{f}_t\right]^{-1} \left[\sum_{t=1}^{T-1} \mathbb{I}\left(t > T_y^e\right) \hat{f}_t y_{t+1}\right], \quad t = 1, \ldots, T, \quad (3)$$

respectively, where $\{w_i\}_{i=1}^N$ is a sequence of weights. Let $\operatorname{diag}(\cdot)$ denote a diagonal matrix of suitable dimension. The following proposition characterizes the expected value of $\hat{\gamma}_2\left(T_y^e\right)$ as $N \to \infty$.

Proposition 3.1 *Given the model in* (1) *and* (2)*, let* $\mathbf{e}_t \sim \text{IID}\left(\mathbf{0}, \sigma_{\mathbf{e}}^2 \mathbf{I}_N\right)$ *and* $(f_t, \varepsilon_{t+1})' \sim \text{IID}\mathcal{N}\left[\mathbf{0}, \operatorname{diag}\left(\sigma_f^2, \sigma_\varepsilon^2\right)\right]$. *Consider* $\hat{\gamma}_2\left(T_y^e\right)$ *as defined in* (3)*, where the sequence of weights* $\{w_i\}_{i=1}^N$ *satisfies* $w_i = O\left(N^{-1}\right)$ *and* $\sum_{i=1}^N w_i = 1$. *Let* $\sum_{i=1}^N w_i \lambda_{j_{\mathbf{x}} i} \to \bar{\lambda}_{\mathbf{w} j_{\mathbf{x}}} \neq 0$, *for* $j_{\mathbf{x}} = 1, 2$. *Then*

$$\begin{aligned} &\lim_{N \to \infty} \mathrm{E}\left[\hat{\gamma}_2\left(T_y^e\right)\right] \\ = \;& \frac{\gamma_2}{\bar{\lambda}_{\mathbf{w}2}} \left\{ \begin{array}{c} \mathbb{I}\left(1 \leq T_{\mathbf{x}}^0 \leq T_y^e\right) \\ +\mathbb{I}\left(T_y^e < T_{\mathbf{x}}^0 \leq T-1\right) \left[\dfrac{T - \left(T_{\mathbf{x}}^0 + 1\right)}{T - \left(T_y^e + 1\right)} + \dfrac{T_{\mathbf{x}}^0 - T_y^e}{T - \left(T_y^e + 1\right)} \dfrac{\bar{\lambda}_{\mathbf{w}2}}{\bar{\lambda}_{\mathbf{w}1}}\right] \end{array} \right\} \\ &+ \frac{\gamma_1 - \gamma_2}{\bar{\lambda}_{\mathbf{w}2}} \left\{ \begin{array}{c} \mathbb{I}\left(1 \leq T_{\mathbf{x}}^0 \leq T_y^e\right) \dfrac{T_y^0 - T_y^e}{T - \left(T_y^e + 1\right)} \\ +\mathbb{I}\left(T_y^e < T_{\mathbf{x}}^0 \leq T-1\right) \left[\dfrac{T_y^0 - \min\left\{T_{\mathbf{x}}^0, T_y^0\right\}}{T - \left(T_y^e + 1\right)} + \dfrac{\min\left\{T_{\mathbf{x}}^0, T_y^0\right\} - T_y^e}{T - \left(T_y^e + 1\right)} \dfrac{\bar{\lambda}_{2\mathbf{w}}}{\bar{\lambda}_{\mathbf{w}1}}\right] \end{array} \right\}. \end{aligned}$$

Proposition 3.1 is informative about the asymptotic bias of $\hat{\gamma}_2\left(T_y^e\right)$ as $N \to \infty$: this is consistent with the analysis we are conducting, which assumes that the time series dimension T is fixed. The stringent assumption on $\mathbf{e}_t$ is imposed for expositional purposes only: Proposition 3.1 would still hold under suitable weaker conditions of time-series and cross-sectional dependence. In order to interpret Proposition 3.1, we consider three mutually exclusive cases: (*i*) $1 \leq T_{\mathbf{x}}^0 \leq T_y^e$; (*ii*) $T_y^e < T_{\mathbf{x}}^0 \leq T-1$ and $T_{\mathbf{x}}^0 \leq T_y^0$; (*iii*) $T_y^e < T_{\mathbf{x}}^0 \leq T-1$ and $T_{\mathbf{x}}^0 > T_y^0$.

If $1 \leq T^0_{\mathbf{x}} \leq T^e_y$ the break in the factor model occurs before the beginning of the estimation window in the forecasting equation and

$$\lim_{N\to\infty} \mathrm{E}\left[\hat{\gamma}_2\left(T^e_y\right)\right] = \frac{1}{\bar{\lambda}_{\mathbf{w}2}}\left[\gamma_2 + (\gamma_1 - \gamma_2)\frac{T^0_y - T^e_y}{T - \left(T^e_y + 1\right)}\right]. \tag{4}$$

Since the factor f_t in (1) is only identified up to a rotation, for $\gamma_1 = \gamma_2$, namely when the forecasting equation in (2) is not subject to structural instability, the right-hand side of (4) is equal to the rotation of γ_2 induced by $\bar{\lambda}^{-1}_{\mathbf{w}2}$ (namely, $\bar{\lambda}^{-1}_{\mathbf{w}2}\gamma_2$). For $\gamma_1 \neq \gamma_2$, the asymptotic bias of $\hat{\gamma}_2\left(T^e_y\right)$ depends on the magnitude of the break, as measured by $|\gamma_1 - \gamma_2|$, and by the ratio between the number of pre-break observations $\left(T^0_y - T^e_y\right)$ and the size of the estimation window $\left[T - \left(T^e_y + 1\right)\right]$. Notice that if f_t was observable and did not have to be estimated, then $\hat{\gamma}_2\left(T^e_y\right)$ and $\lim_{N\to\infty}\mathrm{E}\left[\hat{\gamma}_2\left(T^e_y\right)\right]$ in (3) and (4), respectively, would simplify to

$$\hat{\gamma}_2\left(T^e_y\right) = \left[\sum_{t=1}^{T-1}\mathbb{I}\left(t > T^e_y\right) f_t f_t\right]^{-1}\left[\sum_{t=T-1}^{T-1}\mathbb{I}\left(t > T^e_y\right) f_t y_{t+1}\right]$$

and

$$\mathrm{E}\left[\hat{\gamma}_2\left(T^e_y\right)\right] = \gamma_2 + (\gamma_1 - \gamma_2)\frac{T^0_y - T^e_y}{T - \left(T^e_y + 1\right)}, \tag{5}$$

respectively, where the analytical expression for $\mathrm{E}\left[\hat{\gamma}_2\left(T^e_y\right)\right]$ in (5) is identical to the analogous result in Pesaran and Timmermann (2004).[1]

If $T^e_y < T^0_{\mathbf{x}} \leq T - 1$ and $T^0_{\mathbf{x}} \leq T^0_y$ then $T^e_y < T^0_{\mathbf{x}} \leq T^0_y \leq T - 1$: the estimation window begins before the break in the factor model, which precedes the break in the forecasting equation. The result in Proposition 3.1 simplifies to

$$\begin{aligned}\lim_{N\to\infty} \mathrm{E}\left[\hat{\gamma}_2\left(T^e_y\right)\right] \quad = \quad & \frac{\gamma_2}{\bar{\lambda}_{\mathbf{w}2}}\left[\frac{T - \left(T^0_{\mathbf{x}} + 1\right)}{T - \left(T^e_y + 1\right)} + \frac{T^0_{\mathbf{x}} - T^e_y}{T - \left(T^e_y + 1\right)}\frac{\bar{\lambda}_{\mathbf{w}2}}{\bar{\lambda}_{\mathbf{w}1}}\right] \\ & + \frac{\gamma_1 - \gamma_2}{\bar{\lambda}_{\mathbf{w}2}}\left[\frac{T^0_y - T^0_{\mathbf{x}}}{T - \left(T^e_y + 1\right)} + \frac{T^0_{\mathbf{x}} - T^e_y}{T - \left(T^e_y + 1\right)}\frac{\bar{\lambda}_{\mathbf{w}2}}{\bar{\lambda}_{\mathbf{w}1}}\right].\end{aligned} \tag{6}$$

In the absence of a break in the forecasting model (i.e., $\gamma_1 = \gamma_2$) the result in (6) reduces to

$$\lim_{N\to\infty} \mathrm{E}\left[\hat{\gamma}_2\left(T^e_y\right)\right] = \frac{\gamma_2}{\bar{\lambda}_{\mathbf{w}2}}\left[\frac{T - \left(T^0_{\mathbf{x}} + 1\right)}{T - \left(T^e_y + 1\right)} + \frac{T^0_{\mathbf{x}} - T^e_y}{T - \left(T^e_y + 1\right)}\frac{\bar{\lambda}_{\mathbf{w}2}}{\bar{\lambda}_{\mathbf{w}1}}\right],$$

and the rotation induced around γ_2 depends on the frequency of observations of $\mathbf{x}_t$ before and after the break date $T^0_{\mathbf{x}}$, as given by $\left[T - \left(T^0_{\mathbf{x}} + 1\right)\right]/\left[T - \left(T^e_y + 1\right)\right]$ and $\left(T^0_{\mathbf{x}} - T^e_y\right)/\left[T - \left(T^e_y + 1\right)\right]$, respectively, and it is captured by $\bar{\lambda}_{\mathbf{w}2}/\bar{\lambda}_{\mathbf{w}1}$: when $\bar{\lambda}_{\mathbf{w}1} = \bar{\lambda}_{\mathbf{w}2}$, then $\lim_{N\to\infty}\mathrm{E}\left[\hat{\gamma}_2\left(T^e_y\right)\right] = \gamma_2/\bar{\lambda}_{\mathbf{w}2}$, which is the same it would be if there was no break in the factor model.

