Out-of-sample inference with annual benchmark revisions*

Sílvia Gonçalves,† Michael W. McCracken,‡ and Yongxu Yao§

September 11, 2025

Abstract

This paper examines the properties of out-of-sample predictability tests evaluated with real-time data subject to annual benchmark revisions. The presence of both regular and annual revisions can create time heterogeneity in the moments of the real-time forecast evaluation function, which is not compatible with the standard covariance stationarity assumption used to derive the asymptotic theory of these tests. To cover both regular and annual revisions, we replace this standard assumption with a periodic covariance stationarity assumption that allows for periodic patterns of time heterogeneity. Despite the lack of stationarity, we show that the Clark and McCracken (2009) test statistic is robust to the presence of annual benchmark revisions. A similar robustness property is shared by the bootstrap test of Gonçalves, McCracken, and Yao (2025). Monte Carlo experiments indicate that both tests provide satisfactory finite sample size and power properties even in modest sample sizes. We conclude with an application to U.S. employment forecasting in the presence of real-time data.

*We thank participants at the Real Time Data Conference 2024 (Bank of Canada), the Midwest Econometrics Group 2024 (University of Kentucky), the opening conference of the Aarhus Center for Econometrics (Børkop, 2025), the Dolomite Macro Conference 2025 (Pera di Fassa), the International Association for Applied Econometrics 2025 (Turin), the ERMAS 2025 Conference (Iaşi), and seminar participants at Bundesbank in Frankfurt, St. Louis Fed Brownbag, Federal Reserve Board of Governors in DC, and University of California-Santa Barbara. The views expressed here are those of the individual authors and do not necessarily reflect official positions of the Federal Reserve Bank of St. Louis, the Federal Reserve System, or the Board of Governors.

†McGill University, Department of Economics, 855 Sherbrooke St. W., Montréal, Québec, H3A 2T7, Canada. E-mail: silvia.goncalves@mcgill.ca.

‡Research Division; Federal Reserve Bank of St. Louis; P.O. Box 442; St. Louis, MO 63166; email: michael.w.mccracken@stls.frb.org.

§McGill University, Department of Economics, 855 Sherbrooke St. W., Montréal, Québec, H3A 2T7, Canada. E-mail: yongxu.yao@mail.mcgill.ca

1 Introduction

In this paper, we develop tests of out-of-sample predictability in the presence of annual benchmark revisions. The vast majority of the literature on out-of-sample tests of predictive ability, including but not limited to those developed by Diebold and Mariano (1995), West (1996), Clark and McCracken (2001), Corradi and Swanson (2007), and Giacomini and Rossi (2010), ignores the potential for data revisions. An early counterexample is Clark and McCracken (2009), who derive asymptotics for tests of equal predictive ability between linear models evaluated under quadratic loss. To do so, they limit their analysis to the presence of regular revisions in which a fixed, constant number of the most recent observations is revised every period. More recently, Gonçalves, McCracken and Yao (2025) extend the work of Clark and McCracken (2009) in two ways. First, they derive asymptotic results in a broader framework that allows for more general classes of tests of predictive ability beyond just equal accuracy under quadratic loss. Second, they propose a novel bootstrap approach to inference for out-of-sample tests allowing for regular revisions.

Although Clark and McCracken (2009) and Gonçalves, McCracken and Yao (2025) investigate the role of data revisions on out-of-sample inference, they do not consider the effect of irregular data revisions. This is a non-trivial and important issue since revisions are known to be irregular. In particular, most macroeconomic series contain both regular and annual benchmark revisions. For instance, each month, the Bureau of Labor Statistics (BLS) provides a regular release of nonfarm payroll data in the current employment statistics (CES) report. This monthly report tabulates both the most recent change in payroll and the previous two months' revised changes. Besides these regular monthly releases and revised payroll data, each year the BLS also provides an additional revision to the payroll data that has been released and revised in the past 12 months or more by using the latest Unemployment Insurance tax numbers.

In this paper, we extend the results in Clark and McCracken (2009) and Gonçalves, McCracken and Yao (2025) to a framework that includes both regular and annual benchmark revisions. First, we adapt the results in West (1996) on out-of-sample tests of predictability to a framework that allows for both regular and annual data revisions. Second, we show that the test statistic proposed by Clark and McCracken (2009) to handle regular data revisions is also valid in the presence of benchmark revisions. Third, we show that the bootstrap test of Gonçalves, McCracken and Yao (2025) enjoys a similar robustness property.

Our new results build on a simple observation. When regular and annual benchmark revisions are present, the moments of the testing functions change over time in a periodic fashion. When data revisions are not present, West (1996) assumes these functions are covariance stationary. When only regular revisions are present, Clark and McCracken (2009) and Gonçalves, McCracken and Yao (2025) also assume covariance stationarity. For these latter two papers, the asymptotics follow the style of proof as in West (1996) but with additional details that keep track of the degree of revision within any given vintage. When annual benchmark revisions are present, this approach is no longer tenable.

To understand the issue, consider the following simple example. Suppose that we want to test for zero mean prediction error using an AR(p) model. For concreteness, suppose that the dependent variable is quarterly and exhibits one regular revision each quarter but, in the third quarter, the preceding 4 quarters are fully revised. Let the one-step ahead forecast be evaluated against the first release of the target variable. Whether the presence of the annual revision changes the moment structure of the predictors depends on the lag structure of the AR model. If p is 1, the predictor is always an initial release. If p is 2, the first predictor is always a first release and the second predictor is always once revised. If p is 3, in the first, second, and fourth quarters of the year, the first predictor is the first release while the second and third predictors are once revised. But in the third quarter, the third predictor is now fully revised. Insofar as the revision process changes the moments of the predictors, this implies that the moments of the testing function change during a calendar year — in a periodic fashion.

To cover both regular and annual revisions, we assume that the test function (and its gradient) are periodically stationary rather than covariance stationary. Under additional moment and mixing conditions, we are then able to establish asymptotic normality as in West (1996). The asymptotic variance differs, however, because unconditional heteroskedasticity is now present and the serial correlation structure is now significantly more complex. Despite this, we show that the test statistic suggested in Clark and McCracken (2009) is robust to the presence of annual benchmark revisions. There are two main reasons for this result. First, the standard HAC estimator used to compute the long-run variance of the testing function is still a consistent estimator under the null hypothesis despite the heterogeneity implied by annual revisions (see, e.g., Gallant and White (1988) for the properties of HAC estimators under heterogeneous weak dependence). Second, the contribution of parameter estimation uncertainty to the asymptotic variance is now proportional to the time average of the heterogeneous expectation of the gradient of the test function. Clark and McCracken's test statistic estimates this component by its sample analogue, which is still a consistent estimator under mean heterogeneity of that gradient.

A similar robustness property holds for the bootstrap approach of Gonçalves, McCracken and Yao (2025). By resampling the vector of predictors available in each vintage (jointly with the target dependent variable) using a block bootstrap, this method amounts to resampling the sequence of real-time heterogeneous weakly dependent testing functions. We show that this bootstrap is able to replicate the asymptotic variance of the test statistic under the null hypothesis when the mean of the testing function is zero (a result that follows from Gonçalves and White (2002)).

We provide Monte Carlo evidence on the finite sample efficacy of our analytical results. In these experiments, we allow both regular and annual revisions with an eye towards data that may be monthly or quarterly. This matters because quarterly series like consumption and real GDP as well as monthly series like employment and industrial production exhibit each type of revision. In the context of tests of zero mean prediction error (i.e., our earlier example) and equal forecast accuracy under quadratic loss (which we focus on in our empirical analysis), our simulations validate the robustness of the test statistic from Clark and McCracken (2009) and the bootstrap developed in Gonçalves, McCracken and Yao (2025) by showing that they can provide accurately sized and powerful tests in finite samples.

Finally, we apply our asymptotic and bootstrap results in the context of real-time forecasting of U.S. employment growth. We do so by considering the relative accuracy of a small number of ARX models that augment a benchmark autoregressive model with a measure of slack in the real economy: initial claims, total capacity utilization, or the vacancy rate. Throughout, real-time vintage data is used for all series in order to capture the impact of both regular and annual benchmark revisions. Overall, the benchmark autoregressive model is nominally most accurate at both a shorter 3-month horizon and a longer 12-month horizon. Nevertheless, both asymptotic- and bootstrap-based p-values indicate little statistical difference across the models. The sole exception is the model augmented with the vacancy rate advocated by Birinci et al. (2024) as a measure of labor market slack. This model is significantly less accurate in real time. While speculative, this may be due to the fact that it is released with a delay relative to the other predictors and hence its predictive content may be stale.

Before proceeding, it is worth emphasizing that all of our results are built on the same application of vintage data maintained in our previous work. We assume that at a given forecast origin, the current vintage of the observables is exclusively used to form the forecast. This implies that the parameter estimates are estimated using some data that have just been released, some that have been revised at least once, and some that have been fully revised.

Other methods have also been used in the literature. Koenig, Dolmas, and Piger (2003) and Clements and Galvão (2013) estimate model parameters using observations from multiple historical vintages with a shared level of revision. For example, one might only use the initial release of a series to estimate model parameters. This implies using just one observation each from the current and previous vintages — in contrast to our approach, which uses many values from a single vintage. See Clements and Galvão (2019) for a discussion of the trade-offs between using these two approaches.

The rest of the paper is organized as follows. Section 2 introduces the revision structure with annual revisions. Section 3 introduces the testing framework. Section 4 presents the assumptions. Section 5 shows results for asymptotic inference. Section 6 presents the bootstrap results. Section 7 presents Monte Carlo results. Section 8 applies our theoretical results in the context of real-time forecasting of U.S. employment growth. Section 9 concludes. Appendices A and B contain proofs of the theoretical results whereas additional simulation results are given in Appendix C.

2 Revision structure with annual revisions

The revision structure of this paper builds on Clark and McCracken (2009) and Gonçalves, McCracken and Yao (2025). Similarly to their settings, at each forecast origin $t = R, \ldots, T$, forecasts of a scalar target variable y are made using a finite dimensioned vector of predictors x based on the current vintage of data $\{y_s(t), x_s(t) : s = 1, \ldots, t\}$. The main difference is that we allow for both regular and annual revisions. Take the target variable y as an example. For each observation indexed by s, the first preliminary estimate of y_s is subject to $r_{max} - 1$ revisions to reveal its final estimate. Hence, there are r_{max} versions of y_s in total. We let $y_{s|i}$ be the ith release of y_s. When $i = 1$, $y_{s|1}$ is the first release. When $i = r_{max}$, $y_{s|i}$ is the final release, which we also write as y_s. We use $r - 1$ to represent the number of regular revisions needed on the first estimate of y_s, i.e., $y_{s|1}$, before $y_{s|r_{max}} = y_s$ is revealed. In Clark and McCracken (2009) and Gonçalves, McCracken and Yao (2025), $y_{s|1}$ is subject to $r - 1$ regular revisions only. For this setting, we set $r_{max} = r$. When $y_{s|1}$ is subject to regular and annual revisions, we set $r_{max} = r + r_b - 1$, where $r_b - 1$ represents the number of annual revisions needed on $y_{s|1}$ before we obtain $y_{s|r_{max}} = y_s$. Throughout, for each first release $y_{s|1}$, we assume annual revisions will only take place once all the regular revisions are completed. This means that among the total $r_{max} - 1$ revisions that are to be applied on $y_{s|1}$, the first $r - 1$ revisions are regular revisions, and the last $r_b - 1$ revisions are annual revisions. In other words, to update $y_{s|1}$ to $y_{s|r}$, $y_{s|1}$ receives $r - 1$ regular revisions; to update $y_{s|r}$ to

$y_{s|r_{max}}$, $y_{s|r}$ receives $r_b - 1$ annual revisions.

Regular revisions and annual revisions are learned with different frequencies. Regular revisions are learned every period whereas annual revisions are learned every λ periods. When $\lambda = 1$, annual revisions are structurally indistinguishable from regular revisions. If data have quarterly frequency, $\lambda = 4$ implies that each year we have one annual revision. If data are released with monthly frequency, one annual revision corresponds to $\lambda = 12$. For instance, using our notations, we can describe nonfarm payroll data as monthly released, annually revised real-time data with $r - 1 = 2$, $r_b - 1 \geq 1$, and $\lambda = 12$.

When data are subject to annual revisions, the preliminary observations in the annually revised real-time data set do not resemble the pattern that appears in a data set with regular revisions only. This is because annual revisions are learned every λ periods, and at the time of annual revisions, each of the last $r + (r_b - 1)\lambda - 1$ preliminary observations in the previous vintage column receives one revision. This causes the preliminary observations in the annually revised real-time data set to resemble a pattern of a string of flags.

