ARTICLE IN PRESS

International Journal of Forecasting xxx (xxxx) xxx

Contents lists available at ScienceDirect

International Journal of Forecasting

journal homepage: www.elsevier.com/locate/ijforecast

Forecasting in factor augmented regressions under structural change[☆]

Daniele Massacci[a,b,c,*], George Kapetanios[a]

[a] King's College London, United Kingdom
[b] University of Naples Federico II, Italy
[c] CSEF, Italy

ARTICLE INFO

Keywords:
Factor augmented regression
Structural instability
Out-of-sample forecasts
Estimation window
Cross-sectional averages

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 instability in the factor model *and* in the forecasting equation impacts 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. However, the actual gain depends on the location and the magnitude of the breaks. Finally, an application to out-of-sample stock return forecasting using liquidity proxies illustrates the empirical relevance of our results.

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 contributions above work under the maintained assumption that 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

☆ Insightful comments from Esther Ruiz (the Editor), the Associate Editor, and one anonymous reviewer are gratefully acknowledged. Errors and omissions are the authors' responsibility.
* Correspondence to: King's Business School, Bush House, 30 Aldwych, London, WC2B 4BG, United Kingdom.
E-mail addresses: daniele.massacci@kcl.ac.uk (D. Massacci), george.kapetanios@kcl.ac.uk (G. Kapetanios).

https://doi.org/10.1016/j.ijforecast.2022.12.004
0169-2070/

Please cite this article as: D. Massacci and G. Kapetanios, Forecasting in factor augmented regressions under structural change. International Journal of Forecasting (2023), https://doi.org/10.1016/j.ijforecast.2022.12.004.

(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. (2022).

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.

Estimation and inference in factor augmented regressions in which either the factor model or the forecasting equation (or both) are subject to breaks have been studied within an *in-sample* framework. Corradi and Swanson (2014) develop a test for the joint hypothesis of structural stability of both factor loadings, and factor augmented forecasting model regression coefficients. Massacci (2019) considers estimating the factor augmented regression under structural instability and proposes a break test. However, to our knowledge, the literature is silent regarding the consequence of structural instability on the *out-of-sample* forecasting performance of factor augmented regressions. This is an empirically relevant problem. For example, Chen et al. (2018) argue that aggregate liquidity conditions provide valuable information for out-of-sample stock return forecasting. The aggregate measures they consider are constructed from firm-level counterparts by taking cross-sectional averages under the assumption of an underlying factor structure. This implies that stock return forecasting with liquidity proxies can be implemented using factor augmented regressions. However, liquidity measures have been affected by the minimum tick-size reduction which took place in the New York Stock Exchange in 1997 and 2001. These tick-size reductions may have induced a structural break in the factor model, or the factor augmented regression (or both). This is exactly the problem we are addressing in this paper.

We fill 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's 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 setup of a factor augmented regression with a single latent factor and one break in the factor model and 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. We show how this covariance depends on the choice of the estimation window in relation to the location of the breaks: when the estimation window begins after the break in the factor model, this does not affect the forecast; 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.

Finally, we illustrate the relevance of our results by studying the out-of-sample stock return forecasting ability of the liquidity measures considered in Chen et al. (2018) after the break induced by the tick-size reduction that occurred in 2001. Our empirical findings show that, on this occasion, the expanding window estimation scheme generally produces the most accurate predictions. However, if the forecasts are produced shortly after the break and for a limited time, then the post-break estimation scheme tends to be preferable.

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. Section 5 studies the predictive power of aggregate illiquidity for stock returns under structural instability. Section 6 concludes. Mathematical proofs are provided in Appendix.

2. Setup

We consider the model

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

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

where $\mathbb{I}(\cdot)$ is the indicator function and T denotes the time series dimension. Starting from (1), $\mathbf{x}_t = (x_{1t}, \ldots, x_{Nt})' \in \mathfrak{R}^N$ is the $N \times 1$ vector of observable dependent variables; f_t is the latent factor such that $\mathrm{E}(f_t) = 0$; $\mathbf{e}_t = (e_{1t}, \ldots, e_{Nt})' \in \mathfrak{R}^N$ is the $N \times 1$ vector of idiosyncratic components; $\boldsymbol{\lambda}_{j_{\mathbf{x}}} = \left(\lambda_{j_{\mathbf{x}}1}, \ldots, \lambda_{j_{\mathbf{x}}N}\right)'$ is the $N \times 1$ vector of factor loadings in state $j_{\mathbf{x}} = 1, 2$, whose ith 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), $y_{t+1} \in \mathfrak{R}$ 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). We aim 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. This means we study the case where 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. Then it imputes the estimated factors into the forecasting model: see Bai . 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. 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 requires $N \to \infty$ 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 $1 \le T_y^e \le T_y^0$: 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. Section 3.3 provides a discussion of factor loadings, which play a role in the results in Sections 3.1 and 3.2. Section 3.4 details the rotational indeterminacy problem. Section 3.5 discusses the consequences of factor heteroskedasticity.