[1] See the Proof of Proposition 1 in Appendix A in Pesaran and Timmermann (2004).

Finally, if $T_y^e < T_{\mathbf{x}}^0 \leq T-1$ and $T_{\mathbf{x}}^0 > T_y^0$ then $T_y^e \leq T_y^0 < T_{\mathbf{x}}^0 \leq T-1$: the estimation window starts before the break in the forecasting model, which happens before the break in the factor model. The result stated in Proposition 3.1 simplifies to

$$\lim_{N\to\infty} \mathrm{E}\left[\hat{\gamma}_2\left(T_y^e\right)\right] = \frac{\gamma_2}{\bar{\lambda}_{\mathbf{w}2}}\left[\frac{T-\left(T_{\mathbf{x}}^0+1\right)}{T-\left(T_y^e+1\right)} + \frac{T_{\mathbf{x}}^0 - T_y^e}{T-\left(T_y^e+1\right)}\frac{\bar{\lambda}_{\mathbf{w}2}}{\bar{\lambda}_{\mathbf{w}1}}\right] + \frac{\gamma_1-\gamma_2}{\bar{\lambda}_{\mathbf{w}2}}\left[\frac{T_y^0 - T_y^e}{T-\left(T_y^e+1\right)}\frac{\bar{\lambda}_{\mathbf{w}2}}{\bar{\lambda}_{\mathbf{w}1}}\right], \quad (7)$$

and considerations analogous to those made in the previous case apply.

In conclusion, according to Proposition 3.1, $\lim_{N\to\infty} \mathrm{E}\left[\hat{\gamma}_2\left(T_y^e\right)\right]$ depends on the magnitude of the break in the forecasting equation as measured by $|\gamma_1-\gamma_2|$. An additional source of bias is due to f_t being latent, so that it has to be estimated from a large panel of variables that exhibit a factor structure. This extra source of bias persists even if the forecasting equation does not experience a break and follows from rotational indeterminacy typical of latent factor models: in particular, this bias depends on the relative position of $T_{\mathbf{x}}^0$, T_y^0 and T_y^e. Interestingly, rotational indeterminacy produces biases $\hat{\gamma}_2\left(T_y^e\right)$ only if $\bar{\lambda}_{\mathbf{w}1} \neq \bar{\lambda}_{\mathbf{w}2}$.

3.2 Point forecasts

Given the regression model in (2), the forecast of y_{T+1} at time T is $\hat{y}_{T+1}\left(T_y^e\right) = \hat{\gamma}_2\left(T_y^e\right)\hat{f}_T$, where $\hat{\gamma}_2\left(T_y^e\right)$ and $\hat{f}_T$ are defined in (3). Under the assumptions of Proposition 3.1, $\mathrm{E}\left(y_{T+1}\right) = \lim_{N\to\infty}\mathrm{E}\left[\hat{y}_{T+1}\left(T_y^e\right)\right] = 0$.[2] We thus assess the effect induced by structural instability on $\hat{y}_{T+1}\left(T_y^e\right)$ through the covariance between y_{T+1} and $\hat{y}_{T+1}\left(T_y^e\right)$.

Proposition 3.2 *Given the model in* (1) *and* (2)*, let the assumptions of Proposition 3.1 hold. Then*

$$\lim_{N\to\infty}\mathrm{E}\left[y_{T+1}\hat{y}_{T+1}\left(T_y^e\right)\right] = \gamma_2\sigma_f^2\bar{\lambda}_{\mathbf{w}2}\left\{\lim_{N\to\infty}\mathrm{E}\left[\hat{\gamma}_2\left(T_y^e\right)\right]\right\},$$

where $\lim_{N\to\infty}\mathrm{E}\left[\hat{\gamma}_2\left(T_y^e\right)\right]$ *is given in Proposition 3.1.*

Proposition 3.2 derives the analytical expression for the asymptotic covariance between y_{T+1} and its forecast $\hat{y}_{T+1}\left(T_y^e\right)$ as $N\to\infty$. As in the case of Proposition 3.1, we interpret Proposition 3.2 by considering the same mutually exclusive cases, namely: *(i)* $1 \leq T_{\mathbf{x}}^0 \leq T_y^e$; *(ii)* $T_y^e < T_{\mathbf{x}}^0 \leq T-1$ and $T_{\mathbf{x}}^0 \leq T_y^0$; *(iii)* $T_y^e < T_{\mathbf{x}}^0 \leq T-1$ and $T_{\mathbf{x}}^0 > T_y^0$.

When $1 \leq T_{\mathbf{x}}^0 \leq T_y^e$, taking into account (4), Proposition 3.2 simplifies to

$$\lim_{N\to\infty}\mathrm{E}\left[y_{T+1}\hat{y}_{T+1}\left(T_y^e\right)\right] = \gamma_2\sigma_f^2\left[\gamma_2 + \left(\gamma_1-\gamma_2\right)\frac{T_y^0-T_y^e}{T-\left(T_y^e+1\right)}\right],$$

[2]Under the assumptions of Proposition 3.1, $\hat{\gamma}_2\left(T_y^e\right)$ and $\hat{f}_T$ are independent random variables and

$$\lim_{N\to\infty}\mathrm{E}\left(\hat{f}_T\right) = \lim_{N\to\infty}\mathrm{E}\left(\sum_{i=1}^{N} w_i x_{iT}\right) = 0.$$

which is identical to the homologous finding stated in Proposition 1 in Pesaran and Timmermann (2004). Unlike the result in (4), in this case the asymptotic (as $N \to \infty$) covariance between y_{T+1} and $\hat{y}_{T+1}\left(T_y^e\right)$ does not suffer from the rotational indeterminacy problem induced by the latent factor model.

When $T_y^e < T_{\mathbf{x}}^0 \leq T-1$ and $T_{\mathbf{x}}^0 \leq T_y^0$, from (6) the result in Proposition 3.2 becomes

$$\lim_{N\to\infty} \mathrm{E}\left[y_{T+1}\hat{y}_{T+1}\left(T_y^e\right)\right] = \gamma_2\sigma_f^2 \left\{ \begin{array}{l} \gamma_2 \left[\dfrac{T-\left(T_{\mathbf{x}}^0+1\right)}{T-\left(T_y^e+1\right)} + \dfrac{T_{\mathbf{x}}^0-T_y^e}{T-\left(T_y^e+1\right)}\dfrac{\bar{\lambda}_{\mathbf{w}2}}{\bar{\lambda}_{\mathbf{w}1}}\right] \\ +\left(\gamma_1-\gamma_2\right)\left[\dfrac{T_y^0-T_{\mathbf{x}}^0}{T-\left(T_y^e+1\right)} + \dfrac{T_{\mathbf{x}}^0-T_y^e}{T-\left(T_y^e+1\right)}\dfrac{\bar{\lambda}_{\mathbf{w}2}}{\bar{\lambda}_{\mathbf{w}1}}\right] \end{array} \right\}.$$

In this case, the source of dependence between y_{T+1} and $\hat{y}_{T+1}\left(T_y^e\right)$ induced by $\bar{\lambda}_{\mathbf{w}2}/\bar{\lambda}_{\mathbf{w}1}$ arises. The ratio $\bar{\lambda}_{\mathbf{w}2}/\bar{\lambda}_{\mathbf{w}1}$ plays a role because $T_y^e < T_{\mathbf{x}}^0 \leq T-1$, namely because the break in the factor model occurs after the beginning of the estimation window, and the effects induced by rotational indeterminacy before and after the break do not cancel each other out (unless $\bar{\lambda}_{\mathbf{w}1} = \bar{\lambda}_{\mathbf{w}2}$).