Table 1: Structure of real-time data with no regular revisions ($r - 1 = 0$) and one annual revision ($r_b - 1 = 1$)

	Vintage date (t)								
	t_Y				t_{Y+1}				t_{Y+2}
Obs. s	R	$R+1$	$R+2$	$R+3$	$R+4$	$R+5$	$R+6$	$R+7$	$R+8$
1	y_1	y_1	y_1	y_1	y_1	y_1	y_1	y_1	y_1
2	y_2	y_2	y_2	y_2	y_2	y_2	y_2	y_2	y_2
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
$R-2$	y_{R-2}	y_{R-2}	y_{R-2}	y_{R-2}	y_{R-2}	y_{R-2}	y_{R-2}	y_{R-2}	y_{R-2}
$R-1$	y_{R-1}	y_{R-1}	y_{R-1}	y_{R-1}	y_{R-1}	y_{R-1}	y_{R-1}	y_{R-1}	y_{R-1}
R	$y_{R\|1}$	$y_{R\|1}$	$y_{R\|1}$	$y_{R\|1}$	y_R	y_R	y_R	y_R	y_R
$R+1$		$y_{R+1\|1}$	$y_{R+1\|1}$	$y_{R+1\|1}$	y_{R+1}	y_{R+1}	y_{R+1}	y_{R+1}	y_{R+1}
$R+2$			$y_{R+2\|1}$	$y_{R+2\|1}$	y_{R+2}	y_{R+2}	y_{R+2}	y_{R+2}	y_{R+2}
$R+3$				$y_{R+3\|1}$	y_{R+3}	y_{R+3}	y_{R+3}	y_{R+3}	y_{R+3}
$R+4$					$y_{R+4\|1}$	$y_{R+4\|1}$	$y_{R+4\|1}$	$y_{R+4\|1}$	y_{R+4}
$R+5$						$y_{R+5\|1}$	$y_{R+5\|1}$	$y_{R+5\|1}$	y_{R+5}
$R+6$							$y_{R+6\|1}$	$y_{R+6\|1}$	y_{R+6}
$R+7$								$y_{R+7\|1}$	y_{R+7}
$R+8$									$y_{R+8\|1}$

Consider a simple example where quarterly released data do not contain any regular revisions but are subject to a single annual revision. Table 1 illustrates this revision structure. We use t_Y, t_{Y+1}, and t_{Y+2} to represent dates where annual revisions take place, which in the table correspond to $t = R, R+4, R+8$, respectively. As in the regular revisions case, at each time period, a new preliminary observation is released (this corresponds to the main diagonal observations shaded in dark gray in the table). However, the time it takes for this observation to be updated to a final value depends on how far we are from an annual revision period. Take for instance $t = R$, an annual revision quarter. We observe a first released value of y_R, denoted by $y_{R|1}$. This preliminary value is subject to a single revision which will only take place at time $t = R + 4$, the next annual revision quarter. Thus, each vintage between $t = R$ and $t = R + 4$ will contain $y_{R|1}$ for observation $s = R$ (these correspond to the light gray observations in the table). Similarly, at time $t = R + 1$, the preliminary observation $y_{R+1|1}$ is released, and this will be updated to a final value y_{R+1} only at time $t = R + 4$. If instead the single revision was regular, $y_{R|1}$ and $y_{R+1|1}$ would be updated to final values in the next period $t = R+1$ and $t = R+2$, respectively, which would correspond to having only preliminary observations in the dark gray positions in the table. The fact that the revision is annual creates an additional number of preliminary observations in each vintage, all those observations shaded in light gray.

Consider a more realistic example where data are subject to single regular revisions and single annual revisions. Table 2 illustrates this revision structure. In this example, there are $r_{max} = r+r_b-1 = 3$ releases for each observation indexed by s. For instance, for observations indexed by R, $y_{R|1}, y_{R|2}$, and y_R are the first, second, and final release, respectively. To update $y_{R|1}$ to $y_{R|2}$, $y_{R|1}$ receives one regular revision at time $R + 1$. To update $y_{R|2}$ to y_R, $y_{R|2}$ receives one annual revision at time $R + 4$. The observations in dark gray would be the only preliminary observations in the table if there were no annual revisions. The latter add the observations in light gray.

Table 2: Structure of real-time data with one regular revision ($r-1=1$) and one annual revision ($r_b-1=1$)

	Vintage date (t)								
	t_Y				t_{Y+1}				t_{Y+2}
Obs. s	R	$R+1$	$R+2$	$R+3$	$R+4$	$R+5$	$R+6$	$R+7$	$R+8$
1	y_1	y_1	y_1	y_1	y_1	y_1	y_1	y_1	y_1
2	y_2	y_2	y_2	y_2	y_2	y_2	y_2	y_2	y_2
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
$R-2$	y_{R-2}	y_{R-2}	y_{R-2}	y_{R-2}	y_{R-2}	y_{R-2}	y_{R-2}	y_{R-2}	y_{R-2}
$R-1$	$y_{R-1\|2}$	$y_{R-1\|2}$	$y_{R-1\|2}$	$y_{R-1\|2}$	y_{R-1}	y_{R-1}	y_{R-1}	y_{R-1}	y_{R-1}
R	$y_{R\|1}$	$y_{R\|2}$	$y_{R\|2}$	$y_{R\|2}$	y_R	y_R	y_R	y_R	y_R
$R+1$		$y_{R+1\|1}$	$y_{R+1\|2}$	$y_{R+1\|2}$	y_{R+1}	y_{R+1}	y_{R+1}	y_{R+1}	y_{R+1}
$R+2$			$y_{R+2\|1}$	$y_{R+2\|2}$	y_{R+2}	y_{R+2}	y_{R+2}	y_{R+2}	y_{R+2}
$R+3$				$y_{R+3\|1}$	$y_{R+3\|2}$	$y_{R+3\|2}$	$y_{R+3\|2}$	$y_{R+3\|2}$	y_{R+3}
$R+4$					$y_{R+4\|1}$	$y_{R+4\|2}$	$y_{R+4\|2}$	$y_{R+4\|2}$	y_{R+4}
$R+5$						$y_{R+5\|1}$	$y_{R+5\|2}$	$y_{R+5\|2}$	y_{R+5}
$R+6$							$y_{R+6\|1}$	$y_{R+6\|2}$	y_{R+6}
$R+7$								$y_{R+7\|1}$	$y_{R+7\|2}$
$R+8$									$y_{R+8\|1}$

Comparing the revision structure with regular revisions only with the revisions structures in Tables 1 and 2, which also contain annual revisions, we can see that the number of preliminary observations in each vintage column would have been fixed at r in the regular revisions only case, whereas in Tables 1 and 2, this number can vary from r to $r-1+(r_b-1)\lambda$ depending on the vintage date. Because the number of preliminary observations may vary from one vintage to another in annually revised real-time data, the number of preliminary observations in a vector of vintage t predictors may also change when annually revised real-time data are used to construct forecasts, as we discuss in the next section.

3 The testing framework under annual revisions

3.1 The null hypothesis and the test statistic

Except for the revision structure of the underlying data, our testing framework is identical to that in Gonçalves, McCracken and Yao (2025). Specifically, at each forecast origin $t = R, \dots, T$, the τ-step ahead forecasts are generated by a simple linear model $x_t(t)'\hat{\beta}(t)$, where $\hat{\beta}(t)$ is a recursive OLS estimate based on vintage t data,

$$\hat{\beta}(t) \equiv \Big(\sum_{s=1+\tau+\ddot{r}}^{t} x_{s-\tau}(t)x'_{s-\tau}(t) \Big)^{-1} \sum_{s=1+\tau+\ddot{r}}^{t} x_{s-\tau}(t)y_s(t),$$

and $x_t(t)$ is a vector of vintage t predictors. Within $x_t(t)$, we allow each component to either be a lagged dependent variable or a lagged weakly exogenous variable up to lag order $\ddot{r}$. Because of this, we let the index s in the formula of $\hat{\beta}(t)$ start at $1+\tau+\ddot{r}$. This ensures that the $\ddot{r}^{th}$ lagged component in $x_{s-\tau}(t)$ is observed in the data set. If forecasts are made using an AR(p) forecasting model, then $\ddot{r}=p-1$, and $x_t(t)=(y_t(t),\ldots,y_{t-(p-1)}(t))'$.

The τ-step ahead forecast, $x_t(t)'\hat{\beta}(t)$, is evaluated against $y_{t+\tau|r'}$, the r'th release of the target variable $y_{t+\tau}$, where $r' \in \{1,\ldots,r_{max}\}$. Given a sequence of real-time forecasts, one is interested in testing the scalar null hypothesis

$$\mathsf{H}_0 : Ef(y_{t+\tau|r'}, x_t(t), \beta_0) \equiv Ef_{t+\tau} = 0,$$

where we let $f_{t+\tau} \equiv f(y_{t+\tau|r'}, x_t(t), \beta_0)$ for a known function $f(.)$, and where β_0 is the limit in probability of $\hat{\beta}(t)$.

To test the null hypothesis we form a test statistic based on the finite sample analogue of $f_{t+\tau}$, where β_0 is replaced by $\hat{\beta}(t)$:

$$\hat{S}_P = P^{-1/2}\sum_{t=R}^{T} f(y_{t+\tau|r'}, x_t(t), \hat{\beta}(t)),$$

where $P=T-R+1$.

3.2 The impact of annual revisions on the functional form of $f_{t+\tau}$

As in the regular revisions case, $f_{t+\tau}$ depends on potentially preliminary data either because the target variable $y_{t+\tau|r'}$ is preliminary or the predictors in $x_t(t)$ are preliminary. However, unlike the case where only regular revisions exist, the decomposition of these predictors into final and preliminary values may change across forecast origins when annual revisions are present, as we explain next. This implies that the functional form of $f_{t+\tau}$ can change over time in a way that is not compatible with the standard covariance stationarity assumption.

To describe the periodicity in $f_{t+\tau}$ induced by the periodicity in $x_t(t)$, we introduce the following additional notation. Throughout, for $t=R,\ldots,R+\lambda-1$, the functional form of $x_t(t)$ is defined as $x_t^{(t-R+1)}$. This notation implies that there are (potentially) λ different functional forms for $x_t(t)$ as t varies across $t=R$ through $t=R+\lambda-1$. In particular, we can write

$$x_R(R)=x_R^{(1)}, \quad x_{R+1}(R+1)=x_{R+1}^{(2)}, \quad \ldots, \quad x_{R+\lambda-1}(R+\lambda-1)=x_{R+\lambda-1}^{(\lambda)}.$$

For $t=R+\lambda,\ldots,T$, we write $x_t(t)=x_t^{(j)}$ if the functional form of $x_t(t)$ is identical to $x_{R+j-1}^{(j)}$ where $j=1,\ldots,\lambda$. This means that $x_t(t)=x_t^{(1)}$ if the functional form of $x_t(t)$ is identical to

$x_R^{(1)}$; and $x_t(t) = x_t^{(\lambda)}$ if the functional form of $x_t(t)$ is identical to $x_{R+\lambda-1}^{(\lambda)}$. Now consider how the functional form of $x_t(t)$ evolves with t when a one-step ahead AR(4) forecasting model is applied to data in Table 1. The form of $x_t(t)' = (y_t(t), \ldots, y_{t-3}(t))$ changes as follows:

$$x_t(t) = \begin{cases} (y_{t|1}, y_{t-1}, y_{t-2}, y_{t-3})' \equiv x_t^{(1)}, & t = R + 4n \\ (y_{t|1}, y_{t-1|1}, y_{t-2}, y_{t-3})' \equiv x_t^{(2)}, & t = (R+1) + 4n \\ (y_{t|1}, y_{t-1|1}, y_{t-2|1}, y_{t-3})' \equiv x_t^{(3)}, & t = (R+2) + 4n \\ (y_{t|1}, y_{t-1|1}, y_{t-2|1}, y_{t-3|1})' \equiv x_t^{(4)}, & t = (R+3) + 4n \end{cases}$$

where $n \in \mathbb{N}$.

In this particular example, there are four different versions of $x_t(t)$ and each version of $x_t(t)$ reappears once every 4 periods. For shorter AR forecasting models, some functional forms of $x_t(t)$ may appear more frequently than others over time. For instance, if the forecasting model is AR(2) and the data revisions structure is that in Table 1, we can show that $x_t(t)' \equiv (y_t(t), y_{t-1}(t))$ takes on only two different configurations: $x_t(t) = (y_{t|1}, y_{t-1})' \equiv x_t^{(1)}$ when $t = R + 4n$ with $n \in \mathbb{N}$, whereas $x_t(t) = (y_{t|1}, y_{t-1|1})' \equiv x_t^{(2)} = x_t^{(3)} = x_t^{(4)}$ for all other values of t. In this example, it is clear that $(y_{t|1}, y_{t-1|1})$ appears more frequently than $(y_{t|1}, y_{t-1})$ over time. Nonetheless, for $j = 1, \ldots, 4$, $x_t^{(j)}$ reappears at least one time every four periods. For annually revised real-time data with frequency λ, we can write in general,

$$x_t(t) = \begin{cases} x_t^{(1)}, & t = R + \lambda n \\ x_t^{(2)}, & t = (R+1) + \lambda n \\ & \vdots \\ x_t^{(\lambda)}, & t = (R + \lambda - 1) + \lambda n. \end{cases}$$

Similarly to $x_t(t)$, there are at most λ versions of $f_{t+\tau}$, and for any vintage t, the functional form of $f_{t+\tau}$ reappears at least one time every λ periods. For $n \in \mathbb{N}$, we let

$$f_{t+\tau} = \begin{cases} f(y_{t+\tau|r'}, x_t^{(1)}, \beta_0) \equiv f_{t+\tau}^{(1)}, & t = R + \lambda n \\ f(y_{t+\tau|r'}, x_t^{(2)}, \beta_0) \equiv f_{t+\tau}^{(2)}, & t = (R+1) + \lambda n \\ \quad \vdots \\ f(y_{t+\tau|r'}, x_t^{(\lambda)}, \beta_0) \equiv f_{t+\tau}^{(\lambda)}, & t = (R + \lambda - 1) + \lambda n. \end{cases} \tag{1}$$

Since the functional form of $f_{t+\tau}$ may change with t, the mean and the correlation structure of $f_{t+\tau}$ are not necessarily constant over time. This implies that the standard covariance stationary assumption on $f_{t+\tau}$ used in the existing literature (e.g., see Diebold and Mariano (1995), West (1996), Clark and McCracken (2009) and Gonçalves, McCracken and Yao (2025)) is not suitable when the functional form of $f_{t+\tau}$ is not time invariant. In the next section, we propose a set of assumptions that allows for the periodic heterogeneity of $f_{t+\tau}$ induced by the presence of annual revisions.

4 Assumptions

We adapt the assumptions in Gonçalves, McCracken and Yao (2025) to the presence of annual revisions.