3.1. Cross-sectional average estimation

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

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

respectively, where $\{w_i\}_{i=1}^N$ is a sequence of weights. Let $\mathrm{diag}(\cdot)$ denote a diagonal matrix of suitable dimension and $\mathbf{I}_N$ be the $N \times N$ identity matrix. 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 $\mathbf{e}_t \sim \mathrm{IID}\left(\mathbf{0}, \sigma_{\mathbf{e}}^2 \mathbf{I}_N\right)$ and $(f_t, \varepsilon_{t+1})' \sim \mathrm{IID}\mathcal{N}\left[\mathbf{0}, \mathrm{diag}\left(\sigma_f^2, \sigma_\varepsilon^2\right)\right]$. 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*

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

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. To interpret Proposition 3.1, we consider three mutually exclusive cases: (*i*) $1 \le T_{\mathbf{x}}^0 \le T_y^e$; (*ii*) $T_y^e < T_{\mathbf{x}}^0 \le T-1$ and $T_{\mathbf{x}}^0 \le T_y^0$; (*iii*) $T_y^e < T_{\mathbf{x}}^0 \le T-1$ and $T_{\mathbf{x}}^0 > T_y^0$.

If $1 \le T_{\mathbf{x}}^0 \le T_y^e$ the break in the factor model occurs before the beginning of the estimation window in the forecasting equation and

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

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 $\bar{\lambda}_{\mathbf{w}2}^{-1}$ (namely, $\bar{\lambda}_{\mathbf{w}2}^{-1}\gamma_2$). For $\gamma_1 \neq \gamma_2$, the asymptotic bias of $\hat{\gamma}_2\left(T_y^e\right)$ 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 $\left(T_y^0 - T_y^e\right)$ and the size of the estimation window $\left[T - \left(T_y^e + 1\right)\right]$. Notice that if f_t was observable and did not have to be estimated, then $\hat{\gamma}_2\left(T_y^e\right)$ and $\lim_{N\to\infty} \mathrm{E}\left[\hat{\gamma}_2\left(T_y^e\right)\right]$ in (3) and (4), respectively, would simplify to

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

and

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

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

If $T_y^e < T_{\mathbf{x}}^0 \le T-1$ and $T_{\mathbf{x}}^0 \le T_y^0$ then $T_y^e < T_{\mathbf{x}}^0 \le T_y^0 \le T-1$: 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

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

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

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

and the rotation induced around γ_2 depends on the frequency of observations of $\mathbf{x}_t$ before and after the break date $T_{\mathbf{x}}^0$, as given by $\left[T - \left(T_{\mathbf{x}}^0 + 1\right)\right] / \left[T - \left(T_y^e + 1\right)\right]$ and $\left(T_{\mathbf{x}}^0 - T_y^e\right) / \left[T - \left(T_y^e + 1\right)\right]$, 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 $\lim_{N\to\infty} \mathrm{E}\left[\hat{\gamma}_2\left(T_y^e\right)\right] = \gamma_2 / \bar{\lambda}_{\mathbf{w}2}$, which is the same as in the case with 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

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

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

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

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 $\bar{\lambda}_{\mathbf{w}2}/\bar{\lambda}_{\mathbf{w}1}$ persists.

[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

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

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.

3.3. Factor loadings

In Sections 3.1 and 3.2 we analytically showed that the condition $\bar{\lambda}_{\mathbf{w}2} \neq \bar{\lambda}_{\mathbf{w}1}$ plays a role in determining the bias of the estimator for γ_2, and the covariance between y_{T+1} and its forecast, respectively, as $N \to \infty$. It is, therefore, informative to discuss how relevant the condition $\bar{\lambda}_{\mathbf{w}2} \neq \bar{\lambda}_{\mathbf{w}1}$ is in practice.

To this purpose, we consider the problem of out-of-sample stock return forecasting with aggregate liquidity proxies already mentioned in the introductory Section 1. Following Chen et al. (2018), let us focus on the minimum tick-size reduction that took place on January 29, 2001. Assuming monthly data, it is legitimate to conjecture that from February 2002 onward, this tick-size reduction determined a sizeable *reduction* in the response of firm-level *illiquidity* measures with respect to the underlying latent factor. In terms of the factor model in (1), it is likely that $\bar{\lambda}_{\mathbf{w}2} < \bar{\lambda}_{\mathbf{w}1}$ and more generally that $\bar{\lambda}_{\mathbf{w}2} \neq \bar{\lambda}_{\mathbf{w}1}$. It follows that $\bar{\lambda}_{\mathbf{w}1}$ and $\bar{\lambda}_{\mathbf{w}2}$ do play a role as discussed in Sections 3.1 and 3.2.