Finally, for $T_y^e < T_{\mathbf{x}}^0 \leq T-1$ and $T_{\mathbf{x}}^0 > T_y^0$, from (7) the result in Proposition 3.2 simplifies to

$$\lim_{N\to\infty} \mathrm{E}\left[y_{T+1}\hat{y}_{T+1}\left(T_y^e\right)\right] = \gamma_2\sigma_f^2 \left\{\gamma_2 \left[\frac{T-\left(T_{\mathbf{x}}^0+1\right)}{T-\left(T_y^e+1\right)} + \frac{T_{\mathbf{x}}^0-T_y^e}{T-\left(T_y^e+1\right)}\frac{\bar{\lambda}_{\mathbf{w}2}}{\bar{\lambda}_{\mathbf{w}1}}\right] + \left(\gamma_1-\gamma_2\right)\left[\frac{T_y^0-T_y^e}{T-\left(T_y^e+1\right)}\frac{\bar{\lambda}_{\mathbf{w}2}}{\bar{\lambda}_{\mathbf{w}1}}\right]\right\},$$

and a component in the comovement between y_{T+1} and $\hat{y}_{T+1}\left(T_y^e\right)$ driven by $\bar{\lambda}_{\mathbf{w}2}/\bar{\lambda}_{\mathbf{w}1}$ still persists.

In conclusion, the comovement between y_{T+1} and $\hat{y}_{T+1}\left(T_y^e\right)$, as measured by their asymptotic covariance as $N \to \infty$, depends on the magnitude of the break as captured by $|\gamma_1 - \gamma_2|$. When the break in the factor model occurs after the beginning of the estimation window in the forecasting model, the comovement between y_{T+1} and $\hat{y}_{T+1}\left(T_y^e\right)$ also depends upon the ratio $\bar{\lambda}_{\mathbf{w}2}/\bar{\lambda}_{\mathbf{w}1}$, which is induced by rotational indeterminacy since the estimator $\hat{f}_t$ for f_t in general experiences different rotations around f_t because of the structural break in the factor model.

4 Numerical results

4.1 Data generating process

We consider the data generating process

$$x_{it}^s = \mathbb{I}\left(t \leq T_{\mathbf{x}}^0\right)\lambda_{1i}f_t^s + \mathbb{I}\left(t > T_{\mathbf{x}}^0\right)\lambda_{2i}f_t^s + e_{it}^s, \quad i = 1,\ldots,N, \quad t = 1,\ldots,T,$$

where s denotes the replication, for $s = 1,\ldots,2000$. We consider $N = 100$ and $T = 201$, so that $T-1 = 200$. Define $\delta_{\mathbf{x}i} = \lambda_{1i} - \lambda_{2i}$, for $i = 1,\ldots,N$. We fix λ_{1i} and λ_{2i}, and thus $\delta_{\mathbf{x}i}$, throughout the replications. We generate $\lambda_{2i} \sim \mathcal{N}(1,1)$, and define $\lambda_{1i} = \lambda_{2i} + \delta_{\mathbf{x}i}$. We control for the magnitude of the break by implementing the following scenarios: in Experiment 1 we set $\delta_{\mathbf{x}i} = 0.00, 1.50, 3.00$ for

$i = 1, \ldots, N/2$, and correspondingly $\delta_{\mathbf{x}i} = 0.00, -1.50, -3.00$ for $i = N/2 + 1, \ldots, N$; in Experiment 2a we look at $\delta_{\mathbf{x}i} = \delta_{\mathbf{x}} = 0.00, 1.50, 3.00$; in Experiment 2b we consider $\delta_{\mathbf{x}i} = \delta_{\mathbf{x}} = 0.00, -1.50, -3.00$. Experiment 1 implies that the condition $\bar{\lambda}_{\mathbf{w}2} / \bar{\lambda}_{\mathbf{w}1} = 1$ in Proposition 3.2 is met and the break in the factor model should not affect the produced forecasts. Experiment 2a and Experiment 2b imply that $0 < \bar{\lambda}_{\mathbf{w}2} / \bar{\lambda}_{\mathbf{w}1} \leq 1$ in Proposition 3.2 and thus allow to assess the impact of the break in the factor model on the forecasts. We further control for the break date by setting $T_{\mathbf{x}}^0 = 100, 190$. The factor is generated as $f_t^s \sim \text{IID}\mathcal{N}(0,1)$. The idiosyncratic components are generated as $e_{it}^s = \sigma_{ii}^{1/2} \epsilon_{e,it}^s$, with $\sigma_{ii} \sim \chi^2(1)$ and $\epsilon_{e,it}^s \sim \text{IID}\mathcal{N}(0,1)$, with σ_{ii} fixed throughout the replications.

The data generating process for the target variable is

$$y_{t+1}^s = \mathbb{I}\left(t \leq T_y^0\right) \gamma_1 f_t^s + \mathbb{I}\left(t > T_y^0\right) \gamma_2 f_t^s + \varepsilon_{t+1}^s, \quad t = 1, \ldots, T.$$

The slope coefficients γ_1 and γ_2 are fixed throughout the replications, with $\gamma_2 = 1$ and $\gamma_1 = \gamma_2 + \delta_y$. We control for the magnitude of the break by setting $\delta_y = 0.00, 1.00, 2.00, 3.00$, and for the location of the break by fixing $T_y^0 = 100, 190$. The error term ε_{t+1}^s is generated as $\varepsilon_{t+1}^s \sim \text{IID}\mathcal{N}(0,1)$.

We consider three estimation windows: post-break, with $T_y^e = T_y^0$; expanding, with $T_y^e = 0$; rolling, with $T_y^e = T - 1 - w$ and $w = 50$, so that $T_y^e = 150$. From the discussion in Section 3.2, a necessary condition for $\bar{\lambda}_{\mathbf{w}2} / \bar{\lambda}_{\mathbf{w}1}$ to have an effect on the produced forecast is that $T_y^e < T_{\mathbf{x}}^0 \leq T - 1$. Given our data generating process, $\bar{\lambda}_{\mathbf{w}2} / \bar{\lambda}_{\mathbf{w}1}$ will always impact the forecast in the case of the expanding estimation window, provided that $\bar{\lambda}_{\mathbf{w}2} / \bar{\lambda}_{\mathbf{w}1} \neq 1$. In the case of post-break and rolling estimation windows, the effect induced by $\bar{\lambda}_{\mathbf{w}2} / \bar{\lambda}_{\mathbf{w}1}$ depends on the position of T_y^e relative to $T_{\mathbf{x}}^0$. Also, since we keep λ_{2i} and γ_2 constant, for $i = 1, \ldots, N$, the forecasts produced using the post-break window are independent of the break size in both the factor model and in the factor augmented regression for $T_y^0 \geq T_{\mathbf{x}}^0$.

We evaluate the produced forecasts in terms of the root mean squared forecast error defined as

$$\text{RMSFE}_k = \frac{\sum_{s=1}^S \left[y_{T+1}^s - \hat{y}_{k,T+1}^s\left(T_y^e\right)\right]^2}{S}, \quad k = \text{post-break, expanding, rolling}, \tag{8}$$

where $\hat{y}_{k,T+1}^s\left(T_y^e\right)$ is the forecast made by method k within replication s.

4.2 Results

Table 1 about here

The results from Experiment 1 and collected in Table 1 are consistent with Proposition 3.2. When $\bar{\lambda}_{\mathbf{w}2} / \bar{\lambda}_{\mathbf{w}1} = 1$, the produced forecast is independent of the size of the break in the factor model as measured by $|\delta_{\mathbf{x}}|$. As expected, when $\delta_y = 0$ the expanding window always produces the most accurate forecasts, since it correctly uses all available information. As δ_y increases, the post-break estimation window takes the lead, whereas the expanding window becomes the worst performer, as it is the method that employs the highest amount of wrong information stemming from the observations before the break. Notice that all forecasts deteriorate as T_y^0 increases from $T_y^0 = 100$ to $T_y^0 = 190$, since fewer observations become available in the post-break window, and the expanding and rolling windows use more pre-break observations to estimate the forecasting model.

Table 2a and 2b about here

Table 2a and Table 2b collect results from Experiments 2a and 2b, respectively, and show a different picture compared to Table 1. In both cases, the magnitude of the ratio $\bar{\lambda}_{\mathbf{w}2} / \bar{\lambda}_{\mathbf{w}1}$ declines in $\delta_{\mathbf{x}}$ and the estimator for the factor before the break becomes less precise due to the increased bias induced by rotational indeterminacy. In some cases the conclusions drawn from Experiment 1 are reversed. When $\delta_y > 0$ and $\delta_{\mathbf{x}} > 0$, the post-break estimation window is often dominated by the expanding window (see Table 2a). However, this is not the case for $\delta_{\mathbf{x}} < 0$, when in some instances the rolling window produces the most accurate forecasts (see Table 2b). Therefore, when $0 < \bar{\lambda}_{\mathbf{w}2} / \bar{\lambda}_{\mathbf{w}1} < 1$, the post-break window still has an overall edge with respect to expanding and rolling window estimation methods, although the actual performance needs to be evaluated on a case-by-case basis. In particular, the advantage in terms of forecasting performance the post-break window has when $\bar{\lambda}_{\mathbf{w}2} / \bar{\lambda}_{\mathbf{w}1} = 1$ no longer uniformly holds when $\bar{\lambda}_{\mathbf{w}2} / \bar{\lambda}_{\mathbf{w}1} \neq 1$, in which case the relative performance of the estimation window depends on the interaction between the breaks in the forecasting equation and in the factor model.