Assumption 1 *For each* $j = 1, \ldots, \lambda$ *with* $\lambda < \infty$*, in an open neighborhood* $\mathcal{Z}$ *around* β_0 *and with probability one, (a)* $f_{t+\tau}^{(j)}(\beta)$ *is measurable and twice continuously differentiable. (b) There exists a constant* $D < \infty$ *such that for all* t *and* j, $\sup_{\beta \in \mathcal{Z}} \left| \frac{\partial^2 f_{t+\tau}^{(j)}(\beta)}{\partial \beta \partial \beta'} \right| < m_{t+\tau}$ *with a measurable function* $m_{t+\tau}$ *such that* $E(m_{t+\tau}) < D$*.*

Assumption 1 assumes that each version of $f_{t+\tau}$ is well approximated by a quadratic function in a neighborhood of β_0. When there is only one version of $f_{t+\tau}$, it corresponds to Assumption 1 in Gonçalves, McCracken and Yao (2025).

Assumption 2 *(a) The final-data estimate* $\hat{\beta}_t$ *satisfies* $\hat{\beta}_t - \beta_0 = B(t)H(t)$*, where*

$$B(t) = \left(t^{-1} \sum_{s=1+\tau+\ddot{r}}^{t} x_{s-\tau} x_{s-\tau}' \right)^{-1} \xrightarrow{a.s.} B, \quad H(t) = t^{-1} \sum_{s=1+\tau+\ddot{r}}^{t} h_s$$

with $E(h_s) = 0$, $B = (E(x_s x_s'))^{-1}$*, and* $h_s = x_{s-\tau}(y_s - x_{s-\tau}' \beta_0)$*.*

(b) The real-time data estimate $\hat{\beta}(t)$ *satisfies* $\hat{\beta}(t) - \beta_0 = \hat{B}(t)\hat{H}(t)$*, where*

$$\hat{B}(t) = \left(t^{-1} \sum_{s=1+\tau+\ddot{r}}^{t} x_{s-\tau}(t) x_{s-\tau}(t)' \right)^{-1}, \quad \hat{H}(t) = t^{-1} \sum_{s=1+\tau+\ddot{r}}^{t} h_s(t)$$

with $h_s(t) = x_{s-\tau}(t)(y_s(t) - x_{s-\tau}(t)'\beta_0)$*.*

Assumption 2 is notationally similar to Assumption 2 in Gonçalves, McCracken and Yao (2025). The main difference is that we allow for both regular and annual revisions in $x_{s-\tau}(t)$ and $y_s(t)$. We also set the lower bound of the sums to $1+\tau+\ddot{r}$ rather than $1+\tau$ to explicitly account for the fact that the predictor vector for vintage t may include lagged components up to order $\ddot{r}$. As explained above, the heterogeneity in the real-time function caused by annual revisions is only a concern when $\ddot{r} > r-1$, where $r-1$ denotes the number of regular revisions.

To describe our next assumption, let

$$V_{t+\tau} \equiv (f_{t+\tau}^{(1)}, \ldots, f_{t+\tau}^{(\lambda)})' \quad \text{and} \quad \frac{\partial V_{t+\tau}}{\partial \beta'} \equiv \left(\frac{\partial f_{t+\tau}^{(1)}}{\partial \beta}, \ldots, \frac{\partial f_{t+\tau}^{(\lambda)}}{\partial \beta} \right)'$$

for $t = R, R+1, \ldots, T$. With this notation, define

$$g_{t+\tau} \equiv \Big(\; (V_{t+\tau} - E(V_{t+\tau}))' \quad (\text{vec}(\tfrac{\partial V_{t+\tau}}{\partial \beta'} - E(\tfrac{\partial V_{t+\tau}}{\partial \beta'})')' \quad h_{t+\tau}' \quad x_t' - E(x_t)' \; \Big)'.$$

Note that $g_{t+\tau}$ includes the (centered) vector $V_{t+\tau}$ which collects the λ potentially different versions of $f_{t+\tau}$, as well as (the vectorized version of) its Jacobian matrix, $\frac{\partial V_{t+\tau}}{\partial \beta'}$, centered around its mean $E(\frac{\partial V_{t+\tau}}{\partial \beta'})$ (in addition, $g_{t+\tau}$ includes the scores h_t and the fully revised (centered) predictors x_t).

Our next assumption imposes a covariance stationarity assumption on $g_{t+\tau}$. Given equation (1), this assumption implies that the expected gradient of $f_{t+\tau}$ varies periodically. In particular, for $n \in \mathbb{N}$, we let

$$F_{t+\tau} \equiv E\Big(\frac{\partial f_{t+\tau}}{\partial \beta'}\Big) = \begin{cases} E\Big(\frac{\partial f^{(1)}_{t+\tau}}{\partial \beta'}\Big) \equiv F^{(1)}, & t = R + \lambda n \\ E\Big(\frac{\partial f^{(2)}_{t+\tau}}{\partial \beta'}\Big) \equiv F^{(2)}, & t = (R+1) + \lambda n \\ & \vdots \\ E\Big(\frac{\partial f^{(\lambda)}_{t+\tau}}{\partial \beta'}\Big) \equiv F^{(\lambda)}, & t = (R+\lambda-1) + \lambda n. \end{cases}$$

In the following, we let $\bar{F} = \frac{1}{\lambda}\sum_{j=1}^{\lambda} F^{(j)}$, where $F^{(j)} \equiv E\Big(\frac{\partial f^{(j)}_{t+\tau}}{\partial \beta'}\Big)$.

Assumption 3 *(a) For some $d > 1$ and $\delta > 0$, $\sup_t E\|g_{t+\tau}\|^{4d+\delta} < \infty$, where $\|\cdot\|$ denotes the Euclidean norm. (b) $g_{t+\tau}$ is covariance stationary. (c) $\{g_{t+\tau}\}$ is strong mixing with mixing coefficients of size $-3d/(d-1)$. (d) Ω is positive definite, where $\Omega \equiv \lim_{P,R\to\infty} \mathrm{Var}\Big(P^{-1/2}\sum_{t=R}^{T}(f_{t+\tau} - E(f_{t+\tau})) + \bar{F}BP^{-1/2}\sum_{t=R}^{T} H(t)\Big)$.*

Assumption 3 generalizes Assumption 3 in Gonçalves, McCracken and Yao (2025) to the annual benchmarks revision context. Rather than assuming that $f_{t+\tau}$ and $\frac{\partial f_{t+\tau}}{\partial \beta'}$ are covariance stationary, we impose a stationarity condition on the vector $V_{t+\tau}$ which contains $f^{(j)}_{t+\tau}$ for $j = 1, \ldots, \lambda$ as well as on its Jacobian matrix (in addition, we assume that the scores h_t and the fully revised predictors x_t are also stationary). Under this condition, and given the definition of $f_{t+\tau}$ given in equation (1), we can show that $f_{t+\tau}$ (and its gradient $\frac{\partial f_{t+\tau}}{\partial \beta'}$) are periodically stationary (or periodically correlated) time series with period λ. Following Hurd and Miamee (2007), this means that

$$E(f_{t+\tau}) = E(f_{t+\tau+\lambda}) \;\text{ and }\; Cov(f_{t+\tau+\lambda}, f_{s+\tau+\lambda}) = Cov(f_{t+\tau}, f_{s+\tau})$$

for each $t, s \in \mathbb{Z}$. Hence, Assumption 3 allows the moments and the dependence structure of $f_{t+\tau}$ and $\frac{\partial f_{t+\tau}}{\partial \beta'}$ to vary over time in a periodic manner, consistent with the periodicity induced by annual revisions. A special case of Assumption 3 is Assumption 3 of Gonçalves, McCracken and Yao (2025) where the functional forms of $f_{t+\tau}$ and $\frac{\partial f_{t+\tau}}{\partial \beta'}$ are time-invariant since they only allow regular revisions.

Assumption 4 *For some $d > 1$, $r_{max} < \infty$, and for $i = 1, \ldots, r_{max}$, $(y_{t|i}, x_t', x_{t|i}')'$ is $\mathcal{L}^{4d}$ bounded.*

Assumption 4 is the same as Assumption 4 in Gonçalves, McCracken and Yao (2025) in that we assume that the number of revisions is finite and the preliminary data have slightly more than four finite moments. The main difference is that here we allow for a total of $r_{max} = r - 1 + r_b$ releases of which $r - 1$ are regular and r_b are annual releases. When there are no annual revisions, i.e., $r_b = 1$, Assumption 4 corresponds to Assumption 4 in Gonçalves, McCracken and Yao (2025).

The following assumption is standard and is the same as Assumption 5 in Gonçalves, McCracken and Yao (2025).

Assumption 5 *$P, R \to \infty$ and $\lim_{P,R\to\infty} \frac{P}{R} = \pi$, where $0 \leq \pi < \infty$.*

5 Asymptotic theory

5.1 Asymptotic distribution of $\hat{S}_P$

We define the centered statistic as

$$\hat{S}_P^{\mu} \equiv \hat{S}_P - P^{-1/2} \sum_{t=R}^{T} E(f_{t+\tau}).$$

Following Gonçalves, McCracken and Yao (2025), we first show that $\hat{S}_P^{\mu}$ is asymptotically equivalent to

$$\tilde{S}_P^{\mu} \equiv \tilde{S}_P - P^{-1/2} \sum_{t=R}^{T} E(f_{t+\tau}) \equiv P^{-1/2} \sum_{t=R}^{T} (f_{t+\tau}(\hat{\beta}_t) - E(f_{t+\tau})),$$

where $\tilde{S}_P \equiv P^{-1/2} \sum_{t=R}^{T} f_{t+\tau}(\hat{\beta}_t)$ is a statistic based on the final data estimate $\hat{\beta}_t$, and $f_{t+\tau}(\hat{\beta}_t) = f(y_{t+\tau|r'}, x_t(t)', \hat{\beta}_t)$.

Lemma 5.1 *Under Assumptions 1-5, $\hat{S}_P^{\mu} = \tilde{S}_P^{\mu} + o_p(1)$.*

Lemma 5.1 extends Lemma 4.1 of Gonçalves, McCracken and Yao (2025) to the context of annually revised real-time data. Since the largest number of revisions occur for vintages with annual revisions and Assumption 4 assumes that this number is finite, replacing $\hat{\beta}(t)$ by $\hat{\beta}_t$ when defining the test statistic is asymptotically equivalent.

Our next result derives an asymptotic expansion of $\tilde{S}_P^{\mu}$ analogous to the usual expansion in West (1996). This expansion is based on a second-order mean value expansion of $f_{t+\tau}(\hat{\beta}_t)$

around β_0. Contrary to West (1996) and Clark and McCracken (2009), here we do not assume covariance stationarity. Instead, we rely on Assumption 3 which assumes that $f_{t+\tau}$ and $\frac{\partial f_{t+\tau}}{\partial \beta'}$ are periodically stationary.

Lemma 5.2 *Suppose Assumptions 1-5 hold. Then*

$$\tilde{S}_P^{\mu} = P^{-1/2}\sum_{t=R}^{T}(f_{t+\tau} - E(f_{t+\tau})) + \bar{F}BP^{-1/2}\sum_{t=R}^{T} H(t) + o_p(1) \equiv S_{1P} + \bar{F}BS_{2P} + o_p(1).$$

Lemma 5.2 generalizes Lemma 1 of Clark and McCracken (2009) by allowing for the presence of annual revisions and a general testing function $f_{t+\tau}$. The key difference is that the contribution of parameter estimation uncertainty is now proportional to $\bar{F} \equiv \frac{1}{\lambda}\sum_{j=1}^{\lambda} F^{(j)}$, the average of the periodic means of the gradient of f with respect to β. Without annual revisions and assuming covariance stationarity, $\bar{F} = F \equiv E\big(\frac{\partial f_{t+\tau}}{\partial \beta'}\big)$, but not otherwise. As the proof of Lemma 5.2 in Appendix A.2 shows, we rely crucially on the periodicity of $F_{t+\tau} \equiv E\big(\frac{\partial f_{t+\tau}}{\partial \beta'}\big)$ to prove that $S_D \equiv P^{-1/2}\sum_{t=R}^{T}(F_{t+\tau} - \bar{F})BH(t) = o_p(1)$, a remainder term that is zero under stationarity but not otherwise.

Next, we derive the asymptotic distribution of $\tilde{S}_P^{\mu}$ under our new set of assumptions. We introduce the following notation. Let $\Omega_1 \equiv \lim_{R,P\to\infty} Var(S_{1P})$, $\Omega_2 \equiv \lim_{R,P\to\infty} Var(S_{2P})$, and $\Omega_{12} \equiv \lim_{R,P\to\infty} Cov(S_{1P}, S_{2P})$.

Theorem 5.1 *Let Assumptions 1-5 hold. Then,* $\tilde{S}_P^{\mu} \xrightarrow{d} N(0,\Omega)$, *where* $\Omega = \Omega_1 + \bar{F}B\Omega_2B'\bar{F}' + 2\bar{F}B\Omega_{12}$.

To prove Theorem 5.1, we show in Appendix A.2 that S_{1P} and S_{2P} jointly converge in distribution to a normal random variable centered at zero, with a covariance matrix whose diagonal elements are Ω_1 and Ω_2, respectively, and the off-diagonal element is Ω_{12}. Although this result is similar to West (1996)'s Lemma 4.1, the method of proof is different since we do not assume stationarity.

Asymptotic inference based on Theorem 5.1 requires a consistent estimator of Ω, which we discuss next.

5.2 Consistent variance estimator under annual revisions

We first introduce consistent estimators for B and Ω_2. Note that these two terms only depend on finalized data, hence they are not affected by annual revisions, and their standard consistent estimators can be used. For B, we define $\hat{B} = (\frac{1}{T-\tau-\ddot{r}}\sum_{s=1+\tau+\ddot{r}}^{T} x_{s-\tau}x'_{s-\tau})^{-1}$.