3.4. Rotational indeterminacy

In Sections 3.1 and 3.2 we discussed how rotational indeterminacy, induced by the fact that both the factor and the loadings are latent, affects the results stated in Propositions 3.1 and 3.2, respectively. We now further describe what rotational indeterminacy specifically entails in a one-factor model such as the one we are considering.

Given the model in (1) and the assumptions of Proposition 3.1, from (3) we have that the estimator $\hat{f}_t$ for f_t is such that[3]

$$\hat{f}_t \xrightarrow{p} \mathbb{I}\left(t \leq T_{\mathbf{x}}^0\right) \bar{\lambda}_{\mathbf{w}1} f_t + \mathbb{I}\left(t > T_{\mathbf{x}}^0\right) \bar{\lambda}_{\mathbf{w}2} f_t. \tag{8}$$

Therefore, in the specific case of a single-factor model with one break, rotational indeterminacy takes the form of a regime-specific rescaling of the latent factor f_t.

3.5. Factor heteroskedasticity

For ease of exposition, in Propositions 3.1 and 3.2 we assume $f_t \sim \text{IID}\mathcal{N}\left(0, \sigma_f^2\right)$ and, in particular, that f_t has constant variance equal to σ_f^2. Under this simplifying assumption we can map changes in the variance of the estimator $\hat{f}_t$ for f_t to changes in the average loadings. Formally, from (8) we have

$$\text{Var}\left(\hat{f}_t\right) \to \mathbb{I}\left(t \leq T_{\mathbf{x}}^0\right) \bar{\lambda}_{\mathbf{w}1}^2 \sigma_f^2 + \mathbb{I}\left(t > T_{\mathbf{x}}^0\right) \bar{\lambda}_{\mathbf{w}2}^2 \sigma_f^2,$$

which means that the variance of $\hat{f}_t$ before and after the break is equal to $\bar{\lambda}_{\mathbf{w}1}^2 \sigma_f^2$ and $\bar{\lambda}_{\mathbf{w}2}^2 \sigma_f^2$, respectively. Therefore, under factor homoskedasticity, any change in the variance of $\hat{f}_t$ is proportional to the change in the average factor loadings.[4] However, in the more realistic scenario in which f_t is heteroskedastic, such a mapping is no longer possible: in this case, a change in the variance of $\hat{f}_t$ may be due a change in the average value of the loadings or to a change in the variance of f_t (or both).[5]

4. Numerical results

Out-of-sample forecasts are typically analyzed via the (root) mean squared forecast error. In our setup, as $N \to \infty$, this is defined as

$$\text{MSFE} = \lim_{N\to\infty} \text{E}\left\{\left[y_{T+1} - \hat{y}_{T+1}\left(T_y^e\right)\right]^2\right\} = \text{E}\left(y_{T+1}^2\right) - 2 \lim_{N\to\infty} \text{E}\left[y_{T+1}\hat{y}_{T+1}\left(T_y^e\right)\right] + \lim_{N\to\infty} \text{E}\left\{\left[\hat{y}_{T+1}\left(T_y^e\right)\right]^2\right\}.$$

Given the assumptions of Proposition 3.1, $\text{E}\left(y_{T+1}^2\right) = \gamma_2^2 \sigma_f^2 + \sigma_\varepsilon^2$, and $\lim_{N\to\infty} \text{E}\left[y_{T+1}\hat{y}_{T+1}\left(T_y^e\right)\right]$ is as in Proposition 3.2. However, although it is mathematically possible to obtain the closed form expression for $\lim_{N\to\infty} \text{E}\left\{\left[\hat{y}_{T+1}\left(T_y^e\right)\right]^2\right\}$ is

[3] See Eq. (10) in Appendix.

[4] This intuition underlies the tests for structural stability in large dimensional factor models developed in Chen et al. (2014), and Han and Inoue (2015), according to which structural shifts in factor loadings can be inferred from the covariance structure of the estimated factors provided that the true latent factors are homoskedastic.

[5] This more realistic scenario, in which factor heteroskedasticity is allowed, is accounted for in the test for regime changes developed in Massacci (2021).

difficult to interpret.[6] For this reason, we analyze the root mean squared forecast error by Monte Carlo simulations. In what follows, Section 4.1 presents the data generating process, while Section 4.2 discusses the results.

4.1. Data generating process

We consider the data generating process

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

where s denotes the replication, for $s = 1, \dots, 2000$. We consider $N = 100$ and $T = 201$, so that $T - 1 = 200$.[7] Define $\delta_{\mathbf{x}i} = \lambda_{1i} - \lambda_{2i}$, for $i = 1, \dots, N$. 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 $i = 1, \dots, N/2$, and correspondingly $\delta_{\mathbf{x}i} = 0.00, -1.50, -3.00$ for $i = N/2 + 1, \dots, N$; 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 this allows us 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

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

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 $i = 1, \dots, N$, 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

$$\text{RMSFE}_k = \sqrt{\frac{\sum_{s=1}^{S}\left[y_{T+1}^s - \hat{y}_{k,T+1}^s\left(T_y^e\right)\right]^2}{S}}, \quad k = \text{ post-break, expanding, rolling,} \tag{9}$$

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

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. In contrast, 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 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

[6] Technical details are available upon request.