5 Conclusions

This paper studies out-of-sample forecasting in factor augmented regressions that experience structural instability in the factor model or in the forecasting regression (or both) when the latent factors are estimated by cross-sectional averages and the instability is neglected. We show that a post-break estimation window tends to produce more accurate forecasts than the expanding or the rolling estimation windows, although the actual relative precision depends on the position and the magnitude of the breaks in the factor model and in the forecasting equation. This poses challenges as to how optimally select the estimation window in the forecasting model.

Our work can be extended along several dimensions. We specifically focus on the case in which the forecasting equation has one latent factor and does not include any observable predictor: the more general case with multiple latent factors and observable predictors is an extension worth considering. Also, we estimated the latent factor by cross-sectional averages: it would be interesting to compare this with the asymptotic principal components estimator commonly used in factor augmented regressions. Finally, this paper uses an approach based on unsupervised learning and a comparison with a supervised counterpart in the spirit of Bair *et al.* (2006) is worth considering. All these extensions will be conducted in future research.

A Appendix

Proof of Proposition 3.1. As $N \to \infty$,

$$\begin{aligned} \hat{f}_t &= \mathbb{I}\left(t \le T_{\mathbf{x}}^0\right)\left(\sum_{i=1}^{N} w_i \lambda_{1i}\right) f_t + \mathbb{I}\left(t > T_{\mathbf{x}}^0\right)\left(\sum_{i=1}^{N} w_i \lambda_{2i}\right) f_t + \left(\sum_{i=1}^{N} w_i e_{it}\right) \\ &\xrightarrow{p} \mathbb{I}\left(t \le T_{\mathbf{x}}^0\right) \bar{\lambda}_{\mathbf{w}1} f_t + \mathbb{I}\left(t > T_{\mathbf{x}}^0\right) \bar{\lambda}_{\mathbf{w}2} f_t : \end{aligned}$$

it follows that as $N \to \infty$

$$\begin{aligned} & \hat{\gamma}_2\left(T_y^e\right) \\ \xrightarrow{p} & \left\{\sum_{t=1}^{T-1} \mathbb{I}\left(t > T_y^e\right)\left[\mathbb{I}\left(t \le T_{\mathbf{x}}^0\right) \bar{\lambda}_{\mathbf{w}1} f_t + \mathbb{I}\left(t > T_{\mathbf{x}}^0\right) \bar{\lambda}_{\mathbf{w}2} f_t\right]\left[\mathbb{I}\left(t \le T_{\mathbf{x}}^0\right) \bar{\lambda}_{\mathbf{w}1} f_t + \mathbb{I}\left(t > T_{\mathbf{x}}^0\right) \bar{\lambda}_{\mathbf{w}2} f_t\right]\right\}^{-1} \\ & \times\left\{\sum_{t=1}^{T-1} \mathbb{I}\left(t > T_y^e\right)\left[\mathbb{I}\left(t \le T_{\mathbf{x}}^0\right) \bar{\lambda}_{\mathbf{w}1} f_t + \mathbb{I}\left(t > T_{\mathbf{x}}^0\right) \bar{\lambda}_{\mathbf{w}2} f_t\right]\left[\mathbb{I}\left(t \le T_y^0\right) \gamma_1 f_t + \mathbb{I}\left(t > T_y^0\right) \gamma_2 f_t + \varepsilon_{t+1}\right]\right\} \\ = & \left\{\sum_{t=1}^{T-1} \mathbb{I}\left(t > T_y^e\right)\left[\mathbb{I}\left(t \le T_{\mathbf{x}}^0\right) \bar{\lambda}_{\mathbf{w}1} \bar{\lambda}_{\mathbf{w}1} f_t f_t + \mathbb{I}\left(t > T_{\mathbf{x}}^0\right) \bar{\lambda}_{\mathbf{w}2} \bar{\lambda}_{\mathbf{w}2} f_t f_t\right]\right\}^{-1} \\ & \times\left\{\sum_{t=1}^{T-1} \mathbb{I}\left(t > T_y^e\right)\left[\begin{array}{c} \mathbb{I}\left(t \le T_{\mathbf{x}}^0\right) \mathbb{I}\left(t \le T_y^0\right) \bar{\lambda}_{\mathbf{w}1} f_t \gamma_1 f_t + \mathbb{I}\left(t \le T_{\mathbf{x}}^0\right) \mathbb{I}\left(t > T_y^0\right) \bar{\lambda}_{\mathbf{w}1} f_t \gamma_2 f_t + \mathbb{I}\left(t \le T_{\mathbf{x}}^0\right) \bar{\lambda}_{\mathbf{w}1} f_t \varepsilon_{t+1} \\ +\mathbb{I}\left(t > T_{\mathbf{x}}^0\right) \mathbb{I}\left(t \le T_y^0\right) \bar{\lambda}_{\mathbf{w}2} f_t \gamma_1 f_t + \mathbb{I}\left(t > T_{\mathbf{x}}^0\right) \mathbb{I}\left(t > T_y^0\right) \bar{\lambda}_{\mathbf{w}2} f_t \gamma_2 f_t + \mathbb{I}\left(t > T_{\mathbf{x}}^0\right) \bar{\lambda}_{\mathbf{w}2} f_t \varepsilon_{t+1} \end{array}\right]\right\} \\ = & \mathbb{I}\left(T_{\mathbf{x}}^0 \le T_y^e\right)\left[\sum_{t=1}^{T-1} \mathbb{I}\left(t > T_y^e\right) \mathbb{I}\left(t > T_{\mathbf{x}}^0\right) \bar{\lambda}_{\mathbf{w}2} f_t f_t\right]^{-1} \\ & \times\left\{\sum_{t=1}^{T-1} \mathbb{I}\left(t > T_y^e\right) \mathbb{I}\left(t > T_{\mathbf{x}}^0\right)\left[\mathbb{I}\left(t \le T_y^0\right) \gamma_1 f_t f_t + \mathbb{I}\left(t > T_y^0\right) \gamma_2 f_t f_t + f_t \varepsilon_{t+1}\right]\right\} \\ & +\mathbb{I}\left(T_{\mathbf{x}}^0 > T_y^e\right) \frac{T_{\mathbf{x}}^0 - T_y^e}{T - \left(T_y^e + 1\right)}\left[\sum_{t=1}^{T-1} \mathbb{I}\left(t > T_y^e\right) \mathbb{I}\left(t \le T_{\mathbf{x}}^0\right) \bar{\lambda}_{\mathbf{w}1} f_t f_t\right]^{-1} \\ & \times\left\{\sum_{t=1}^{T-1} \mathbb{I}\left(t > T_y^e\right) \mathbb{I}\left(t \le T_{\mathbf{x}}^0\right)\left[\mathbb{I}\left(t \le T_y^0\right) \gamma_1 f_t f_t + \mathbb{I}\left(t > T_y^0\right) \gamma_2 f_t f_t + f_t \varepsilon_{t+1}\right]\right\} \\ & +\mathbb{I}\left(T_{\mathbf{x}}^0 > T_y^e\right) \frac{T - \left(T_{\mathbf{x}}^0 + 1\right)}{T - \left(T_y^e + 1\right)}\left[\sum_{t=1}^{T-1} \mathbb{I}\left(t > T_y^e\right) \mathbb{I}\left(t > T_{\mathbf{x}}^0\right) \bar{\lambda}_{\mathbf{w}2} f_t f_t\right]^{-1} \\ & \times\left\{\sum_{t=1}^{T-1} \mathbb{I}\left(t > T_y^e\right) \mathbb{I}\left(t > T_{\mathbf{x}}^0\right)\left[\mathbb{I}\left(t \le T_y^0\right) \gamma_1 f_t f_t + \mathbb{I}\left(t > T_y^0\right) \gamma_2 f_t f_t + f_t \varepsilon_{t+1}\right]\right\}, \end{aligned}$$

and since $\mathbb{I}\left(t > T_y^0\right) = 1 - \mathbb{I}\left(t \leq T_y^0\right)$, as $N \to \infty$