Under Assumptions 2 and 4, $\hat{B} \xrightarrow{p} B$. For Ω_2, we follow West (1996) by writing Ω_2 as the product of a scale factor and the long run variance of h_s, i.e.,

$$\Omega_2 = 2\Pi \cdot \lim_{R,P\to\infty} Var(P^{-1/2} \sum_{s=R}^{T} h_s),$$

where $\Pi \equiv (1 - \pi^{-1} \ln(1+\pi)$. This product representation greatly simplifies the estimation of Ω_2, since $\hat{\pi} = \frac{P}{R} \to \pi$ implying that $\hat{\Pi} \to \Pi$, and consistent kernel estimators are available for the long run variance of h_s. Utilizing the product representation of Ω_2, we propose the following kernel-based estimator,

$$\hat{\Omega}_2 = 2\hat{\Pi} \cdot \frac{1}{P} \sum_{t=R}^{T} \sum_{s=R}^{T} \hat{h}_t \hat{h}_s' K((t-s)/b),$$

where $K(\cdot)$ is the kernel function, $b \equiv b_P$ is the kernel bandwidth, and $\hat{h}_t \equiv h(\hat{\beta}_T)$ with $\hat{\beta}_T$ the OLS estimate of β_0 defined in Assumption 2.

Next we present consistent estimators for $\bar{F} \equiv \lambda^{-1} \sum_{j=1}^{\lambda} F^{(j)}$, Ω_1, and Ω_{12}. Note that these three terms depend on the testing function $f_{t+\tau}$ which may not be covariance stationary in our context.

For $\bar{F}$, we define $\hat{\bar{F}} = P^{-1} \sum_{t=R}^{T} \partial f_{t+\tau}(\hat{\beta}_T)/\partial\beta'$, where $\partial f_{t+\tau}(\hat{\beta}_T)/\partial\beta'$ is the (row) vector of partial derivatives of $f_{t+\tau}$ with respect to β evaluated at $\hat{\beta}_T$. This is the same estimator as proposed by Clark and McCracken (2009), but its limit in probability is different under annual revisions. In particular, it is equal to $\bar{F}$ here as opposed to F under covariance stationarity of $f_{t+\tau}$.

To estimate Ω_1, we propose

$$\hat{\Omega}_1 = \frac{1}{P} \sum_{t=R}^{T} \sum_{s=R}^{T} (\hat{f}_t - \bar{f})(\hat{f}_s - \bar{f}) K((t-s)/b),$$

where $\hat{f}_t \equiv f_t(\hat{\beta}_T)$ and $\bar{f} \equiv P^{-1} \sum_{t=R}^{T} \hat{f}_t$. Although this is the same estimator as in West (1996) without revisions, and as in Clark and McCracken (2009) with regular revisions only, we do not require covariance stationarity for its validity, thus accommodating annual revisions. In particular, we show in Appendix A that $\hat{\Omega}_1$ is consistent for Ω_1 under the null hypothesis that $E(f_{t+\tau}) = 0$, a result that follows from de Jong and Davidson (2000) (see also Gallant and White (1988)) for general heterogeneous weakly dependent arrays.

Estimation of Ω_{12} is more involved. It depends on the long run covariance between $f_{t+\tau}$ and $H(t) \equiv t^{-1} \sum_{s=1+\tau+\ddot{r}}^{t} h_s$. When f_t is covariance stationary, we can write Ω_{12} as the product of Π and the long run covariance between f_t and h_s, a result which follows by

Lemma A6 in West (1996) and which motivates the standard estimator of Ω_{12} given by

$$\hat{\Omega}_{12} = \hat{\Pi} \cdot \frac{1}{P} \sum_{t=R}^{T} \sum_{s=R}^{T} (\hat{f}_t - \bar{f}) \hat{h}_s' K((t-s)/b).$$

Clark and McCracken (2009) rely on this estimator assuming that f_t depends on data subject to regular revisions only, which justifies their stationarity condition on f_t. However, when f_t is subject to annual revisions, stationarity may not hold, and we cannot directly apply Lemma A6 in West (1996). In order to obtain West's (1996) product representation for Ω_{12} and justify the standard estimator $\hat{\Omega}_{12}$, we exploit the periodic stationarity assumption on f_t and impose the following restriction on the dependence between f_t and h_s.

Assumption 6 *$E\big((f_s - E(f_s))h_t'\big) = 0$ for $|s-t| > M$ where M is a positive integer.*

Assumption 6 allows for the testing function f_t and the scores h_s to be correlated up to some finite lag order M. Although this is a more restrictive dependence condition than in West (1996) and Clark and McCracken (2009), it greatly simplifies the proof of the following result, which is the analogue of Lemma A6 in West (1996) under annual revisions.

Lemma 5.3 *Let Assumptions 1-6 hold. Then,*

$$\Omega_{12} = \Pi \cdot \lim_{R,P\to\infty} P^{-1} \sum_{t=R}^{T} \sum_{s=R}^{T} E\big((f_{t+\tau} - E(f_{t+\tau}))h_s'\big).$$

Lemma 5.3 justifies estimating Ω_{12} using $\hat{\Omega}_{12}$ defined above. Combining these results, our proposed estimator of Ω is

$$\hat{\Omega} = \hat{\Omega}_1 + \hat{\bar{F}}\hat{B}\hat{\Omega}_2\hat{B}'\hat{\bar{F}}' + 2\hat{\bar{F}}\hat{B}\hat{\Omega}_{12}'.$$

The following assumption provides sufficient conditions on $K(\cdot)$ and b for the consistency of the kernel-based estimators of Ω_1, Ω_2, and Ω_{12}. These conditions are similar to those used by de Jong and Davidson (2000), whose Theorem 2.2 is used when proving Theorem 5.2 below.

Assumption 7 *(a) Let $K(x)$ be a continuous kernel function such that for all real scalars x, $|K(x)| \leq 1$, $K(x) = K(-x)$ and $K(0) = 1$, $\int_{-\infty}^{\infty} |K(x)|dx < \infty$, and $\int_{-\infty}^{\infty} |\psi(\xi)|d\xi < \infty$, where $\psi(\xi) = (2\pi)^{-1} \int_{-\infty}^{\infty} K(x)e^{i\xi x}dx$. (b) $b \equiv b_P \to \infty$, $P^{-1/2}b_P \to 0$.*

Theorem 5.2 *Let Assumptions 1-7 hold, and assume that there exists $\mathcal{N}$, an open neighborhood around β_0, such that $E \sup_{\beta\in\mathcal{N}} \|\frac{\partial f_t}{\partial \beta'}\|^2 < \infty$. Then, under H_0, $\hat{\Omega} \xrightarrow{p} \Omega$.*

Theorem 5.2 shows the consistency of $\hat{\Omega}$ under H_0 when $E(f_{t+\tau}) = 0$. Under the alternative hypothesis, $\hat{\Omega}$ is not necessarily consistent for Ω unless $Ef_{t+\tau}$ is time invariant. The reason is that a bias term may appear in $\hat{\Omega}_1$ (and hence in $\hat{\Omega}$), as implied by Gallant and White (1988)'s Theorem 6.8. Because this bias is $O(b)$, the studentized statistic is of order $O_p(\sqrt{P/b})$, which diverges under the condition that $P^{-1/2}b \to 0$. This ensures the consistency of the standard test statistic.

6 Bootstrap inference

The main contribution of this section is to prove the validity of the bootstrap method proposed by Gonçalves, McCracken and Yao (2025) when applied to annually revised real-time data. In particular, we show here that this method is robust to the periodic heterogeneity in $f_{t+\tau}$ implied by the annual revisions. It can successfully replicate the asymptotic expansion in Lemma 5.2 and the asymptotic variance Ω described in Theorem 5.1 under the null hypothesis.

For completeness, we describe Gonçalves, McCracken and Yao (2025)'s bootstrap algorithm next. This algorithm consists of two applications of the moving blocks bootstrap (MBB): one to the first part of the sample ending at observation R, and another to the out-of-sample period starting after R. The application of the MBB guarantees that the serial dependence within each vintage is preserved. Because the same random indices are used to reshuffle the data across different vintages, the method preserves the dependence across vintages, including the dependence in the revisions.

In the following, ℓ denotes the block size used in applying the MBB.

Bootstrap algorithm (Gonçalves, McCracken and Yao (2025))

1. Let $R-(1+\tau+\ddot{r})+1 = k_1\ell$ and generate $I_1, \ldots, I_{k_1} \sim$ i.i.d. Uniform on $\{1+\tau+\ddot{r}, \ldots, R-\ell+1\}$. Then, for each $i = 1, \ldots, k_1$ and $j = 1, \ldots, \ell$, set $I_i+(j-1) = \gamma_{1+\tau+\ddot{r}+(i-1)\ell+(j-1)}$ and let

$$\{\gamma_s : s = 1+\tau+\ddot{r}, \ldots, R\} = \{\gamma_{1+\tau+\ddot{r}+(i-1)l+(j-1)} : i = 1, \ldots, k_1;\ j = 1, \ldots, \ell\}.$$

Let $T+\tau-(R+1)+1 = k_2\ell$ and generate $J_1, \ldots, J_{k_2} \sim$ i.i.d Uniform on $\{R+\tau, \ldots, T+\tau-\ell+1\}$. For each $i = 1, \ldots, k_2$ and $j = 1, \ldots, \ell$, set $J_i + (j-1) = \eta_{R+1+(i-1)\ell+(j-1)}$, and let

$$\{\eta_s : s = R+1, \ldots, T+\tau\} = \{\eta_{R+1+(i-1)\ell+(j-1)} : i = 1, \ldots, k_2;\ j = 1, \ldots, \ell\}.$$

2. For $t = R, \ldots, T$, set

$$(y_s^*, x_{s-\tau}^{*\prime}) = \begin{cases} (y_{\gamma_s}, x'_{\gamma_s-\tau}) & \text{if } 1+\tau+\ddot{r} \leq s \leq R \\ (y_{\eta_s}, x'_{\eta_s-\tau}) & \text{if } R+1 \leq s \leq t, \end{cases}$$

   and compute

$$\hat{\beta}_t^* = \left(t^{-1} \sum_{s=1+\tau+\ddot{r}}^{t} x_{s-\tau}^* x_{s-\tau}^{*\prime}\right)^{-1} t^{-1} \sum_{s=1+\tau+\ddot{r}}^{t} x_{s-\tau}^* y_s^*.$$

3. For $t = R, \ldots, T$, let

$$(y_{t+\tau|r'}^*, x_t^*(t)) = (y_{\eta_{t+\tau}|r'}, x_{\eta_{t+\tau}-\tau}(\eta_{t+\tau} - \tau)),$$

   and set

$$f_{t+\tau}^*(\hat{\beta}_t^*) \equiv f(y_{t+\tau|r'}^*, x_t^*(t), \hat{\beta}_t^*).$$

4. Compute

$$\tilde{S}_P^* \equiv P^{-1/2} \sum_{t=R}^{T} (f_{t+\tau}^*(\hat{\beta}_t^*) - f_{t+\tau}(\bar{\beta}_t)),$$

   where $\bar{\beta}_t = \frac{R}{t}\hat{\beta}_R + \frac{t-R}{t}\hat{\beta}_P$, with

$$\hat{\beta}_R = \left(R^{-1} \sum_{s=1+\tau+\ddot{r}}^{R} x_{s-\tau} x'_{s-\tau}\right)^{-1} R^{-1} \sum_{s=1+\tau+\ddot{r}}^{R} x_{s-\tau} y_s$$

   and

$$\hat{\beta}_P = \left(P^{-1} \sum_{s=R+\tau}^{T+\tau} x_{s-\tau} x'_{s-\tau}\right)^{-1} P^{-1} \sum_{s=R+\tau}^{T+\tau} x_{s-\tau} y_s.$$

5. Reject $\mathsf{H}_0 : E(f_{t+\tau}) = 0$ at level α if $|\tilde{S}_P| \geq \hat{c}_{1-\alpha}$, where $\hat{c}_{1-\alpha}$ is the $100(1-\alpha)^{th}$ percentile of the bootstrap distribution of $|\tilde{S}_P^*|$.

The first application of the MBB in step 1 draws random indices γ_t from a set of integers ending at R. These indices are used to reshuffle the first part of the sample (e.g., in Table 1, the first R rows), setting all observations as final. Although this does not mimic exactly the nature of the first part of the sample (e.g., the R^{th} row in Table 1 also contains preliminary values for the first four vintages), these data are used only in the estimation of β_0. Given our assumption of a finite number of revisions, setting those observations in the bootstrap as final does not have an impact on the test statistic asymptotically (this explains also why in step 2 we estimate β_0 using the final data estimator $\hat{\beta}_t^*$ rather than the real-time data estimator $\hat{\beta}^*(t)$).

The second application of the bootstrap in step 1 draws random indices η_t for $t \geq T+1$ from a set of integers starting at $R+\tau$, for which both final and preliminary observations are available. These indices are used twice, first in step 2 to obtain the bootstrap data for $t \geq R+1$ used in computing $\hat{\beta}_t^*$, setting these observations as final; and second in step 3 to obtain the bootstrap analogue $f_{t+\tau}^*(\hat{\beta}_t^*)$ of the testing function $f_{t+\tau}(\hat{\beta}_t)$ for $t \geq R$. In obtaining $f_{t+\tau}^*(\hat{\beta}_t^*) \equiv f(y_{t+\tau|r'}^*, x_t^*(t), \hat{\beta}_t^*)$, we resample the "pairs" $(y_{t+\tau|r'}^*, x_t^*(t)')$, where $x_t^*(t)$ is a resampled version of the vintage-t predictor vector $x_t(t)$ which contains a mix of final and preliminary observations. Gonçalves, McCracken and Yao (2025) showed the validity of this method under regular revisions only. In this case, $x_t(t)$ has a time-invariant structure, and its resampled version $x_t^*(t)$ mimics the structure of $x(t)$ for each t.