[7] Within a given replication s, consistent estimation of the factor f_t^s (up to a rotation) requires $N \to \infty$. In our Monte Carlo setup, $N = 100$ cross-sectional observations are sufficient to estimate the factor accurately. In practice, empirical researchers may encounter situations where N is much larger. For example, in the empirical analysis conducted in Section 5, the number of firms is likely to be in the order of thousands.

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	1.0614	1.0560	1.0796	1.0614	1.0560	1.0796
0.00	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
	0.00	1.0614	1.2858	1.0796	1.0614	1.2858	1.0796
1.00	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
	0.00	1.0614	2.0268	1.0796	1.0614	2.0268	1.0796
2.00	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
	0.00	1.0614	3.2791	1.0796	1.0614	3.2791	1.0796
3.00	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	1.2193	1.0560	1.0796	1.2193	1.0560	1.0796
0.00	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
	0.00	1.2193	1.9312	1.7342	1.2193	1.9312	1.7342
1.00	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
	0.00	1.2193	4.6567	3.7817	1.2193	4.6567	3.7817
2.00	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
	0.00	1.2193	9.2327	7.2220	1.2193	9.2327	7.2220
3.00	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 (9) for Experiment 1, whose data generating process and results are described in Section 4.1 and in Section 4.2, respectively.

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 this case, the relative performance of the estimation window depends on the interaction between the breaks in the forecasting equation and the factor model.

5. Predicting stock returns with aggregate illiquidity

Chen et al. (2018) show that aggregate illiquidity contains valuable information for out-of-sample forecasting of equity returns. In particular, they argue that the tick-size reduction occurring in 2001 affected the predictive content of illiquidity measures as it induced a structural break in illiquidity proxies. We investigate the consequences of such a break on the stock return out-of-sample forecasting performance of illiquidity measures under post-break, expanding, and rolling window estimation schemes. In what follows, Section 5.1 describes the data and the model specification, while Section 5.2 discusses the results.

5.1. Data, model specification and forecasting method

Based on the specification in (1) and (2), we model stock returns in excess of the risk-free rate through the factor-augmented predictive regression

$$\begin{aligned} illiq_{it} &= \mathbb{I}\left(t \leq T_{\mathbf{x}}^0\right)\lambda_{1i}f_t + \mathbb{I}\left(t > T_{\mathbf{x}}^0\right)\lambda_{2i}f_t + e_{it}, \\ r_{t+1} &= \mathbb{I}\left(t \leq T_y^0\right)\gamma_1 f_t + \mathbb{I}\left(t > T_y^0\right)\gamma_2 f_t + \varepsilon_{t+1}, \end{aligned} \qquad i = 1,\ldots,N_t, \quad t = 1,\ldots,T, \quad 1 < T_{\mathbf{x}}^0, T_y^0 < T,$$

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	0.00	1.0614	1.0560	1.0796	1.0614	1.0560	1.0796
	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
1.00	0.00	1.0614	1.2858	1.0796	1.0614	1.2858	1.0796
	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
2.00	0.00	1.0614	2.0268	1.0796	1.0614	2.0268	1.0796
	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
3.00	0.00	1.0614	3.2791	1.0796	1.0614	3.2791	1.0796
	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	0.00	1.2193	1.0560	1.0796	1.2193	1.0560	1.0796
	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
1.00	0.00	1.2193	1.9312	1.7342	1.2193	1.9312	1.7342
	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
2.00	0.00	1.2193	4.6567	3.7817	1.2193	4.6567	3.7817
	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
3.00	0.00	1.2193	9.2327	7.2220	1.2193	9.2327	7.2220
	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 (9) for Experiment 2a, whose data generating process and results are described in Section 4.1 and in Section 4.2, respectively.

where the cross-sectional dimension N_t is given by the number of firms in the sample at time t, and it is allowed to be time-dependent.[8]

We consider data collected at a monthly frequency. The variable $illiq_{it}$ is one of the following firm-level measures considered in Chen et al. (2018): $roll_{it}$ by Roll (1984); cs_{it} by Corwin and Schultz (2012); fht_{it} by Fong et al. (2017); $tick_{it}$ by Holden (2009) and Goyenko et al. (2009); $zeros_{it}$ by Lesmond et al. (1999); ami_{it} by Amihud (2002); $amito_{it}$ as in Brennan et al. (2013); PS_{it} by Pástor and Stambaugh (2003); hm_{it} as in Hou and Moskowitz (2005). All quantities are constructed in such a way that they provide a measure of illiquidity. For each time period t, the aggregate illiquidity measure is the factor estimate $\hat{f}_t$ obtained as the equal-weighted cross-sectional average $\hat{f}_t = N_t^{-1} \sum_{i=1}^{N_t} illiq_{it}$.[9]

We assume $T_{\mathbf{x}}^0 = T_y^0$, namely, the break in the factor model occurs at the same time as the break in the factor-augmented regression. We make the common breakpoint be January 2001, which is the month at the end of which the tick-size reduction was implemented.[10] Consistently with our analytical and numerical results in Sections 3 and 4, respectively, we study post-break, expanding, and rolling window estimation schemes. As in Timmermann (2008), the expanding window estimation scheme uses an initial window of 120 monthly observations, and it begins in February 1991. In light of this choice, the rolling window scheme adopts an estimation window of fixed length equal to 120 monthly observations.