$$
\begin{aligned}
& \hat{\gamma}_2\left(T_y^e\right) \\
\xrightarrow{p}\; & \mathbb{I}\left(T_{\mathbf{x}}^0 \leq T_y^e\right)\left[\sum_{t=1}^{T-1} \mathbb{I}\left(t > T_y^e\right)\mathbb{I}\left(t > T_{\mathbf{x}}^0\right) f_t f_t\right]^{-1} \\
& \times\left\{\sum_{t=1}^{T-1} \mathbb{I}\left(t > T_y^e\right)\mathbb{I}\left(t > T_{\mathbf{x}}^0\right)\left[\mathbb{I}\left(t \leq T_y^0\right)\frac{\gamma_1}{\bar{\lambda}_{\mathbf{w}2}} f_t f_t + \mathbb{I}\left(t > T_y^0\right)\frac{\gamma_2}{\bar{\lambda}_{\mathbf{w}2}} f_t f_t + \frac{1}{\bar{\lambda}_{\mathbf{w}2}} f_t \varepsilon_{t+1}\right]\right\} \\
& + \mathbb{I}\left(T_{\mathbf{x}}^0 > T_y^e\right)\frac{T_{\mathbf{x}}^0 - T_y^e}{T - \left(T_y^e + 1\right)}\left[\sum_{t=1}^{T-1} \mathbb{I}\left(t > T_y^e\right)\mathbb{I}\left(t \leq T_{\mathbf{x}}^0\right) f_t f_t\right]^{-1} \\
& \times\left\{\sum_{t=1}^{T-1} \mathbb{I}\left(t > T_y^e\right)\mathbb{I}\left(t \leq T_{\mathbf{x}}^0\right)\left[\mathbb{I}\left(t \leq T_y^0\right)\frac{\gamma_1}{\bar{\lambda}_{\mathbf{w}1}} f_t f_t + \mathbb{I}\left(t > T_y^0\right)\frac{\gamma_2}{\bar{\lambda}_{\mathbf{w}1}} f_t f_t + \frac{1}{\bar{\lambda}_{\mathbf{w}1}} f_t \varepsilon_{t+1}\right]\right\} \\
& + \mathbb{I}\left(T_{\mathbf{x}}^0 > T_y^e\right)\frac{T - \left(T_{\mathbf{x}}^0 + 1\right)}{T - \left(T_y^e + 1\right)}\left[\sum_{t=1}^{T-1} \mathbb{I}\left(t > T_y^e\right)\mathbb{I}\left(t > T_{\mathbf{x}}^0\right) f_t f_t\right]^{-1} \\
& \times\left\{\sum_{t=1}^{T-1} \mathbb{I}\left(t > T_y^e\right)\mathbb{I}\left(t > T_{\mathbf{x}}^0\right)\left[\mathbb{I}\left(t \leq T_y^0\right)\frac{\gamma_1}{\bar{\lambda}_{\mathbf{w}2}} f_t f_t + \mathbb{I}\left(t > T_y^0\right)\frac{\gamma_2}{\bar{\lambda}_{\mathbf{w}2}} f_t f_t + \frac{1}{\bar{\lambda}_{\mathbf{w}2}} f_t \varepsilon_{t+1}\right]\right\} \\
= \; & \mathbb{I}\left(T_{\mathbf{x}}^0 \leq T_y^e\right)\frac{\gamma_2}{\bar{\lambda}_{\mathbf{w}2}} \\
& + \mathbb{I}\left(T_{\mathbf{x}}^0 \leq T_y^e\right)\frac{\gamma_1 - \gamma_2}{\bar{\lambda}_{\mathbf{w}2}}\left[\sum_{t=1}^{T-1} \mathbb{I}\left(t > T_y^e\right) f_t f_t\right]^{-1}\left[\sum_{t=1}^{T-1} \mathbb{I}\left(t > T_y^e\right)\mathbb{I}\left(t \leq T_y^0\right) f_t f_t\right] \\
& + \mathbb{I}\left(T_{\mathbf{x}}^0 \leq T_y^e\right)\frac{1}{\bar{\lambda}_{\mathbf{w}2}}\left[\sum_{t=1}^{T-1} \mathbb{I}\left(t > T_y^e\right) f_t f_t\right]^{-1}\left[\sum_{t=1}^{T-1} \mathbb{I}\left(t > T_y^e\right) f_t \varepsilon_t\right] \\
& + \mathbb{I}\left(T_{\mathbf{x}}^0 > T_y^e\right)\frac{\gamma_2}{\bar{\lambda}_{\mathbf{w}1}}\frac{T_{\mathbf{x}}^0 - T_y^e}{T - \left(T_y^e + 1\right)} \\
& + \mathbb{I}\left(T_{\mathbf{x}}^0 > T_y^e\right)\frac{\gamma_1 - \gamma_2}{\bar{\lambda}_{\mathbf{w}1}}\frac{T_{\mathbf{x}}^0 - T_y^e}{T - \left(T_y^e + 1\right)}\left[\sum_{t=1}^{T-1} \mathbb{I}\left(t > T_y^e\right)\mathbb{I}\left(t \leq T_{\mathbf{x}}^0\right) f_t f_t\right]^{-1}\left[\sum_{t=1}^{T-1} \mathbb{I}\left(t > T_y^e\right)\mathbb{I}\left(t \leq T_{\mathbf{x}}^0\right)\mathbb{I}\left(t \leq T_y^0\right) f_t f_t\right] \\
& + \mathbb{I}\left(T_{\mathbf{x}}^0 > T_y^e\right)\frac{1}{\bar{\lambda}_{\mathbf{w}1}}\frac{T_{\mathbf{x}}^0 - T_y^e}{T - \left(T_y^e + 1\right)}\left[\sum_{t=1}^{T-1} \mathbb{I}\left(t > T_y^e\right)\mathbb{I}\left(t \leq T_{\mathbf{x}}^0\right) f_t f_t\right]^{-1}\left[\sum_{t=1}^{T-1} \mathbb{I}\left(t > T_y^e\right)\mathbb{I}\left(t \leq T_{\mathbf{x}}^0\right) f_t \varepsilon_t\right] \\
& + \mathbb{I}\left(T_{\mathbf{x}}^0 > T_y^e\right)\frac{\gamma_2}{\bar{\lambda}_{\mathbf{w}2}}\frac{T - \left(T_{\mathbf{x}}^0 + 1\right)}{T - \left(T_y^e + 1\right)} \\
& + \mathbb{I}\left(T_{\mathbf{x}}^0 > T_y^e\right)\frac{\gamma_1 - \gamma_2}{\bar{\lambda}_{\mathbf{w}1}}\frac{T - \left(T_{\mathbf{x}}^0 + 1\right)}{T - \left(T_y^e + 1\right)}\left[\mathbb{I}\left(t > T_{\mathbf{x}}^0\right) f_t f_t\right]^{-1}\left[\sum_{t=1}^{T-1} \mathbb{I}\left(t > T_{\mathbf{x}}^0\right)\mathbb{I}\left(t \leq T_y^0\right) f_t f_t\right] \\
& + \mathbb{I}\left(T_{\mathbf{x}}^0 > T_y^e\right)\frac{1}{\bar{\lambda}_{\mathbf{w}2}}\frac{T - \left(T_{\mathbf{x}}^0 + 1\right)}{T - \left(T_y^e + 1\right)}\left[\sum_{t=1}^{T-1} \mathbb{I}\left(t > T_{\mathbf{x}}^0\right) f_t f_t\right]^{-1}\left[\sum_{t=1}^{T-1} \mathbb{I}\left(t > T_{\mathbf{x}}^0\right) f_t \varepsilon_t\right].
\end{aligned}
$$

Simplifying terms, we get that as $N \to \infty$

$$
\begin{aligned}
& \hat{\gamma}_2\left(T_y^e\right) \\
\xrightarrow{p}\; & \frac{\gamma_2}{\bar{\lambda}_{\mathbf{w}2}}\frac{T - \left(\max\left\{T_{\mathbf{x}}^0, T_y^e\right\} + 1\right)}{T - \left(T_y^e + 1\right)} + \mathbb{I}\left(T_{\mathbf{x}}^0 > T_y^e\right)\frac{\gamma_2}{\bar{\lambda}_{\mathbf{w}1}}\frac{T_{\mathbf{x}}^0 - T_y^e}{T - \left(T_y^e + 1\right)} \\
& + \frac{\gamma_1 - \gamma_2}{\bar{\lambda}_{\mathbf{w}2}}\frac{T - \left(\max\left\{T_{\mathbf{x}}^0, T_y^e\right\} + 1\right)}{T - \left(T_y^e + 1\right)}\left[\sum_{t=1}^{T-1} \mathbb{I}\left(t > \max\left\{T_{\mathbf{x}}^0, T_y^e\right\}\right) f_t f_t\right]^{-1}\left[\sum_{t=1}^{T-1} \mathbb{I}\left(t > \max\left\{T_{\mathbf{x}}^0, T_y^e\right\}\right)\mathbb{I}\left(t \leq T_y^0\right) f_t f_t\right] \\
& + \mathbb{I}\left(T_{\mathbf{x}}^0 > T_y^e\right)\frac{\gamma_1 - \gamma_2}{\bar{\lambda}_{\mathbf{w}1}}\frac{T_{\mathbf{x}}^0 - T_y^e}{T - \left(T_y^e + 1\right)}\left[\sum_{t=1}^{T-1} \mathbb{I}\left(t > T_y^e\right)\mathbb{I}\left(t \leq T_{\mathbf{x}}^0\right) f_t f_t\right]^{-1}\left[\sum_{t=1}^{T-1} \mathbb{I}\left(t > T_y^e\right)\mathbb{I}\left(t \leq \min\left\{T_{\mathbf{x}}^0, T_y^0\right\}\right) f_t f_t\right] \\
& + \frac{1}{\bar{\lambda}_{\mathbf{w}2}}\frac{T - \left(\max\left\{T_{\mathbf{x}}^0, T_y^e\right\} + 1\right)}{T - \left(T_y^e + 1\right)}\left[\sum_{t=1}^{T-1} \mathbb{I}\left(t > \max\left\{T_{\mathbf{x}}^0, T_y^e\right\}\right) f_t f_t\right]^{-1}\left[\sum_{t=1}^{T-1} \mathbb{I}\left(t > \max\left\{T_{\mathbf{x}}^0, T_y^e\right\}\right) f_t \varepsilon_{t+1}\right] \\
& + \mathbb{I}\left(T_{\mathbf{x}}^0 > T_y^e\right)\frac{1}{\bar{\lambda}_{\mathbf{w}1}}\frac{T_{\mathbf{x}}^0 - T_y^e}{T - \left(T_y^e + 1\right)}\left[\sum_{t=1}^{T-1} \mathbb{I}\left(t > T_y^e\right)\mathbb{I}\left(t \leq T_{\mathbf{x}}^0\right) f_t f_t\right]^{-1}\left[\sum_{t=1}^{T-1} \mathbb{I}\left(t > T_y^e\right)\mathbb{I}\left(t \leq T_{\mathbf{x}}^0\right) f_t \varepsilon_{t+1}\right].
\end{aligned}
$$