We show next that this method is also robust to annual revisions. This is despite the fact that the bootstrap does not mimic exactly the pattern of $x_t(t)$ when there are annual revisions. To see this, consider Table 1 and suppose that the target variable is y_{t+1} and the predictor is $x_t(t) = y_{t-1}(t)$ (as it would be the case if we relied on a restricted AR(2) model that only uses the twice-lagged value of the dependent variable as a predictor). The presence of annual revisions implies that $x_t(t)$ alternates between $y_{t-1|1}$ (when there are no annual revisions) and y_{t-1} (otherwise). According to step 3, the bootstrap analogue of $(y_{t+1}, x_t(t))$ is $(y_{t+1}^*, x_t^*(t)) = (y_{\eta_{t+1}}, x_{\eta_{t+1}-1}(\eta_{t+1}-1))$ for any $t \geq R$. Hence, for $t = R$, $x_R^*(R) = x_{\eta_{R+1}-1}(\eta_{R+1}-1)$. Suppose we draw η_{R+1} according to step 1 and we obtain $\eta_{R+1} = R+4$. The bootstrap analogue of $x_R(R) = y_{R-1}$ (a final value) is $x_{R+3}(R+3) = y_{R+2|1}$, a preliminary value. Hence, for $t = R$ we have replaced a final value of the predictor by a preliminary value. If instead we draw $\eta_{R+1} = R+5$, we obtain $x_R^*(R) = y_{R+3}$, a final value. Thus, depending on the realization of η_{R+1}, this bootstrap may or may not enforce the correct classification of $x_R(R)$ as a final value. Note that if only regular revisions exist (as in Gonçalves, Yao and McCracken (2025)), $x_t(t) = y_{t-1}$ for all t and its bootstrap analogue is always a final value.

Although Gonçalves, Yao and McCracken (2025)'s bootstrap does not always replicate the classification of the predictor into final or preliminary, it allows for both. In particular, by bootstrapping the pairs $(y_{t+\tau}, x_t(t))$, step 3 is equivalent to resampling the (heterogeneous) testing function $f_{t+\tau}$. As shown by Gonçalves and White (2002), the moving blocks bootstrap variance (and distribution) of a sample mean is consistent for heterogeneous weakly dependent arrays provided a certain mean heterogeneity condition holds (see their Assumption 2.2). In our context, this condition is satisfied under the null since $Ef_{t+\tau} = 0$. Hence, the moving blocks bootstrap in step 3 is able to replicate Ω_1 under the null. This in turn implies that the MBB is able to replicate the asymptotic distribution of $\tilde{S}_P$ under the null.

To describe these results, we let $f^*_{t+\tau} \equiv f(y^*_{t+\tau|r'}, x^*_t(t), \beta_0)$, $S^*_{1P} = P^{-1/2}\sum_{t=R}^{T}(f^*_{t+\tau} - f_{t+\tau})$, and

$$S^*_{2P} = a_{R,0}P^{-1/2}\sum_{s=1+\tau+\ddot{r}}^{R}(h^*_s - \bar{h}_R) + P^{-1/2}\sum_{i=1}^{P-1} a_{R,i}(h^*_{R+i} - \bar{h}_P) \equiv S^*_{2P.1} + S^*_{2P.2},$$

where $h^*_t = x^*_{t-\tau}(y^*_t - x^{*\prime}_{t-\tau}\beta_0)$, $\bar{h}_R = (R - \tau - \ddot{r})^{-1}\sum_{s=1+\tau+\ddot{r}}^{R} h_s$ and $\bar{h}_P = P^{-1}\sum_{s=R+\tau}^{T+\tau} h_s$.

As usual in the bootstrap literature, in the following we let P^* denote the bootstrap probability measure, conditional on the original sample. Appendix B.1 contains a description of our bootstrap notation, including the definition of $o^*_p(1)$.

Lemma 6.1 *Under Assumptions 1-5 and letting* $\ell \to \infty$ *such that* $\ell/\min\{\sqrt{R}, \sqrt{P}\} \to 0$,

$$\tilde{S}^*_P \equiv P^{-1/2}\sum_{t=R}^{T}(f^*_{t+\tau}(\hat{\beta}^*_t) - f_{t+\tau}(\bar{\beta}_t)) = S^*_{1P} + \bar{F}BS^*_{2P} + o^*_p(1).$$

Lemma 6.1 shows that the MBB replicates the asymptotic expansion of $\tilde{S}^\mu_P$ derived in Lemma 5.2. This result is instrumental in proving the consistency of the bootstrap distribution.

Theorem 6.1 *Suppose Assumptions 1-6 hold, and* $\ell \to \infty$ *such that* $\ell/\min\{\sqrt{R}, \sqrt{P}\} \to 0$. *Then, under* H_0, $\sup_{u\in\mathbb{R}}\left|P^*\left(\tilde{S}^*_P \le u\right) - \Pr\left(\tilde{S}_P \le u\right)\right| \to_p 0$.

Theorem 6.1 ensures that our bootstrap test has correct asymptotic size. To ensure that the bootstrap test has power, it suffices to show that the bootstrap statistic $\tilde{S}^*_P$ diverges at a smaller rate than $\tilde{S}_P$. Since the latter diverges at rate $O_p(P^{1/2})$, this requires that $\tilde{S}^*_P$ diverges at rate $P^{1/2-\delta}$, for some $\delta > 0$. This is true in our context if $Ef_{t+\tau}$ is time invariant under the alternative hypothesis, a result that follows from from Gonçalves and White (2002)'s Lemma 2.1. More generally, if $Ef_{t+\tau}$ is time varying, the bootstrap variance of $\tilde{S}^*_P$ contains a bias term that diverges at rate $O(\ell)$ but since $\ell = o(P^{1/2})$ by assumption, $\tilde{S}^*_P$ diverges at a smaller rate than $\tilde{S}_P$.

7 Monte Carlo simulations

In this section, we examine the finite sample performance of four out-of-sample tests in the context of the zero-mean prediction error test (Zero MPE) and the equal mean square error test (Equal MSE) with annually revised real-time data. The first test is the t-test of Diebold and Mariano (1995), labeled t_{DM}. It takes the form $t_{DM} = \hat{\Omega}_1^{-1/2}\hat{S}_P$. The second test is the t-test of Clark and McCracken (2009), labeled t_{CM}, implemented using $\hat{\Omega}$ as described in Section 5. The third test is a restricted version of the bootstrap algorithm in Gonçalves, McCracken and Yao (2025) which only replicates the structure of regular revisions, ignoring the presence of annual revisions. For instance, in the example described in Section 6 (see also below), it sets the bootstrap predictor always as final, i.e. $x_t^*(t) = y_{t-1}^* = y_{\eta_{t+1}-1}$ for all $t \geq R$ in Step 3. We label this test $Bootstrap_R$. The fourth test is the original bootstrap algorithm introduced in Gonçalves, McCracken and Yao (2025), labeled by *Bootstrap.*

7.1 Zero MPE experiment

The design of the experiment is comparable to that in Section 6.3.2 in Gonçalves, McCracken and Yao (2025). Specifically, we consider two AR models for forecasting y_{t+1}, one is an AR(1) model with the once lagged y value as predictor, and the other is a restricted AR(2) model where the predictor is the twice lagged y value (which is the example described in Section 6 when discussing the bootstrap). The data generating process of these two models can be described as

$$y_t = x_{t-1}\beta_0 + e_t + v_t \quad \text{and} \quad y_{t|1} = y_t - v_t + w_t,$$

where $e_t \sim \text{i.i.d.}N(0, \sigma_e^2)$, $v_t \sim \text{i.i.d.}N(0, \sigma_v^2)$, and $w_t \sim \text{i.i.d.}N(\mu_w, \sigma_w^2)$. We set $x_{t-1} = y_{t-1}$ for the AR(1) model and $x_{t-1} = y_{t-2}$ for the AR(2) model. For both models, we let $\sigma_e^2 = 0.3, \quad \sigma_v^2 = 0.2, \quad \sigma_w^2 = 0.2, \quad \mu_w = 0.85$. Under the null hypothesis, $\beta_0 = 0$, which implies $y_t \sim \text{i.i.d.}N(0, \sigma^2)$, with $\sigma^2 = \sigma_e^2 + \sigma_v^2$. We set $\beta_0 = 0.5$ under the alternative hypothesis. The first release $y_{t|1}$ is only subject to one annual revision, which takes place every λ periods. In our experiments, we set $\lambda = 1, 4, 12$. When $\lambda = 1$, this revision structure is equivalent to a single regular revision. When $\lambda = 4$, it represents quarterly released real-time data with a single annual revision, as in Table 1. When $\lambda = 12$, it corresponds to a monthly released real-time data with a single annual revision. The null hypothesis of interest takes the form

$$\mathsf{H}_0 : Ef_{t+1} = E(y_{t+1} - x_t(t)'\beta_0) = 0.$$

For the AR(1) model, $x_t(t) = y_t(t) = y_{t|1}$ for any t. This means that the functional form of f_{t+1} is time-invariant, i.e., $f_{t+1} = f_{t+1}^{(1)} = \ldots = f_{t+1}^{(\lambda)}$, and $F^{(1)} = F^{(2)} = \ldots = F^{(\lambda)} = \bar{F}$. For this reason, the asymptotic results in Clark and McCracken (2009) and Gonçalves, McCracken and Yao (2025) apply; the *Bootstrap* algorithm is valid and numerically equivalent to its restricted version $Bootstrap_R$.

Contrary to the AR(1) experiment, for $\lambda > 1$, the functional form of f_{t+1} changes across t in the AR(2) experiment. This is because in this case, $x_t(t) = y_{t-1}(t)$, and the functional form of $y_{t-1}(t)$ depends on t. Using Table 1 as an example, we see that in annual revision periods, $y_{t-1}(t) = y_{t|1}$ and $f_{t+1} = y_{t+1} - y_{t|1}\beta_0 \equiv f_{t+1}^{(1)}$, while in the remaining periods, $y_{t-1}(t) = y_t$ and $f_{t+1} = y_{t+1} - y_t\beta_0 \equiv f_{t+1}^{(2)} = f_{t+1}^{(3)} = f_{t+1}^{(4)}$. Hence, f_{t+1} is heterogenous, alternating between $f_{t+1}^{(1)}$ and $f_{t+1}^{(2)}$ every 4 periods. Note that in this case, $\bar{F} = \frac{1}{4}\sum_{j=1}^{4} F^{(j)}$, where $F^{(1)} = E\partial f_{R+1}^{(1)}/\partial\beta = -Ey_{R-1}(R) = -Ey_{R-1}$, and $F^{(2)} = F^{(3)} = F^{(4)} = E\partial f_{R+4}^{(4)}/\partial\beta = -Ey_{R+2}(R+3) = -Ey_{R+2|1}$. Under the null hypothesis, $Ey_{R-1} = 0$, and $Ey_{R+2|1} = \mu_w$. Note also that when $\lambda > 1$, *Bootstrap*$_R$ is different from *Bootstrap* because the first algorithm enforces the regular revision structure by setting $(y_{t+1}^*, x_t^*(t)) = (y_{\eta_{t+1}}, y_{\eta_{t+1}-1})$, whereas the *Bootstrap* algorithm sets $(y_{t+1}^*, x_t^*(t)) = (y_{\eta_{t+1}}, y_{\eta_{t+1}-1}(\eta_{t+1}))$. *Bootstrap*$_R$ effectively amounts to resampling only one version $f_{t+1}^{(1)}$ of f_{t+1}, thus disregarding its heterogeneity. Instead, *Bootstrap* reshuffles f_{t+1} in its entirety and only this method is valid asymptotically under annual revisions (except when $\lambda = 1$, in which case the two methods coincide).

Table 3 contains the results for nominal level $\alpha = 0.05$, based on 10,000 Monte Carlo replications and 499 bootstrap replications each. We set $R = 80$ and $P = 80$. Since $f_{t+1} = y_{t+1} \sim$ i.i.d.$N(0, \sigma^2)$ under the null, we set the block size $\ell = 1$ and the bandwidth parameter $b = 1$. Note also that Assumption 6 is automatically satisfied in this example.

The left panel of Table 3 shows that the Diebold and Mariano test is particularly oversized. This is because it does not account for parameter estimation uncertainty, which should not be ignored when the underlying data are subject to revisions. We also see that *Bootstrap*$_R$ is oversized when the AR(2) model is used on data with $\lambda = 4, 12$. Under these settings, f_{t+1} is heterogeneous, and resampling only one version of f_{t+1} ($f_{t+1}^{(1)}$, the one corresponding to predictor $x_t(t) = y_{t-1}$) is not valid. In contrast, both the Clark and McCracken test t_{CM} and the Gonçalves, McCracken and Yao (2025) *Bootstrap* test correct the size distortions brought by annual revisions. Their performance is very similar, both yielding slightly oversized tests when $\lambda = 12$. As expected, *Bootstrap* is identical to *Bootstrap*$_R$ when the forecasting model is AR(1) or when the forecasting model is AR(2) and $\lambda = 1$.

The right panel of Table 3 shows the power results. Except for the AR(2) forecasting model with $\lambda = 1$, all four tests have power larger than 0.88. When the model is AR(2) and $\lambda = 1$, we have that

$$Ef_{t+1}(\beta) = E(y_{t+1} - y_{t-1}(t)\beta) = E(y_{t+1} - y_{t-1}\beta) = 0 - 0 \cdot \beta = 0 \quad \text{for any } \beta < \infty.$$

Hence, setting β to 0.5 does not constitute a deviation from the null.