8 It is easy to show that the analytical and numerical results in Sections 3 and 4, respectively, remain valid when the cross-sectional dimension of the factor model is time-dependent. Intuitively, this is because the cross-sectional average estimator of Pesaran (2006) does not use information stemming from the time-series dimension of the underlying panel. Also, for internal consistency with the factor augmented regression in (2), we consider a forecasting equation for r_{t+1} that does not include an intercept. Adding an intercept would lead to qualitatively similar results and analogous recommendations.

9 The aggregate illiquidity measures are kindly made publicly available at https://sites.google.com/site/yongchenfinance/.

10 The implementation date for the tick-size reduction we consider is January 29, 2001.

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 (9) for Experiment 2b, whose data generating process and results are described in Section 4.1 and in Section 4.2, respectively.

We evaluate the out-of-sample forecasts through the root mean squared forecast error (RMSFE), whose performance is studied numerically in Section 4. We consider four post-break evaluation windows: (*i*) between 10 and 20 months after the break; (*ii*) between 10 and 60 months after the break; (*iii*) between 100 and 110 months after the break; (*iv*) between 100 and 150 months after the break. This combination of evaluation windows allows us to evaluate the out-of-sample performance across estimation windows depending on the starting point after the break (namely, 10 and 100 months) and the length of the window for a given starting point (namely, 10 and 50 months).

5.2. Empirical results

The empirical results are collected in Table 3, which displays the RMSFE (in percentage terms) for all estimation windows, illiquidity measures, and evaluation windows we consider.

When the forecasts are evaluated for a short period after the break (namely, between 10 and 20 months after the tick-size reduction), the post-break estimation scheme generally has the edge over the expanding and rolling window counterparts, as it tends to have lower RMSFE. This is because only few observations are available after the break, and both the expanding and the rolling window schemes include misleading information stemming from the pre-break period, which biases the forecasts.

In all remaining post-break evaluation windows, the expanding window scheme tends to provide more accurate forecasts than the other two forecasting methods. This could be due to two simultaneous effects: (*i*) enough post-break observations are available, which reduces the bias of the forecasts; and (*ii*) the break induced by the tick-size reduction is such that the pre-break observations reduce the variance of the forecasts.

In conclusion, expanding window estimation generally provides the most accurate forecasts, except for short periods of time after the tick-size reduction, in which case the post-break scheme is preferable.

Table 3
Out-of-sample stock return-forecasting, RMSFE(%).

Panel A: Post-break										
Begin	End	*roll*	*cs*	*fht*	*tick*	*zeros*	*ami*	*amito*	*PS*	*hm*
10	20	5.332	5.376	5.377	5.375	5.404	5.635	5.422	6.959	5.603
10	60	3.974	3.999	4.008	3.999	4.021	3.879	4.009	4.956	4.143
100	110	4.272	4.278	4.321	4.310	4.394	4.216	4.392	4.791	4.668
100	150	4.400	4.404	4.423	4.417	4.456	4.435	4.468	4.613	4.599
Panel B: Expanding										
Begin	End	*roll*	*cs*	*fht*	*tick*	*zeros*	*ami*	*amito*	*PS*	*hm*
10	20	5.908	5.900	5.942	5.935	5.942	5.680	5.851	5.644	5.813
10	60	3.836	3.834	3.852	3.847	3.864	3.897	3.829	4.052	3.845
100	110	3.950	3.964	3.997	3.979	4.090	4.277	4.045	4.537	4.318
100	150	4.282	4.290	4.306	4.299	4.343	4.440	4.325	4.525	4.456
Panel C: Rolling										
Begin	End	*roll*	*cs*	*fht*	*tick*	*zeros*	*ami*	*amito*	*PS*	*hm*
10	20	5.871	5.864	5.893	5.892	5.872	5.671	5.816	5.651	5.731
10	60	3.837	3.836	3.853	3.848	3.865	3.864	3.833	4.075	3.871
100	110	4.284	4.289	4.328	4.316	4.398	4.212	4.386	4.639	4.676
100	150	4.370	4.373	4.395	4.389	4.426	4.438	4.418	4.556	4.574

This table displays the RMSFE (in percentage terms) for the out-of-sample forecasting analysis in Section 5.