It follows that

$$
\begin{aligned}
& \mathrm{E}\left[\hat{\gamma}_2\left(T_y^e\right)\right] \\
\to\; & \frac{\gamma_2}{\bar{\lambda}_{\mathbf{w}2}}\frac{T - \left(\max\left\{T_{\mathbf{x}}^0, T_y^e\right\} + 1\right)}{T - \left(T_y^e + 1\right)} + \mathbb{I}\left(T_{\mathbf{x}}^0 > T_y^e\right)\frac{\gamma_2}{\bar{\lambda}_{\mathbf{w}1}}\frac{T_{\mathbf{x}}^0 - T_y^e}{T - \left(T_y^e + 1\right)} \\
& + \frac{\gamma_1 - \gamma_2}{\bar{\lambda}_{\mathbf{w}2}}\frac{T - \left(\max\left\{T_{\mathbf{x}}^0, T_y^e\right\} + 1\right)}{T - \left(T_y^e + 1\right)}\mathrm{E}\left\{\left[\sum_{t=1}^{T-1} \mathbb{I}\left(t > \max\left\{T_{\mathbf{x}}^0, T_y^e\right\}\right) f_t f_t\right]^{-1}\left[\sum_{t=1}^{T-1} \mathbb{I}\left(t > \max\left\{T_{\mathbf{x}}^0, T_y^e\right\}\right)\mathbb{I}\left(t \leq T_y^0\right) f_t f_t\right]\right\} \\
& + \mathbb{I}\left(T_{\mathbf{x}}^0 > T_y^e\right)\frac{\gamma_1 - \gamma_2}{\bar{\lambda}_{\mathbf{w}1}}\frac{T_{\mathbf{x}}^0 - T_y^e}{T - \left(T_y^e + 1\right)}\mathrm{E}\left\{\left[\sum_{t=1}^{T-1} \mathbb{I}\left(t > T_y^e\right)\mathbb{I}\left(t \leq T_{\mathbf{x}}^0\right) f_t f_t\right]^{-1}\left[\sum_{t=1}^{T-1} \mathbb{I}\left(t > T_y^e\right)\mathbb{I}\left(t \leq \min\left\{T_{\mathbf{x}}^0, T_y^0\right\}\right) f_t f_t\right]\right\}.
\end{aligned}
$$

For $1 \leq T_{\mathbf{x}}^{0} \leq T_{y}^{e}$, using the same results employed on pp. $421-422$ in the Proof of Proposition 1 of Pesaran and Timmermann (2004),

$$
\begin{aligned}
& \lim_{N\to\infty} \mathrm{E}\left[\hat{\gamma}_{2}\left(T_{y}^{e}\right)\right] \\
= & \frac{\gamma_{2}}{\bar{\lambda}_{\mathbf{w}2}} \frac{T-\left(T_{y}^{e}+1\right)}{T-\left(T_{y}^{e}+1\right)} + \frac{\gamma_{1}-\gamma_{2}}{\bar{\lambda}_{\mathbf{w}2}} \frac{T-1-\left(T_{y}^{e}+1\right)-1}{T-1-\left(T_{y}^{e}+1\right)-1} \mathrm{E}\left\{\left[\sum_{t=1}^{T-1} \mathbb{I}\left(t>T_{y}^{e}\right) f_{t} f_{t}\right]^{-1}\left[\sum_{t=1}^{T-1} \mathbb{I}\left(t>T_{y}^{e}\right) \mathbb{I}\left(t \leq T_{y}^{0}\right) f_{t} f_{t}\right]\right\} \\
= & \frac{1}{\bar{\lambda}_{\mathbf{w}2}}\left[\gamma_{2}+\left(\gamma_{1}-\gamma_{2}\right) \frac{T_{y}^{0}-T_{y}^{e}}{T-\left(T_{y}^{e}+1\right)}\right],
\end{aligned}
$$

which gives (4). Following analogous arguments, for $T_{y}^{e}<T_{\mathbf{x}}^{0}<T_{y}^{0}$,

$$
\begin{aligned}
& \lim_{N\to\infty} \mathrm{E}\left[\hat{\gamma}_{2}\left(T_{y}^{e}\right)\right] \\
= & \left[\frac{\gamma_{2}}{\bar{\lambda}_{\mathbf{w}2}} \frac{T-\left(T_{\mathbf{x}}^{0}+1\right)}{T-\left(T_{y}^{e}+1\right)}+\frac{\gamma_{2}}{\bar{\lambda}_{\mathbf{w}1}} \frac{T_{\mathbf{x}}^{0}-T_{y}^{e}}{T-\left(T_{y}^{e}+1\right)}\right] \\
& +\left[\frac{\gamma_{1}-\gamma_{2}}{\bar{\lambda}_{\mathbf{w}2}} \frac{T-\left(T_{\mathbf{x}}^{0}+1\right)}{T-\left(T_{y}^{e}+1\right)} \frac{T_{y}^{0}-T_{\mathbf{x}}^{0}}{T-\left(T_{\mathbf{x}}^{0}+1\right)}+\frac{\gamma_{1}-\gamma_{2}}{\bar{\lambda}_{\mathbf{w}1}} \frac{T_{\mathbf{x}}^{0}-T_{y}^{e}}{T-\left(T_{y}^{e}+1\right)}\right] \\
= & \frac{\gamma_{2}}{\bar{\lambda}_{\mathbf{w}2}}\left[\frac{T-\left(T_{\mathbf{x}}^{0}+1\right)}{T-\left(T_{y}^{e}+1\right)}+\frac{T_{\mathbf{x}}^{0}-T_{y}^{e}}{T-\left(T_{y}^{e}+1\right)} \frac{\bar{\lambda}_{\mathbf{w}2}}{\bar{\lambda}_{\mathbf{w}1}}\right]+\frac{\gamma_{1}-\gamma_{2}}{\bar{\lambda}_{\mathbf{w}2}}\left[\frac{T_{y}^{0}-T_{\mathbf{x}}^{0}}{T-\left(T_{y}^{e}+1\right)}+\frac{T_{\mathbf{x}}^{0}-T_{y}^{e}}{T-\left(T_{y}^{e}+1\right)} \frac{\bar{\lambda}_{\mathbf{w}2}}{\bar{\lambda}_{\mathbf{w}1}}\right],
\end{aligned}
$$

which is equal to (6). Finally, for $T_{y}^{0} \leq T_{\mathbf{x}}^{0} \leq T-1$,