Table 3: Size and power results of zero MPE experiment

Tests	$\lambda=1$	4	12	$\lambda=1$	4	12
	size: AR(1)			power: AR(1)		
t_{DM}	0.157	0.159	0.207	0.978	0.983	0.992
t_{CM}	0.057	0.052	0.076	0.955	0.965	0.980
$Bootstrap_R$	0.053	0.052	0.072	0.958	0.967	0.982
$Boootstrap$	0.053	0.052	0.072	0.958	0.967	0.982
	size: AR(2)			power: AR(2)		
t_{DM}	0.055	0.108	0.162	0.057	0.930	0.987
t_{CM}	0.055	0.049	0.064	0.057	0.887	0.966
$Bootstrap_R$	0.053	0.106	0.162	0.056	0.942	0.992
$Bootstrap$	0.053	0.046	0.063	0.056	0.883	0.966

7.2 Equal MSE experiment

In this experiment, we let $y_{t+1} = z_{1,t}\beta_{1,0} + z_{2,t}\beta_{2,0} + e_{y,t+1} + v_{y,t+1}$, where $e_{y,t+1} \sim$ i.i.d. $N(0, \sigma^2_{e,y})$ is the error term and $v_{y,t+1} \sim$ i.i.d. $N(0, \sigma^2_{v,y})$ is the news term. For $i = 1, 2$, we let

$$z_{i,t} = \begin{cases} x_{i,t} & \text{if the data generating process is DL(1),} \\ x_{i,t-1} & \text{if the data generating process is DL(2),} \end{cases}$$

where $DL(q)$ denotes a distributed lag model of order q. We generate $x_{i,t} = e_{x_i,t} + v_{x_i,t}$ with $e_{x_i,t} \sim$ i.i.d. $N(0, \sigma^2_{e,x})$ and $v_{x_i,t} \sim$ i.i.d. $N(0, \sigma^2_{v,x})$. The revisions structure is similar to that in the zero MPE experiment. Specifically, at each time t, there are two estimates for $x_{1,t}$, $x_{2,t}$ and y_t, the preliminary estimate, and the final estimate. The final estimates are only observed after annual revisions, which take place every λ periods. We set $\lambda = 1, 4, 12$. For $i = 1, 2$, we let

$$x_{i,t|1} = x_{i,t} - v_{x_i,t} + w_{x_i,t}, \text{ and } \; y_{t|1} = y_t - v_{y,t} + w_{y,t},$$

where the noise term of x is i.i.d. $N(0, \sigma^2_{w,x})$, and the noise term of y is i.i.d. $N(0, \sigma^2_{w,y})$. For both DL(1) and DL(2) models, we let $\sigma^2_{e,y} = 0.1$, $\sigma^2_{v,y} = 0.9$, $\sigma^2_{w,y} = 0.2$, $\sigma^2_{e,x} = 1.7$, $\sigma^2_{v,x} = 0.3$, $\sigma^2_{w,x} = 4$. The null hypothesis of interest is

$$\mathsf{H}_0 : Ef_{t+1} = E((y_{t+1} - z_{1,t}(t)\beta_{1,0})^2 - (y_{t+1} - z_{2,t}(t)\beta_{2,0})^2) = 0,$$

where $z_{i,t}(t) = x_{i,t}(t)$ if the data generating process is DL(1) and $z_{i,t} = x_{i,t-1}(t)$ if the data generating process is DL(2). We let $\beta_{1,0} = \beta_{2,0} = 0.4$ under the null and set $\beta_{2,0}$ to 1 under the alternative hypothesis.

Note that for the DL(1) model, $x_{i,t}(t) = x_{i,t|1}$ for any t and λ. Hence, for this model, the functional form of f_{t+1}, $E\partial f_{t+1}/\partial\beta_{1,0}$ and $E\partial f_{t+1}/\partial\beta_{2,0}$ do not depend on t nor on the value of λ. Furthermore, as explained in Clark and McCracken (2009), using $x_{i,t|1}$ as a predictor

Table 4: Size and power results of equal MSE experiment

Tests	$\lambda=1$	4	12	$\lambda=1$	4	12
	size: DL(1)			power: DL(1)		
t_{DM}	0.104	0.099	0.088	0.505	0.456	0.370
t_{CM}	0.045	0.039	0.027	0.435	0.380	0.272
$Bootstrap_R$	0.041	0.031	0.017	0.409	0.326	0.178
Bootstrap	0.041	0.031	0.017	0.409	0.326	0.178
	size: DL(2)			power: DL(2)		
t_{DM}	0.073	0.083	0.086	0.986	0.195	0.270
t_{CM}	0.069	0.043	0.033	0.986	0.152	0.196
$Bootstrap_R$	0.058	0.301	0.305	0.985	0.518	0.612
Bootstrap	0.058	0.036	0.020	0.985	0.128	0.129

may result in a non zero $\bar{F}$. This means that Ω is of the long form given in Theorem 5.1. In contrast, when the forecasting model is DL(2) and $\lambda = 1$, $x_{i,t-1}(t) = x_{i,t-1}$ for all t. Hence, in this case, $\bar{F}$ is zero, and $\Omega = \Omega_1$.

For the DL(2) model, deriving the form of Ω is more complicated when $\lambda > 1$. Specifically, for $\lambda = 4$ and 12, we have that

$$f_{t+1} = \begin{cases} ((y_{t+1} - x_{1,t-1|1}\beta_{1,0})^2 - (y_{t+1} - x_{2,t-1|1}\beta_{2,0})^2) & \text{if at time } t \text{ there are no annual revisions,} \\ ((y_{t+1} - x_{1,t-1}\beta_{1,0})^2 - (y_{t+1} - x_{2,t-1}\beta_{2,0})^2) & \text{if at time } t \text{ there are annual revisions,} \end{cases}$$

implying that for $i = 1, 2$,

$$E\partial f_{t+1}/\partial \beta_{i,t} = \begin{cases} (-1)^i 2E((y_{t+1} - x_{i,t-1|1}\beta_{i,0})x_{i,t-1|1}) \neq 0 & \text{if at time } t \text{ there are no annual revisions,} \\ (-1)^i 2E((y_{t+1} - x_{i,t-1}\beta_{i,0})x_{i,t-1}) = 0 & \text{if at time } t \text{ there are annual revisions.} \end{cases}$$

After some algebra, we can show that $\bar{F} = \frac{\lambda-1}{\lambda}(-2\sigma^2_{w,x}\beta_{1,0}, 2\sigma^2_{w,x}\beta_{2,0})$. Thus, except if $\lambda = 1$, $\Omega \neq \Omega_1$ and we need to account for parameter estimation uncertainty.

Table 4 contains results for $\alpha = 0.05$, obtained with 10,000 Monte Carlo replications and 499 bootstrap replications each as in the previous section. Similarly, we set $R = P = 80$. Following Gonçalves, McCracken and Yao (2025), we let $\ell = \lfloor \min\{R^{1/3}, P^{1/3}\} \rfloor$ ensuring that $\ell / \min\{\sqrt{R}, \sqrt{P}\} \to 0$ as $\ell \to \infty$. The left panel of Table 4 shows tests size. Diebold and Mariano's (1995) tests are oversized, all well above 8%, except for the DL(2) model with $\lambda = 1$. In this particular case, the parameter estimation uncertainty does not contribute to the overall variance, implying that t_{DM} is asymptotically valid. The size results of Clark and McCracken (2009)'s test range from 2.7% to 6.9%, confirming the robustness of t_{CM} to annual revisions. The size results of $Bootstrap_R$ test range from 1.7% to 30.5%. As explained previously, this test is not valid when $\lambda > 1$ because it incorrectly enforces the regular revision structure on the bootstrap samples, failing to replicate the correct asymptotic variance when the forecasting model is DL(2). In contrast, the *Bootstrap* test produces much more reasonable size results, which range from 1.7% to 5.8%. The right panel of Table 4 shows the power results. All tests are negatively affected by annual revisions, but power increases as we increase P. See Table C.1 in Appendix C, which contains additional simulation results for $P = 320$.

8 Forecasting U.S. employment growth

In this section, we apply our bootstrap and the test statistic developed in Clark and McCracken (2009) to tests of equal forecast accuracy in the context of employment growth forecasting. While there are many existing papers that forecast employment growth, including Rapach and Strauss (2008, 2010) and Borup and Schütte (2022), most consider only current vintage data and hence do not address a more realistic situation in which data are subject to revision.[1]

With that in mind we compare the relative accuracy of a small handful of linear forecasting models for forecasting U.S. employment growth at 3- and 12-month horizons. The benchmark model is an AR(5) where the number of lags was selected using BIC. The competing models each take the form of an AR(5) augmented with a single lag of either initial unemployment claims (claims; monthly average of weekly data), total capacity utilization (TCU), or the vacancy rate (v/u).[2] Four of the series (employment, claims, TCU, vacancy) are subject to both regular and annual revisions. In contrast, the number of unemployed is only subject to an annual revision. All series are seasonally adjusted by the source. Vintages of TCU were obtained from the RTDSM hosted by the Federal Reserve Bank of Philadelphia and all other vintages from ALFRED hosted by the Federal Reserve Bank of St. Louis. Among our series, vacancies has the most limited history. Data for vacancies was first released by the BLS in December of 2000 based on the Job Openings and Labor Transition Survey (JOLTS). ALFRED has vintages associated with JOLTS starting in 2010:09 with observations dating back to 2000:12. Earlier vintages could be manually constructed using source data from the BLS. However, since we want a substantial number of observations available for parameter estimation at our first forecast origin, we have chosen to simply use the vintages in ALFRED.

[1] A notable exception is Barnichon and Nekarda (2012).

[2] In unreported results we also consider the vacancy-to-labor force ratio and the vacancy-to-(unemployed + all employed who make a job-to-job transition)-ratio, the latter of which is motivated by Birinci et al. (2024). In all experiments the vacancy rate provided more accurate forecasts and so for brevity we only report those results.

Within each vintage t we apply the following data transformations. Annualized monthly employment growth ($y_t^{[1]} = y_t$) is constructed as twelve times the log first difference of total nonfarm payroll employment. 3- and 12-month employment growth are constructed in the obvious way as $y_t^{[3]} = \sum_{j=0}^{2} y_{t-j}/3$ and $y_t^{[12]} = \sum_{j=0}^{11} y_{t-j}/12$ respectively. TCU and claims are transformed to be twelve-times the monthly first differences.

Constructing the vacancy rate requires a bit more thought. The issue is that the JOLTS data are released with a two-month lag. For many macroeconomic series like unemployment, the data is released with only a one-month lag so that the most recently reported value is associated with the previous month. In contrast, the most recent value for vacancies is from two-months previous. To insure that the concept of a vacancy rate is properly defined we construct it as $(v/u)_s(t) = v_s(t)/u_{s-1}(t)$ for all observations $s = 1, \ldots, t$ in vintage t.

For each forecast horizon $\tau = 3, 12$, the three OLS estimated forecasting models take the form

$$y_t^{[\tau]}(t) = \beta_0 + \sum_{j=1}^{5} \beta_j y_{t-\tau+1-j}(t) + \beta_6 x_{t-\tau}(t) + u_t^{[\tau]}(t)$$

where x is either omitted or denotes *claims*, *TCU*, or v/u. As noted earlier, due to data limitations associated with JOLTS data, we restrict attention to a sample starting in 2000:12 and hence all OLS regressions are estimated on a sample ranging from 2000:12 through a given forecast origin. Throughout, all forecasts are evaluated against the initial release $y_{t+\tau|1}^{[\tau]}$.

Table 5 provides the results of our forecasting exercise. The first two columns denote the six pairwise model comparisons while the remaining columns distinguish the forecast horizon. For each permutation of model comparison and horizon we report three numbers. The first denotes the ratio of root mean squared errors (RMSE) such that a value less than one favors model 1. The second number denotes the percentile bootstrap p-value (in parentheses) associated with the test of equal forecast accuracy under quadratic loss. For the same test, the third number is the p-value (in brackets) implied by the asymptotic distribution of the test statistic delineated in Clark and McCracken (2009).

Across both horizons there is little difference in forecast accuracy across models. The benchmark autoregressive model is nominally the most accurate in almost all cases but the gains are small and nearly always statistically insignificant when using either the asymptotic or bootstrap p-values. The sole exception arises at the 3-month horizon when using the vacancy rate as a predictor. Here the benchmark model is roughly 5 percent more accurate and significantly so at conventional levels. Interestingly, the only other signs of statistical significance arise when comparing the claims and TCU models to the vacancy rate. This is particularly true in the final row where we see that at the 3-month horizon, the TCU-

Table 5: **Application to Forecasting Employment Growth**

Model 1	**Model 2**	**Total Employment Growth**	
		$\boldsymbol{\tau = 3}$	$\boldsymbol{\tau = 12}$
AR(5)	AR(5) + claims	0.983	1.003
		(0.713)	(0.655)
		[0.649]	[0.691]
AR(5)	AR(5) + TCU	0.992	0.996
		(0.905)	(0.581)
		[0.633]	[0.558]
AR(5)	AR(5) + v/u	0.948	0.990
		(0.097)	(0.183)
		[0.007]	[0.302]
AR(5) + claims	AR(5) + TCU	1.009	0.993
		(0.757)	(0.621)
		[0.796]	[0.609]
AR(5) + claims	AR(5) + v/u	0.964	0.987
		(0.099)	(0.293)
		[0.331]	[0.301]
AR(5) + TCU	AR(5) + v/u	0.956	0.993
		(0.087)	(0.445)
		[0.026]	[0.510]

Notes: For each pairwise comparison, the table presents: the ratio of root mean squared errors, the p-value associated with a test of equal forecast accuracy based on our percentile bootstrap (in parentheses), and the p-value for the same test based on the asymptotic distribution associated with the test statistic delineated in Clark and McCracken (2009) (in square brackets). RMSE ratios less (greater) than one favor model 1 (2). Results are provided for an initial window size $R = 115$ and horizons $\tau = 3$ and $\tau = 12$, across 999 bootstrap replications. The forecast origins range from July 2010 to August 2024 for $\tau = 3$ and from July 2010 to November 2023 for $\tau = 12$.

augmented model is significantly better based on either set of p-values. While speculative, it may be that the delayed data release associated with JOLTS may have an impact on the short run predictive content of the vacancy rate for employment growth.