Our findings should be interpreted with a caveat. Given the purpose of our work, which studies the consequences of neglecting structural instability when this occurs towards the end of the sample, the results collected in Table 3 rely only on point estimates. However, no formal test for differences in forecasting performance could be conducted. This implies that the relative advantages we have documented could be statistically insignificant in practice.

6. Conclusions

This paper studies out-of-sample forecasting in factor augmented regressions that experience structural instability in the factor model or 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 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.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Appendix

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

$$\begin{aligned}\hat{f}_t &= \mathbb{I}\left(t \le T_{\mathbf{x}}^0\right)\left(\sum_{i=1}^{N} w_i\lambda_{1i}\right)f_t + \mathbb{I}\left(t > T_{\mathbf{x}}^0\right)\left(\sum_{i=1}^{N} w_i\lambda_{2i}\right)f_t + \left(\sum_{i=1}^{N} w_i e_{it}\right) \\ &\xrightarrow{p} \mathbb{I}\left(t \le T_{\mathbf{x}}^0\right)\bar{\lambda}_{\mathbf{w}1}f_t + \mathbb{I}\left(t > T_{\mathbf{x}}^0\right)\bar{\lambda}_{\mathbf{w}2}f_t :\end{aligned} \tag{10}$$

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 \leq 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 \leq 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 \leq 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 \leq 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 \leq 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}{l}\mathbb{I}\left(t \leq T_{\mathbf{x}}^0\right)\mathbb{I}\left(t \leq T_y^0\right)\bar{\lambda}_{\mathbf{w}1} f_t \gamma_1 f_t + \mathbb{I}\left(t \leq 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 \leq 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 \leq 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 \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)\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 \leq 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 \leq 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 \leq T_{\mathbf{x}}^0\right)\left[\mathbb{I}\left(t \leq 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 \leq 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)}
\end{aligned}
$$

$$
\begin{aligned}
&+\mathbb{I}\left(T_{\mathbf{x}}^{0}>T_{y}^{e}\right)\frac{\gamma_1-\gamma_2}{\bar{\lambda}_{\mathbf{w}1}}\frac{T_{\mathbf{x}}^{0}-T_{y}^{e}}{T-\left(T_{y}^{e}+1\right)}\left[\sum_{t=1}^{T-1}\mathbb{I}\left(t>T_{y}^{e}\right)\mathbb{I}\left(t\le T_{\mathbf{x}}^{0}\right)f_tf_t\right]^{-1}\left[\sum_{t=1}^{T-1}\mathbb{I}\left(t>T_{y}^{e}\right)\mathbb{I}\left(t\le T_{\mathbf{x}}^{0}\right)\mathbb{I}\left(t\le T_{y}^{0}\right)f_tf_t\right]\\
&+\mathbb{I}\left(T_{\mathbf{x}}^{0}>T_{y}^{e}\right)\frac{1}{\bar{\lambda}_{\mathbf{w}1}}\frac{T_{\mathbf{x}}^{0}-T_{y}^{e}}{T-\left(T_{y}^{e}+1\right)}\left[\sum_{t=1}^{T-1}\mathbb{I}\left(t>T_{y}^{e}\right)\mathbb{I}\left(t\le T_{\mathbf{x}}^{0}\right)f_tf_t\right]^{-1}\left[\sum_{t=1}^{T-1}\mathbb{I}\left(t>T_{y}^{e}\right)\mathbb{I}\left(t\le T_{\mathbf{x}}^{0}\right)f_t\varepsilon_t\right]\\
&+\mathbb{I}\left(T_{\mathbf{x}}^{0}>T_{y}^{e}\right)\frac{\gamma_2}{\bar{\lambda}_{\mathbf{w}2}}\frac{T-\left(T_{\mathbf{x}}^{0}+1\right)}{T-\left(T_{y}^{e}+1\right)}\\
&+\mathbb{I}\left(T_{\mathbf{x}}^{0}>T_{y}^{e}\right)\frac{\gamma_1-\gamma_2}{\bar{\lambda}_{\mathbf{w}1}}\frac{T-\left(T_{\mathbf{x}}^{0}+1\right)}{T-\left(T_{y}^{e}+1\right)}\left[\mathbb{I}\left(t>T_{\mathbf{x}}^{0}\right)f_tf_t\right]^{-1}\left[\sum_{t=1}^{T-1}\mathbb{I}\left(t>T_{\mathbf{x}}^{0}\right)\mathbb{I}\left(t\le T_{y}^{0}\right)f_tf_t\right]\\
&+\mathbb{I}\left(T_{\mathbf{x}}^{0}>T_{y}^{e}\right)\frac{1}{\bar{\lambda}_{\mathbf{w}2}}\frac{T-\left(T_{\mathbf{x}}^{0}+1\right)}{T-\left(T_{y}^{e}+1\right)}\left[\sum_{t=1}^{T-1}\mathbb{I}\left(t>T_{\mathbf{x}}^{0}\right)f_tf_t\right]^{-1}\left[\sum_{t=1}^{T-1}\mathbb{I}\left(t>T_{\mathbf{x}}^{0}\right)f_t\varepsilon_t\right].
\end{aligned}
$$