$$
\begin{aligned}
& \lim_{N\to\infty} \mathrm{E}\left[\hat{\gamma}_{2}\left(T_{y}^{e}\right)\right] \\
= & \frac{\gamma_{2}}{\bar{\lambda}_{\mathbf{w}2}}\left[\frac{T-\left(T_{\mathbf{x}}^{0}+1\right)}{T-\left(T_{y}^{e}+1\right)}+\frac{T_{\mathbf{x}}^{0}-T_{y}^{e}}{T-\left(T_{y}^{e}+1\right)} \frac{\bar{\lambda}_{\mathbf{w}2}}{\bar{\lambda}_{\mathbf{w}1}}\right]+\frac{\gamma_{1}-\gamma_{2}}{\bar{\lambda}_{1\mathbf{w}}} \frac{T_{\mathbf{x}}^{0}-T_{y}^{e}}{T-\left(T_{y}^{e}+1\right)} \frac{T_{y}^{0}-\left(T_{y}^{e}+1\right)+1}{T_{\mathbf{x}}^{0}-\left(T_{y}^{e}+1\right)+1} \\
= & \frac{\gamma_{2}}{\bar{\lambda}_{\mathbf{w}2}}\left[\frac{T-\left(T_{\mathbf{x}}^{0}+1\right)}{T-\left(T_{y}^{e}+1\right)}+\frac{T_{\mathbf{x}}^{0}-T_{y}^{e}}{T-\left(T_{y}^{e}+1\right)} \frac{\bar{\lambda}_{\mathbf{w}2}}{\bar{\lambda}_{\mathbf{w}1}}\right]+\frac{\gamma_{1}-\gamma_{2}}{\bar{\lambda}_{\mathbf{w}2}}\left[\frac{T_{y}^{0}-T_{y}^{e}}{T-\left(T_{y}^{e}+1\right)} \frac{\bar{\lambda}_{\mathbf{w}2}}{\bar{\lambda}_{\mathbf{w}1}}\right],
\end{aligned}
$$

which gives (7). The result stated in Proposition 3.1 then follows. ■

Proof of Proposition 3.2. Consider

$$
\begin{aligned}
\lim_{N\to\infty} \mathrm{E}\left[y_{T+1}, \hat{y}_{T+1}\left(T_{y}^{e}\right)\right] & = \lim_{N\to\infty} \mathrm{E}\left[\left(\gamma_{2} f_{T}+\varepsilon_{T+1}\right) \hat{\gamma}_{2}\left(T_{y}^{e}\right) \hat{f}_{T}\right] \\
& = \lim_{N\to\infty} \mathrm{E}\left[\gamma_{2} f_{T} \hat{f}_{T} \hat{\gamma}_{2}\left(T_{y}^{e}\right)\right] \\
& = \lim_{N\to\infty} \mathrm{E}\left\{\gamma_{2} f_{T}\left[\sum_{i=1}^{N} w_{i}\left(\lambda_{2i} f_{T}+e_{it}\right)\right] \hat{\gamma}_{2}\left(T_{y}^{e}\right)\right\} \\
& = \gamma_{2} \mathrm{E}\left(f_{T}^{2}\right) \lim_{N\to\infty}\left(\sum_{i=1}^{N} w_{i} \lambda_{2i}\right) \lim_{N\to\infty} \mathrm{E}\left[\hat{\gamma}_{2}\left(T_{y}^{e}\right)\right] \\
& = \gamma_{2} \sigma_{f}^{2} \bar{\lambda}_{\mathbf{w}2} \lim_{N\to\infty} \mathrm{E}\left[\hat{\gamma}_{2}\left(T_{y}^{e}\right)\right],
\end{aligned}
$$

which completes the proof of Proposition 3.2. ■

References

[1] Bai, J. (2003), "Inferential Theory for Factor Models of Large Dimensions," *Econometrica*, 71 (1), 135 − 171.

[2] Bai, J., and P. Perron (1998), "Estimating and Testing Linear Models with Multiple Structural Changes," *Econometrica*, 66 (1), 47 − 78.

[3] Bai, J., and S. Ng (2002), "Determining the Number of Factors in Approximate Factor Models," *Econometrica*, 70 (1), 191 − 221.

[4] Bai, J., and S. Ng (2006), "Confidence Intervals for Diffusion Index Forecasts and Inference for Factor-Augmented Regressions," *Econometrica*, 74 (4), 1133 – 1150.

[5] Bair, E., T. Hastie, D. Paul, and R. Tibshirani (2006), "Prediction by Supervised Principal Components," *Journal of the American Statistical Association*, 101 (473), 119 – 137.

[6] Baetje, F., and L. Menkhoff (2016), "Equity Premium Prediction: Are Economic and Technical Indicators Unstable?", *International Journal of Forecasting*, 32 (4), 1193 – 1207.

[7] Çakmakli, C., and D. van Dijk, (2016), "Getting the Most out of Macroeconomic Information for Predicting Excess Stock Returns", *International Journal of Forecasting*, 32 (3), 650 – 668.

[8] Corradi, V., and N. R. Swanson (2014), "Testing for Structural Stability of Factor Augmented Forecasting models," *Journal of Econometrics*, 182 (1), 100 – 118.

[9] Farmer, L., L. Schmidt, and A. Timmermann (2021), "Pockets of Predictability", Working Paper, University of Virginia, MIT and UCSD.

[10] Forni, M., A. Giovannelli, M. Lippi, and S. Soccorsi (2018), "Dynamic Factor Model with Infinite Dimensional Factor Space: Forecasting," *Journal of Applied Econometrics*, 33 (5), 625 – 642.

[11] Giacomini, R., and H. White (2006), "Tests of Conditional Predictive Ability," *Econometrica*, 74 (6), 1545 – 1578.

[12] Giovannelli, A., D. Massacci, and S. Soccorsi (2021), "Forecasting Stock Returns with Large Dimensional Factor Models," *Journal of Empirical Finance*, 63, 252 – 269.

[13] Gonçalves, S., M. W. McCracken, and B. Perron (2017), "Tests of Equal Accuracy for Nested Models with Estimated Factors", *Journal of Econometrics*, 198 (2), 231 – 252.

[14] Ludvigson, S. C., and S. Ng (2007), "The Empirical Risk-Return Relation: A Factor Analysis Approach", *Journal of Financial Economics*, 83 (1), 171 – 222.

[15] Ludvigson, S. C., and S. Ng (2009), "Macro Factors in Bond Risk Premia", *Review of Financial Studies*, 22 (12), 5027 – 5067.

[16] Massacci, D. (2019), "Unstable Diffusion Indexes: With an Application to Bond Risk Premia," *Oxford Bulletin of Economics and Statistics*, 81 (6), 1376 – 1400.

[17] Neely, C. J., D. E. Rapach, J. Tu, and G. Zhou (2014), "Forecasting the Equity Risk Premium: The Role of Technical Indicators," *Management Science*, 60 (7), 1772 – 1791.

[18] Pástor, L., and R. F. Stambaugh (2001), "The Equity Premium and Structural Breaks," *Journal of Finance*, 56 (4), 1207 – 1239.

[19] Paye, B. S., and A. Timmermann (2006), "Instability of Return Prediction Models," *Journal of Empirical Finance*, 13 (3), 274 – 315.

[20] Pesaran, M. H. (2006), "Estimation and Inference in Large Heterogeneous Panels with a Multifactor Error Structure," *Econometrica*, 74 (4), 967 – 1012.

[21] Pesaran, M. H., and A. Timmermann (2004), "How Costly is to Ignore Breaks when Forecasting the Direction of a Time Series?," *International Journal of Forecasting*, 20 (3), 411 – 425.

[22] Pesaran, M. H., and A. Timmermann (2007), "Selection of Estimation Window in the Presence of Breaks", *Journal of Econometrics*, 137 (1), 134 – 161.

[23] Rapach, D. E., and M. E. Wohar (2006), "Structural Breaks and Predictive Regression Models of Aggregate U.S. Stock Returns," *Journal of Financial Econometrics*, 4 (2), 238 – 274.

[24] Rossi, B. (2021), "Forecasting in the Presence of Instabilities: How Do We Know Whether Economic Models Work and How to Improve Them," *Journal of Economic Literature*, Forthcoming.

[25] Stock, J. H., and M. W. Watson (1996), "Evidence on Structural Instability in Macroeconomic Time Series Relations," *Journal of Business and Economic Statistics*, 14 (1), 11 – 30.

[26] Stock, J. H., and M. W. Watson (2002a), "Forecasting Using Principal Components from a Large Number of Predictors," *Journal of the American Statistical Association*, 97 (460), 1167 – 1179.

[27] Stock, J. H., and M. W. Watson (2002b), "Macroeconomic Forecasting Using Diffusion Indexes", *Journal of Business and Economic Statistics*, 20 (2), 147 – 162.

[28] Timmermann, A. (2001), "Structural Breaks, Incomplete Information, and Stock Prices," *Journal of Business and Economic Statistics*, 19 (3), 299 – 314.

[29] Timmermann, A. (2008), "Elusive Return Predictability," *International Journal of Forecasting*, 24 (1), 1 – 18.

[30] Timmermann, A. (2018), "Forecasting Methods in Finance," *Annual Review of Financial Economics*, 10, 449 – 479.

[31] Welch, I., and A. Goyal (2008), "A Comprehensive Look at the Empirical Performance of Equity Premium Prediction," *Review of Financial Studies*, 21 (4), 1455 – 1508.