9 Conclusions

This paper derives the limiting distributions for West-type out-of-sample predictability tests when the underlying data are subject to annual revisions. Specifically, we show that these tests are still asymptotically normal with annually revised real-time data, but the asymptotic variance may be different than that obtained with regular revisions only. We then show that both the t-test of Clark and McCracken (2009) and the bootstrap test of Gonçalves, McCracken and Yao (2025) are robust to the change in asymptotic variance caused by the presence of annual revisions. Monte Carlo simulations confirm our analytical results. We conclude with an application to employment growth forecasting in the presence of real-time vintage data that exhibits both regular and annual benchmark revisions.

A Appendix: Asymptotic Theory

Throughout this Appendix and the next, we will write GMY(2025) to refer to Conçalves, McCracken and Yao (2025). For simplicity, we treat β as a scalar and focus on the case of a single model; we also let $T - R + 1 = P = N\lambda$, where λ is the periodicity of the annual revisions. Following West (1996), we write "$\sup_t$" to mean "$\sup_{R \leq t \leq T}$".

A.1 Auxiliary lemmas

Here, we provide several auxiliary lemmas, followed by their proofs.

Lemma A.1 *Under Assumptions 1-5,*

(a) $\sup_t |\hat{B}(t) - B(t)| = o_p(1)$.

(b) $P^{-1/2} \sum_{t=R}^{T} |\hat{H}(t) - H(t)| = o_p(1)$.

(c) *For any* $0 \leq a < 1/2$, $\sup_t P^a |\hat{H}(t) - H(t)| = o_p(1)$.

(d) *For any* $0 \leq a < 1/2$, $\sup_t |P^a(\hat{\beta}(t) - \beta_0)| = o_p(1)$.

Lemma A.2 *Under Assumptions 1-5,*

(a) $P^{-1/2} \sum_{t=R}^{T} (\frac{\partial f_{t+\tau}}{\partial \beta} - F_{t+\tau}) B H(t) = o_p(1)$.

(b) $P^{-1/2} \sum_{t=R}^{T} F_{t+\tau} (B(t) - B) H(t) = o_p(1)$.

(c) $P^{-1/2} \sum_{t=R}^{T} (\frac{\partial f_{t+\tau}}{\partial \beta} - F_{t+\tau})(B(t) - B) H(t) = o_p(1)$.

(d) $\tilde{S}_P^{\mu} = P^{-1/2} \sum_{t=R}^{T} (f_{t+\tau} - E f_{t+\tau}) + F_{t+\tau} B H(t)) + o_p(1)$.

Lemma A.3 *Define* $\hat{Z}_t \equiv (\hat{f}_t - \bar{f}, \hat{h}_t)'$, *with* $\hat{f}_t \equiv f(\hat{\beta}_T)$, $\bar{f} \equiv P^{-1} \sum_{t=R}^{T} \hat{f}_t$, *and* $\hat{h}_t \equiv h(\hat{\beta}_T)$. *Then, under Assumptions 1-7, and if* H_0 *is true,*

$$\frac{1}{P} \sum_{t=R}^{T} \sum_{s=R}^{T} \hat{Z}_t \hat{Z}_s' K((t-s)/b_P) - \Omega_Z \xrightarrow{p} 0,$$

where

$$\Omega_Z \equiv \begin{pmatrix} Var(P^{-1/2} \sum_{t=R}^{T} f_t) & Cov(P^{-1/2} \sum_{t=R}^{T} f_t, P^{-1/2} \sum_{s=R}^{T} h_s) \\ Cov(P^{-1/2} \sum_{t=R}^{T} f_t, P^{-1/2} \sum_{s=R}^{T} h_s) & Var(P^{-1/2} \sum_{t=R}^{T} h_t) \end{pmatrix}.$$

Proof of Lemma A.1. The proof of this result follows the same arguments as that of Lemma A.1 in GMY(2025). The main difference is that we now rely on Assumption 4 which assumes that the total number of revisions r_{max} is finite, where $r_{max} = r + r_b - 1 = r$ only when there are no annual revisions, i.e. $r_b - 1 = 0$. ■

Proof of Lemma A.2. We adapt the proof of Lemma A4 in West (1996) to allow for periodic stationarity. The crucial insight is that we can write $\sum_{t=R}^{T}$ as a double sum $\sum_{j=0}^{\lambda-1}\sum_{n=0}^{N-1}$, where we assume for simplicity that $P = N\lambda$. Since $\lambda < \infty$, we can then apply the arguments in West (1996) to each summand indexed by j by relying on our new set of assumptions, in particular Assumption 3.

Part (a). Define $D_t \equiv \partial f_{t+\tau}/\partial\beta - F_{t+\tau}$, where $F_{t+\tau} \equiv E(\partial f_{t+\tau}/\partial\beta)$. For $j = 0, 1, \ldots, \lambda$, let $f_{t+\tau}^{(j+1)}$ denote the j^{th} version of $f_{t+\tau}$ and define $D_t^{(j+1)} = \frac{\partial f_{t+\tau}^{(j+1)}}{\partial\beta} - F^{(j+1)}$, where $F^{(j+1)} \equiv E(\partial f_{t+\tau}^{(j+1)}/\partial\beta)$. Note that under Assumption 3, $F^{(j+1)}$ does not depend on t. Using the definition of $H(t) \equiv t^{-1}\sum_{s=1+\tau+\ddot{r}}^{t} h_s$, we can write

$$\begin{aligned}
P^{-1/2}\sum_{t=R}^{T} D_{t+\tau}BH(t) &= P^{-1/2}\sum_{t=R}^{T}(D_{t+\tau}Bt^{-1}(h_{1+\tau+\ddot{r}} + \ldots + h_t)) \\
&= \sum_{j=0}^{\lambda-1} P^{-1/2}\sum_{n=0}^{N-1}(R+n\lambda+j)^{-1}D_{R+\tau+n\lambda+j}(h_{R+n\lambda+j} + \ldots + h_{1+\tau+\ddot{r}}) \\
&\equiv \sum_{j=0}^{\lambda-1}\mathcal{A}^{(j+1)}.
\end{aligned}$$

We can show that for each $j = 0, \ldots, \lambda - 1$ with $\lambda < \infty$, $\mathcal{A}^{(j+1)} = o_p(1)$ by Chebyshev's inequality. For instance, consider $j = 0$. It suffices to show that (i) $E\mathcal{A}^{(1)} \to 0$ and (ii) $Var(\mathcal{A}^{(1)}) \to 0$. To show (i), let $\gamma_i^{(1)} \equiv ED_t^{(1)}Bh_{t-i} = ED_t^{(1)}h_{t-i}$, where we redefine Bh_t as h_t. With this notation, we can write

$$\begin{aligned}
E\mathcal{A}^{(1)} &= P^{-1/2}\sum_{n=0}^{N-1}(R+n\lambda)^{-1}(\gamma_{\tau}^{(1)} + \gamma_{\tau+1}^{(1)} + \ldots + \gamma_{R+n\lambda-1-\ddot{r}}^{(1)}) \\
&\leq P^{-1/2}\sum_{n=0}^{N-1}(R+n\lambda)^{-1}\sum_{i=\tau}^{\infty}|\gamma_i^{(1)}| \\
&\leq P^{-1/2}\sum_{n=0}^{N-1}\sum_{j=0}^{\lambda-1}(R+n\lambda+j)^{-1}\sum_{i=0}^{\infty}|\gamma_i^{(1)}|,
\end{aligned}$$

where $P^{-1/2}\sum_{n=0}^{N-1}\sum_{j=0}^{\lambda-1}(R+n\lambda+j)^{-1} = P^{-1/2}\sum_{t=R}^{T}t^{-1}$ which converges to zero by Lemma A1(a) in West (1996), and $\sum_{i=0}^{\infty}|\gamma_i^{(1)}| < \infty$ by Lemma A2 in West (1996). Note in particular the use of the stationarity and mixing assumptions on $\frac{\partial f_{t+\tau}^{(j+1)}}{\partial\beta}$ for $j = 0$ implied by Assumption 3. The proof of (ii) follows similarly as in West (1996) using Assumption 3(a) which bounds the fourth moments of $(D_t^{(1)}, h_t)'$.

Part (b). Note that

$$P^{-1/2}\sum_{t=R}^{T}F_{t+\tau}(B(t)-B)H(t)\leq\sup_t|B(t)-B|P^{-1/2}\sum_{t=R}^{T}|F_{t+\tau}H(t)|$$
$$\leq\sup_t|B(t)-B|\sup_{1\leq j\leq\lambda}|F^{(j)}|P^{-1/2}\sum_{t=R}^{T}|H(t)|$$

where $\sup_t|B(t)-B|=o_p(1)$ by Assumption 2, $\sup_{1\leq j\leq\lambda}|F^{(j)}|=O(1)$ by Assumption 3 and $\lambda<\infty$, and $P^{-1/2}\sum_{t=R}^{T}|H(t)|=O_p(1)$ by the proof of Lemma A4 (c) in West (1996).

Part (c). First, write

$$P^{-1/2}\sum_{t=R}^{T}(\frac{\partial f_{t+\tau}}{\partial\beta}-F_{t+\tau})(B(t)-B)H(t)\leq\sup_t|B(t)-B|P^{-1/2}\sum_{t=R}^{T}|(\frac{\partial f_{t+\tau}}{\partial\beta}-F_{t+\tau})H(t)|,$$

where $\sup_t|B(t)-B|=o_p(1)$ by Assumption 2. The result follows if $P^{-1/2}\sum_{t=R}^{T}|(\frac{\partial f_{t+\tau}}{\partial\beta}-F_{t+\tau})H(t)|=O_p(1)$. Using the periodicity assumption on f_t (see equation (1) in the main text), we can write

$$P^{-1/2}\sum_{t=R}^{T}\Big|(\frac{\partial f_{t+\tau}}{\partial\beta}-F_{t+\tau})H(t)\Big|=P^{-1/2}\sum_{n=0}^{N-1}\sum_{j=0}^{\lambda-1}\Big|(\frac{\partial f^{(j+1)}_{R+\tau+n\lambda+j}}{\partial\beta}-F^{(j+1)})H(R+n\lambda+j)\Big|$$
$$\leq\sum_{j=0}^{\lambda-1}P^{-1/2}\sum_{t=R}^{T}\Big|(\frac{\partial f^{(j+1)}_{t+\tau}}{\partial\beta}-F^{(j+1)})H(t)\Big|,$$

where $\lambda<\infty$, and $P^{-1/2}\sum_{t=R}^{T}\Big|(\frac{\partial f^{(j+1)}_{t+\tau}}{\partial\beta}-F^{(j+1)})H(t)\Big|=O_p(1)$ by logic such as that in the proof of Lemma A4 (b) in West (1996), applied here to each $j=0,1,\ldots,\lambda$.

Part (d). Applying a second-order mean value expansion of $f_{t+\tau}(\hat\beta_t)$ around β_0,

$$\tilde S^{\mu}_P=P^{-1/2}\sum_{t=R}^{T}(f_{t+\tau}-Ef_{t+\tau})+\xi_1+\xi_2,$$

with $\xi_1=P^{-1/2}\sum_{t=R}^{T}\frac{\partial f_{t+\tau}}{\partial\beta}(\hat\beta_t-\beta_0)$ and $\xi_2=0.5P^{-1/2}\sum_{t=R}^{T}\frac{\partial^2}{\partial\beta^2}f_{t+\tau}(\tilde\beta_t)(\hat\beta_t-\beta_0)^2$, where $\tilde\beta_t$ lies between $\hat\beta_t$ and β_0. The result follows if $\xi_1=P^{-1/2}\sum_{t=R}^{T}F_{t+\tau}BH(t)+o_p(1)$ and $\xi_2=o_p(1)$. To show $\xi_2=o_p(1)$, note that

$$|\xi_2|\leq 0.5(\sup_t\big|P^{1/4}(\hat\beta_t-\beta_0)\big|)^2P^{-1}\sum_{t=R}^{T}\big|\frac{\partial^2}{\partial\beta^2}f_{t+\tau}(\tilde\beta_t)\big|,$$

where $\sup_t|P^{1/4}(\hat\beta_t-\beta_0)|=o_p(1)$ by Lemma A3 (b) in West (1996). The result follows since we can show that $P^{-1}\sum_{t=R}^{T}|\frac{\partial^2}{\partial\beta^2}f_{t+\tau}(\tilde\beta_t)|=O_p(1)$, as we argue next. Using equation (1) in the main text, we can write

$$P^{-1}\sum_{t=R}^{T}|\frac{\partial^2}{\partial\beta^2}f_{t+\tau}(\tilde\beta_t)|=P^{-1}\sum_{j=0}^{\lambda-1}\sum_{n=0}^{N-1}|\frac{\partial^2}{\partial\beta^2}f^{(j+1)}_{R+\tau+n\lambda+j}(\tilde\beta_{R+n\lambda+j})|.$$

By Assumption 1, and the fact that $\tilde{\beta}_t \xrightarrow{p} \beta_0$, for $j = 0, \ldots, \lambda-1$, we can bound $|\frac{\partial^2}{\partial \beta^2} f_{t+\tau}^{(j+1)}(\tilde{\beta}_t)|$ by $\sup_{\beta \in \mathcal{Z}} |\frac{\partial^2}{\partial \beta^2} f_{t+\tau}^{(j+1)}| < m_{t+\tau}$. The result follows by Markov's inequality since $Em_{t+\tau} < D < \infty$ for all $j = 0, 1, \ldots, \lambda$. For ξ_1, adding and subtracting appropriately yields

$$\xi_1 = P^{-1/2} \sum_{t=R}^{T} \frac{\partial f_{t+\tau}}{\partial \beta} (\hat{\beta}_t - \beta_0) = \sum_{i=1}^{4} \xi_{1.i}$$

where

$$\xi_{1.1} = P^{-1/2} \sum_{t=R}^{T} F_{t+\tau} BH(t), \quad \xi_{1.2} = P^{-1/2} \sum_{t=R}^{T} (\frac{\partial f_{t+\tau}}{\partial \beta} - F_{t+\tau}) BH(t)$$

$$\xi_{1.3} = P^{-1/2} \sum_{t=R}^{T} F_{t+\tau} (B(t) - B) H(t), \text{ and } \xi_{1.4} = P^{-1/2} \sum_{t=R}^{T} (\frac{\partial f_{t+\tau}}{\partial \beta} - F_{t+\tau}) (B(t) - B) H(t)$$