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_tf_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\le T_{y}^{0}\right)f_tf_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\le T_{\mathbf{x}}^{0}\right)f_tf_t\right]^{-1}\left[\sum_{t=1}^{T-1}\mathbb{I}\left(t>T_{y}^{e}\right)\mathbb{I}\left(t\le\min\left\{T_{\mathbf{x}}^{0},T_{y}^{0}\right\}\right)f_tf_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_tf_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\le T_{\mathbf{x}}^{0}\right)f_tf_t\right]^{-1}\left[\sum_{t=1}^{T-1}\mathbb{I}\left(t>T_{y}^{e}\right)\mathbb{I}\left(t\le T_{\mathbf{x}}^{0}\right)f_t\varepsilon_{t+1}\right].
\end{aligned}
$$

It follows that

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

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

$$
\begin{aligned}
&\lim_{N\to\infty}\mathrm{E}\left[\hat{\gamma}_2\left(T_{y}^{e}\right)\right]\\
=\;&\frac{\gamma_2}{\bar{\lambda}_{\mathbf{w}2}}\frac{T-\left(T_{y}^{e}+1\right)}{T-\left(T_{y}^{e}+1\right)}+\frac{\gamma_1-\gamma_2}{\bar{\lambda}_{\mathbf{w}2}}\frac{T-1-\left(T_{y}^{e}+1\right)-1}{T-1-\left(T_{y}^{e}+1\right)-1}\mathrm{E}\left\{\left[\sum_{t=1}^{T-1}\mathbb{I}\left(t>T_{y}^{e}\right)f_tf_t\right]^{-1}\left[\sum_{t=1}^{T-1}\mathbb{I}\left(t>T_{y}^{e}\right)\mathbb{I}\left(t\le T_{y}^{0}\right)f_tf_t\right]\right\}\\
=\;&\frac{1}{\bar{\lambda}_{\mathbf{w}2}}\left[\gamma_2+\left(\gamma_1-\gamma_2\right)\frac{T_{y}^{0}-T_{y}^{e}}{T-\left(T_{y}^{e}+1\right)}\right],
\end{aligned}
$$

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 $T_y^0 \le T_{\mathbf{x}}^0 \le T-1$,

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

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

Amihud, Y. (2002). Illiquidity and stock returns: Cross-section and time-series effects. *Journal of Financial Markets*, *5*(1), 31–56.
Baetje, F., & Menkhoff, L. (2016). Equity premium prediction: Are economic and technical indicators unstable? *International Journal of Forecasting*, *32*(4), 1193–1207.
Bai, J. (2003). Inferential theory for factor models of large dimensions. *Econometrica*, *71*(1), 135–171.
Bai, J., & Ng, S. (2002). Determining the number of factors in approximate factor models. *Econometrica*, *70*(1), 191–221.
Bai, J., & Ng, S. (2006). Confidence intervals for diffusion index forecasts and inference for factor-augmented regressions. *Econometrica*, *74*(4), 1133–1150.
Bai, J., & Perron, P. (1998). Estimating and testing linear models with multiple structural changes. *Econometrica*, *66*(1), 47–78.
Bair, E., Hastie, T., Paul, D., & Tibshirani, R. (2006). Prediction by supervised principal components. *Journal of the American Statistical Association*, *101*(473), 119–137.
Brennan, M., Huh, S. W., & Subrahmanyam, A. (2013). An analysis of the Amihud illiquidity premium. *Review of Asset Pricing Studies*, *3*(1), 133–176.
Çakmakli, C., & van Dijk, D. (2016). Getting the most out of macroeconomic information for predicting excess stock returns. *International Journal of Forecasting*, *32*(3), 650–668.
Chen, L., Dolado, J. J., & Gonzalo, J. (2014). Detecting big structural breaks in large factor models. *Journal of Econometrics*, *180*(1), 30–48.
Chen, Y., Eaton, G. W., & Paye, B. S. (2018). Micro(structure) before macro? The predictive power of aggregate illiquidity for stock returns and economic activity. *Journal of Financial Economics*, *30*(1), 48–73.
Corradi, V., & Swanson, N. R. (2014). Testing for structural stability of factor augmented forecasting models. *Journal of Econometrics*, *182*(1), 100–118.
Corwin, S. A., & Schultz, P. (2012). A simple way to estimate bid–ask spreads from daily high and low prices. *Journal of Finance*, *67*(2), 719–760.
Farmer, L., Schmidt, L., & Timmermann, A. (2022). Pockets of predictability. *Journal of Finance*, forthcoming.
Fong, K. Y. L., Holden, C. W., & Trzcinka, C. A. (2017). What are the best liquidity proxies for global research? *Review of Finance*, *21*(4), 1355–1401.
Forni, M., Giovannelli, A., Lippi, M., & Soccorsi, S. (2018). Dynamic factor model with infinite dimensional factor space: Forecasting. *Journal of Applied Econometrics*, *33*(5), 625–642.
Giacomini, R., & White, H. (2006). Tests of conditional predictive ability. *Econometrica*, *74*(6), 1545–1578.
Giovannelli, A., Massacci, D., & Soccorsi, S. (2021). Forecasting stock returns with large dimensional factor models. *Journal of Empirical Finance*, *63*, 252–269.
Gonçalves, S., McCracken, M. W., & Perron, B. (2017). Tests of equal accuracy for nested models with estimated factors. *Journal of Econometrics*, *198*(2), 231–252.