Table 1: **Experiment 1, RMSFE, $\bar{\lambda}_{\mathbf{w}2}/\bar{\lambda}_{\mathbf{w}1} = 1$**

Panel A: $T_y^0 = 100$							
		$T_{\mathbf{x}}^0 = 100$			$T_{\mathbf{x}}^0 = 190$		
		Post-break	Expanding	Rolling	Post-break	Expanding	Rolling
δ_y	$\lvert\delta_{\mathbf{x}}\rvert$						
0.00	0.00	1.0614	1.0560	1.0796	1.0614	1.0560	1.0796
	1.50	1.0614	1.0560	1.0796	1.0614	1.0560	1.0796
	3.00	1.0614	1.0560	1.0796	1.0614	1.0560	1.0796
1.00	0.00	1.0614	1.2858	1.0796	1.0614	1.2858	1.0796
	1.50	1.0614	1.2858	1.0796	1.0614	1.2858	1.0796
	3.00	1.0614	1.2858	1.0796	1.0614	1.2858	1.0796
2.00	0.00	1.0614	2.0268	1.0796	1.0614	2.0268	1.0796
	1.50	1.0614	2.0268	1.0796	1.0614	2.0268	1.0796
	3.00	1.0614	2.0268	1.0796	1.0614	2.0268	1.0796
3.00	0.00	1.0614	3.2791	1.0796	1.0614	3.2791	1.0796
	1.50	1.0614	3.2791	1.0796	1.0614	3.2791	1.0796
	3.00	1.0614	3.2791	1.0796	1.0614	3.2791	1.0796

Panel B: $T_y^0 = 190$							
		$T_{\mathbf{x}}^0 = 100$			$T_{\mathbf{x}}^0 = 190$		
		Post-break	Expanding	Rolling	Post-break	Expanding	Rolling
δ_y	$\lvert\delta_{\mathbf{x}}\rvert$						
0.00	0.00	1.2193	1.0560	1.0796	1.2193	1.0560	1.0796
	1.50	1.2193	1.0560	1.0796	1.2193	1.0560	1.0796
	3.00	1.2193	1.0560	1.0796	1.2193	1.0560	1.0796
1.00	0.00	1.2193	1.9312	1.7342	1.2193	1.9312	1.7342
	1.50	1.2193	1.9312	1.7342	1.2193	1.9312	1.7342
	3.00	1.2193	1.9312	1.7342	1.2193	1.9312	1.7342
2.00	0.00	1.2193	4.6567	3.7817	1.2193	4.6567	3.7817
	1.50	1.2193	4.6567	3.7817	1.2193	4.6567	3.7817
	3.00	1.2193	4.6567	3.7817	1.2193	4.6567	3.7817
3.00	0.00	1.2193	9.2327	7.2220	1.2193	9.2327	7.2220
	1.50	1.2193	9.2327	7.2220	1.2193	9.2327	7.2220
	3.00	1.2193	9.2327	7.2220	1.2193	9.2327	7.2220

This table displays the RMSFE as defined in (8) for Experiment 1, whose data generating process and results are described in Section 4.1 and in Section 4.2, respectively.

Table 2a: Experiment 2a, RMSFE, $0 < \bar{\lambda}_{\mathbf{w}2} / \bar{\lambda}_{\mathbf{w}1} < 1$, $\delta_{\mathbf{x}} = 0.00, 1.50, 3.00$

Panel A: $T_y^0 = 100$

		$T_{\mathbf{x}}^0 = 100$			$T_{\mathbf{x}}^0 = 190$		
		Post-break	Expanding	Rolling	Post-break	Expanding	Rolling
δ_y	$\delta_{\mathbf{x}}$						
	0.00	1.0614	1.0560	1.0796	1.0614	1.0560	1.0796
0.00	1.50	1.0614	1.3194	1.0796	1.4061	1.4108	1.3963
	3.00	1.0614	1.5550	1.0796	1.6190	1.6222	1.6111
	0.00	1.0614	1.2858	1.0796	1.0614	1.2858	1.0796
1.00	1.50	1.0614	1.0746	1.0796	1.4061	1.1957	1.3963
	3.00	1.0614	1.2625	1.0796	1.6190	1.4313	1.6111
	0.00	1.0614	2.0268	1.0796	1.0614	2.0268	1.0796
2.00	1.50	1.0614	1.0857	1.0796	1.4061	1.0766	1.3963
	3.00	1.0614	1.0930	1.0796	1.6190	1.2790	1.6111
	0.00	1.0614	3.2791	1.0796	1.0614	3.2791	1.0796
3.00	1.50	1.0614	1.3526	1.0796	1.4061	1.0536	1.3963
	3.00	1.0614	1.0467	1.0796	1.6190	1.1655	1.6111

Panel B: $T_y^0 = 190$

		$T_{\mathbf{x}}^0 = 100$			$T_{\mathbf{x}}^0 = 190$		
		Post-break	Expanding	Rolling	Post-break	Expanding	Rolling
δ_y	$\delta_{\mathbf{x}}$						
	0.00	1.2193	1.0560	1.0796	1.2193	1.0560	1.0796
0.00	1.50	1.2193	1.3194	1.0796	1.2193	1.4108	1.3963
	3.00	1.2193	1.5550	1.0796	1.2193	1.6222	1.6111
	0.00	1.2193	1.9312	1.7342	1.2193	1.9312	1.7342
1.00	1.50	1.2193	1.0492	1.7342	1.2193	1.0814	1.0860
	3.00	1.2193	1.2102	1.7342	1.2193	1.2901	1.2871
	0.00	1.2193	4.6567	3.7817	1.2193	4.6567	3.7817
2.00	1.50	1.2193	1.2689	3.7817	1.2193	1.0998	1.1040
	3.00	1.2193	1.0556	3.7817	1.2193	1.1883	1.1886
	0.00	1.2193	9.2327	7.2220	1.2193	9.2327	7.2220
3.00	1.50	1.2193	1.9784	7.2220	1.2193	1.4660	1.4504
	3.00	1.2193	1.0913	7.2220	1.2193	1.0461	1.0497

This table displays the RMSFE as defined in (8) for Experiment 2a, whose data generating process and results are described in Section 4.1 and in Section 4.2, respectively.

Table 2b: Experiment 2b, RMSFE, $0 < \bar{\lambda}_{\mathbf{w}2} / \bar{\lambda}_{\mathbf{w}1} < 1$, $\delta_{\mathbf{x}} = 0.00, -1.50, -3.00$

Panel A: $T_y^0 = 100$

		$T_{\mathbf{x}}^0 = 100$			$T_{\mathbf{x}}^0 = 190$		
		Post-break	Expanding	Rolling	Post-break	Expanding	Rolling
δ_y	$\delta_{\mathbf{x}}$						
0.00	0.00	1.0614	1.0560	1.0796	1.0614	1.0560	1.0796
	-1.50	1.0614	1.3413	1.0796	6.1589	8.4087	3.9456
	-3.00	1.0614	2.5212	1.0796	3.3949	3.4676	3.2516
1.00	0.00	1.0614	1.2858	1.0796	1.0614	1.2858	1.0796
	-1.50	1.0614	1.8546	1.0796	6.1589	14.7800	3.9456
	-3.00	1.0614	3.7394	1.0796	3.3949	4.4453	3.2516
2.00	0.00	1.0614	2.0268	1.0796	1.0614	2.0268	1.0796
	-1.50	1.0614	2.6258	1.0796	6.1589	23.1370	3.9456
	-3.00	1.0614	5.3211	1.0796	3.3949	5.5895	3.2516
3.00	0.00	1.0614	3.2791	1.0796	1.0614	3.2791	1.0796
	-1.50	1.0614	3.6550	1.0796	6.1589	33.4810	3.9456
	-3.00	1.0614	7.2664	1.0796	3.3949	6.9002	3.2516

Panel B: $T_y^0 = 190$

		$T_{\mathbf{x}}^0 = 100$			$T_{\mathbf{x}}^0 = 190$		
		Post-break	Expanding	Rolling	Post-break	Expanding	Rolling
δ_y	$\delta_{\mathbf{x}}$						
0.00	0.00	1.2193	1.0560	1.0796	1.2193	1.0560	1.0796
	-1.50	1.2193	1.3413	1.0796	1.2193	8.4087	3.9456
	-3.00	1.2193	2.5212	1.0796	1.2193	3.4676	3.2516
1.00	0.00	1.2193	1.9312	1.7342	1.2193	1.9312	1.7342
	-1.50	1.2193	1.0960	1.7342	1.2193	22.308	9.2407
	-3.00	1.2193	3.0835	1.7342	1.2193	5.4797	5.0765
2.00	0.00	1.2193	4.6567	3.7817	1.2193	4.6567	3.7817
	-1.50	1.2193	1.2314	3.7817	1.2193	43.416	17.260
	-3.00	1.2193	3.7398	3.7817	1.2193	8.0949	7.4469
3.00	0.00	1.2193	9.2327	7.2220	1.2193	9.2327	7.2220
	-1.50	1.2193	1.7475	7.2220	1.2193	71.735	28.003
	-3.00	1.2193	4.4899	7.2220	1.2193	11.313	10.363

This table displays the RMSFE as defined in (8) for Experiment 2b, whose data generating process and results are described in Section 4.1 and in Section 4.2, respectively.