Goyenko, R. Y., Holden, C. W., & Trzcinka, C. A. (2009). Do liquidity measures measure liquidity? *Journal of Financial Economics*, *92*(2), 153–181.
Han, X., & Inoue, A. (2015). Tests for parameter instability in dynamic factor models. *Econometric Theory*, *31*(5), 1117–1152.
Holden, C. W. (2009). New low-frequency spread measures. *Journal of Financial Markets*, *12*(4), 778–813.
Hou, K., & Moskowitz, T. J. (2005). Market frictions, price delay, and the cross-section of expected returns. *Review of Financial Studies*, *18*(3), 981–1020.
Lesmond, D. A., Ogden, J. P., & Trzcinka, C. A. (1999). A new estimate of transaction costs. *Review of Financial Studies*, *12*(5), 1113–1141.
Ludvigson, S. C., & Ng, S. (2007). The empirical risk-return relation: A factor analysis approach. *Journal of Financial Economics*, *83*(1), 171–222.
Ludvigson, S. C., & Ng, S. (2009). Macro factors in bond risk premia. *Review of Financial Studies*, *22*(12), 5027–5067.
Massacci, D. (2019). Unstable diffusion indexes: With an application to bond risk premia. *Oxford Bulletin of Economics and Statistics*, *81*(6), 1376–1400.
Massacci, D. (2021). Testing for regime changes in portfolios with a large number of assets: A robust approach to factor heteroskedasticity. *Journal of Financial Econometrics*, http://dx.doi.org/10.1093/jjfinec/nbaa046, available at.
Neely, C. J., Rapach, D. E., Tu, J., & Zhou, G. (2014). Forecasting the equity risk premium: The role of technical indicators. *Management Science*, *60*(7), 1772–1791.
Pástor, L., & Stambaugh, R. F. (2001). The equity premium and structural breaks. *Journal of Finance*, *56*(4), 1207–1239.
Pástor, L., & Stambaugh, R. F. (2003). Liquidity risk and expected stock returns. *Journal of Political Economy*, *111*(3), 642–685.
Paye, B. S., & Timmermann, A. (2006). Instability of return prediction models. *Journal of Empirical Finance*, *13*(3), 274–315.
Pesaran, M. H. (2006). Estimation and inference in large heterogeneous panels with a multifactor error structure. *Econometrica*, *74*(4), 967–1012.
Pesaran, M. H., & Timmermann, A. (2004). How costly is to ignore breaks when forecasting the direction of a time series? *International Journal of Forecasting*, *20*(3), 411–425.
Pesaran, M. H., & Timmermann, A. (2007). Selection of estimation window in the presence of breaks. *Journal of Econometrics*, *137*(1), 134–161.
Rapach, D. E., & Wohar, M. E. (2006). Structural breaks and predictive regression models of aggregate U.S. stock returns. *Journal of Financial Econometrics*, *4*(2), 238–274.
Roll, R. (1984). A simple implicit measure of the effective bid–ask spread in an efficient market. *Journal of Finance*, *39*(4), 1127–1139.
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*, *59*(4), 1135–1190.
Stock, J. H., & Watson, M. W. (1996). Evidence on structural instability in macroeconomic time series relations. *Journal of Business and Economic Statistics*, *14*(1), 11–30.
Stock, J. H., & Watson, M. W. (2002a). Forecasting using principal components from a large number of predictors. *Journal of the American Statistical Association*, *97*(460), 1167–1179.
Stock, J. H., & Watson, M. W. (2002b). Macroeconomic forecasting using diffusion indexes. *Journal of Business and Economic Statistics*, *20*(2), 147–162.
Timmermann, A. (2001). Structural breaks, incomplete information, and stock prices. *Journal of Business and Economic Statistics*, *19*(3), 299–314.
Timmermann, A. (2008). Elusive return predictability. *International Journal of Forecasting*, *24*(1), 1–18.
Timmermann, A. (2018). Forecasting methods in finance. *Annual Review of Financial Economics*, *10*, 449–479.
Welch, I., & Goyal, A. (2008). A comprehensive look at the empirical performance of equity premium prediction. *Review of Financial Studies*, *21*(4), 1455–1508.