The result follows by parts (a), (b) and (c) since $\xi_{1.i} = o_p(1)$ for $i = 2, 3, 4$, respectively. ■

Proof of Lemma A.3. Let

$$\tilde{\Omega}_Z = P^{-1} \sum_{t=R}^{T} \sum_{s=R}^{T} \tilde{Z}_t \tilde{Z}_s' K((s-t)/b_P) \quad \text{and} \quad \hat{\Omega}_Z = P^{-1} \sum_{t=R}^{T} \sum_{s=R}^{T} \hat{Z}_t \hat{Z}_s' K((s-t)/b_P),$$

where $\tilde{Z}_t = (\hat{f}_t, \hat{h}_t)' \equiv (f_t(\hat{\beta}_T), h_t(\hat{\beta}_T))'$ and $\hat{Z}_t = (\hat{f}_t - \bar{f}, \hat{h}_t)'$, with $\bar{f} = P^{-1} \sum_{t=R}^{T} \hat{f}_t$. Note that $\tilde{Z}_t$ differs from $\hat{Z}_t$ because of the demeaning of $\hat{f}_t$. The proof contains two steps: (i) show that $\tilde{\Omega}_Z - \Omega_Z \to_p 0$; (ii) show that $\hat{\Omega}_Z - \tilde{\Omega}_Z \to_p 0$. Starting with (i), define $X_t(\hat{\beta}_T) \equiv P^{-1/2} \tilde{Z}_t$ and $X_t(\beta_0) \equiv P^{-1/2} Z_t(\beta_0)$ where $Z_t(\beta_0) \equiv (f_t, h_t)'$. We next show that $X_t \equiv X_t(\beta_0)$ satisfies Assumptions 1-4 of Theorem 2.2 of de Jong and Davidson (2000), which implies (i). First, note that X_t has mean zero by Assumption 2, which implies $Eh_s = 0$, and by imposing H_0 since then $Ef_t = 0$. Second, note that their Assumption 1 is verified by our Assumption 7. Next, we can show that their Assumption 2 holds for X_t. In particular, we can show that X_t is L_2-NED (near epoch dependent) on a mixing process such that the regularity conditions in de Jong and Davidson's (2000) Assumption 2 are verified. To see this, let $\mathcal{V}_t = (f_t^{(1)}, f_t^{(2)}, \ldots, f_t^{(\lambda)}, h_t)'$. Under our Assumption 3, $\mathcal{V}_t$ is a strong mixing (covariance stationary) process of size $-\psi = -3d/(d-1)$, where $d > 1$. Letting $r = 1 + \epsilon$ for small enough $\epsilon > 0$ implies that $-\psi = -3(1+\epsilon)/\epsilon$, which is larger in absolute value than $-(2+\epsilon)/\epsilon$, the size condition imposed by Assumption 2 in de Jong and Davidson (2000). Hence, $\mathcal{V}_t$ is a strong mixing process satisfying the size conditions imposed on the array V_{nt} defined in Jong and Davidson (2000). Recall that X_t is said to be L_2-NED on $\mathcal{V}_t$ if or $m \geq 0$, $\|X_t - E(X_t|\mathcal{V}_{t-m}, \ldots, \mathcal{V}_{t+m}\|_2 \leq d_t v(m)$, where d_t is a nonstochastic sequence and $v(m) \to 0$ as $m \to \infty$. Since X_t is measurable with respect to $\mathcal{V}_{t-m}, \ldots, \mathcal{V}_{t+m}$, $E(X_t|\mathcal{V}_{t-m}, \ldots, \mathcal{V}_{t+m}) = 0$ for any m, implying that X_t is automatically NED on the mixing process $\mathcal{V}_t$. We can set the constant d_t that appears in Assumption 2 of de Jong and Davidson (2000) to any arbitrary value, including $d_t = P^{-1/2}$. Similarly, we can take $c_t = P^{-1/2}$ since $\|X_t/c_t\|_2$ is uniformly bounded under Assumption 3. This implies that conditions (2.6) and (2.7) of de Jong and Davidson (2000) are verified in our context. Assumption 3 of de Jong and Davidson (2000) follows if $b_P \to \infty$ such that $P^{-1}b_P \to 0$ given that $c_t = P^{-1/2}$. This condition is ensured by our rate condition in Assumption 7. Finally, we verify Assumption 4 of de Jong and Davidson (2000). Part (a) requires $\sqrt{T}(\hat{\beta}_T - \beta_0) = O_p(1)$, which follows under our Assumptions 2 and 3. Part (b) requires that $P^{-1/2}\sum_{t=R}^{T} E(\partial X_t(\beta)/\partial\beta)$ is continuous at β_0 uniformly in P. Since $X_t(\beta) = P^{-1/2}(f_t(\beta), h_t(\beta))'$ with $h_t(\beta) = x_{t-\tau}(y_t - x_{t-\tau}\beta)$, it follows that $\partial X_t(\beta)/\partial\beta = (\partial f_t(\beta)/\partial\beta, -x_{t-\tau}^2)'$ when β is a scalar. Hence, condition (b) holds in our context if $P^{-1}\sum_{t=R}^{T} E(\partial f_t(\beta)/\partial\beta)$ is continuous at β_0 uniformly in P. Given the periodic stationarity of $\partial f_t(\beta)/\partial\beta$, this condition holds if $\lambda^{-1}\sum_{j=1}^{\lambda} E(\partial f_t^{(j)}(\beta)/\partial\beta)$ is continuous at β_0. This follows from the twice continuity assumption on $f_t^{(j)}(\beta)$ imposed by Assumption 1(a) and the dominance condition of Assumption 1(b). To end the proof, we verify equations (2.8) and (2.10) in Assumption 4 of de Jong and Davidson (2000). Starting with (2.8), note that in our context this condition is

$$\begin{aligned}
&\lim_{R,P\to\infty} \sup_{R\leq T} \sum_{t=R}^{T} E \sup_{\beta\in\mathcal{N}} \|\partial X_t(\beta)/\partial\beta\|^2 \\
&\lim_{R,P\to\infty} \sup_{R\leq T} \frac{1}{P}\sum_{t=R}^{T} E \sup_{\beta\in\mathcal{N}} ((\partial f_t(\beta)/\partial\beta)^2 + x_{t-\tau}^4) \\
&\leq \lim_{R,P\to\infty} \sup_{R\leq T} \frac{1}{P}\sum_{t=R}^{T} E \sup_{\beta\in\mathcal{N}} (\partial f_t(\beta)/\partial\beta)^2 + \lim_{R,P\to\infty} \sup_{R\leq T} \frac{1}{P}\sum_{t=R}^{T} E(x_t^4) < \infty,
\end{aligned}$$

since for $t = R, \ldots, T$, $E(x_t^4) < \infty$ by Assumption 4 and we assume that $E\sup_{\beta\in\mathcal{N}}(\partial f_t/\partial\beta)^2 < \infty$. Finally, equation (2.10) holds by letting $P^{-1/2}b_P \to 0$ as in Assumption 7(b), and applying Markov's inequality to bound

$$\sup_{\beta\in\mathcal{N}} |\sum_{t=R}^{T} \partial X_t'(\beta)/\partial\beta \cdot \partial X_t(\beta)/\partial\beta| \leq \frac{1}{P}\sum_{t=R}^{T} \sup_{\beta\in\mathcal{N}} (\partial f_t(\beta)/\partial\beta)^2 + \frac{1}{P}\sum_{t=R}^{T} x_t^4 = O_p(1).$$

To prove step (ii), note that $\hat{\Omega}_Z - \tilde{\Omega}_Z$ is a matrix whose elements are all zero except for the element (1,1) which equals $-\bar{f}P^{-1}\sum_{s,t=R}^{T}(\hat{f}_t + \hat{f}_s)K((s-t)/b_P) + \bar{f}^2 P^{-1}\sum_{s,t=R}^{T} K((s-t)/b_P)$. Since $P^{-1}\sum_{s,t=R}^{T} K((s-t)/b_P) = O(b_P)$ and $\bar{f} = O_p(P^{-1/2})$ (when $E(f_t) = 0$), the last term is of order $O_p(b_P/P) = o_p(1)$ under our assumptions on the bandwidth. A similar argument applies to the first term, which implies that $\hat{\Omega}_Z - \tilde{\Omega}_Z = o_p(1)$. ■

A.2 Proofs of asymptotic results in the paper

Proof of Lemma 5.1. By two mean value expansions of $f_{t+\tau}(\hat{\beta}_t)$ and $f_{t+\tau}(\hat{\beta}(t))$ around β_0, respectively, we can write $\hat{S}_P^\mu - \tilde{S}_P^\mu = P^{-1/2}\sum_{t=R}^T F_{t+\tau}(\hat{\beta}(t)-\hat{\beta}_t) + o_p(1)$. The result then follows by showing that $P^{-1/2}\sum_{t=R}^T F_{t+\tau}(\hat{\beta}(t)-\hat{\beta}_t) = o_p(1)$. In particular, using the definitions of $\hat{\beta}_t$ and $\hat{\beta}(t)$, we can write

$$P^{-1/2}\sum_{t=R}^T F_{t+\tau}(\hat{\beta}(t)-\hat{\beta}_t) = \sum_{i=1}^3 \mathcal{C}_i,$$

where $\mathcal{C}_1 = P^{-1/2}\sum_{t=R}^T F_{t+\tau}(\hat{B}(t)-B(t))H(t)$, $\mathcal{C}_2 = P^{-1/2}\sum_{t=R}^T F_{t+\tau}B(t)(\hat{H}(t)-H(t))$, and $\mathcal{C}_3 = P^{-1/2}\sum_{t=R}^T F_{t+\tau}(\hat{B}(t)-B(t))(\hat{H}(t)-H(t))$. The result follows by showing that $\mathcal{C}_i = o_p(1)$ for $i=1,2,3$, which follows from Lemma A.1 and standard inequalities under our assumptions. ■

Proof of Lemma 5.2. Using Lemma A.2 (d), and adding and subtracting appropriately,

$$\tilde{S}_P^\mu = P^{-1/2}\sum_{t=R}^T (f_{t+\tau} - Ef_{t+\tau}) + \bar{F}BP^{-1/2}\sum_{t=R}^T H(t) + S_D + o_p(1),$$

where $S_D \equiv P^{-1/2}\sum_{t=R}^T (F_{t+\tau}-\bar{F})BH(t)$. We complete the proof by showing that $S_D = o_p(1)$ under our assumptions. Note that under covariance stationarity of $f_{t+\tau}$ and $\partial f_{t+\tau}/\partial\beta'$, $S_D = 0$ because $F_{t+\tau} = F = \bar{F}$, but this is not necessarily true under our new set of assumptions. As it turns out, the special periodic structure of $F_{t+\tau}$ (according to which it can take up to λ different values, where $\lambda < \infty$) is crucial in showing that $S_D = o_p(1)$. More specifically, using the definition of $H(t) \equiv t^{-1}\sum_{s=1+\tau+\ddot{r}}^t h_s$, we can write

$$\begin{aligned}
S_D &\equiv P^{-1/2}\sum_{t=R}^T (F_{t+\tau}-\bar{F})BH(t) \\
&= P^{-1/2}\Big[R^{-1}(F_{R+\tau}-\bar{F})B(\sum_{s=1+\tau+\ddot{r}}^R h_s) + (R+1)^{-1}(F_{R+1+\tau}-\bar{F})B\big((\sum_{s=1+\tau+\ddot{r}}^R h_s) + h_{R+1}\big) + \ldots\Big] \\
&= \mathcal{D}_{R,0}BP^{-1/2}\sum_{s=1+\tau+\ddot{r}}^R h_s + P^{-1/2}\sum_{i=1}^{P-1}\mathcal{D}_{R,i}Bh_{R+i} \\
&\equiv S_{D.1} + S_{D.2},
\end{aligned}$$

where $\mathcal{D}_{R,i} = (R+i)^{-1}(F_{R+\tau+i}-\bar{F}) + \ldots + T^{-1}(F_{T+\tau}-\bar{F})$. This is similar to West (1996)'s decomposition of $\sum_{t=R}^T H(t) = a_{R,0}[\sum_{s=1+\tau+\ddot{r}}^R h_s] + a_{R,1}h_{R+1} + \ldots + a_{R,P-1}h_T$, $a_{R,i} = (R+i)^{-1} + \ldots + T^{-1}$, but our weights $\mathcal{D}_{R,i}$ include the factors $(F_{R+\tau+i}-\bar{F})$.

Next, we show that both $S_{D.1}$ and $S_{D.2}$ vanish asymptotically. For $S_{D.1} = o_p(1)$, we only need to show that $\mathcal{D}_{R,0} = o(1)$ since $BP^{-1/2}\sum_{s=1+\tau+\ddot{r}}^R h_s = O_p(1)$ by the strong mixing assumption on h_t of Assumption 3. To show that $\mathcal{D}_{R,0} = o(1)$ we use the periodicity of $F_{t+\tau}$. Specifically, we can write

$$\mathcal{D}_{R,0} = \sum_{t=R}^{T} t^{-1}(F_{t+\tau}-\bar{F}) = \sum_{n=0}^{N-1}\sum_{j=0}^{\lambda-1}(R+n\lambda+j)^{-1}(F_{R+\tau+n\lambda+j}-\bar{F}) = \sum_{j=0}^{\lambda-1}(F^{(j+1)}-\bar{F})\sum_{n=0}^{N-1}(R+n\lambda+j)^{-1}.$$

Note that $\sum_{j=0}^{\lambda-1}(F^{(j+1)} - \bar{F}) = 0$ by definition of $\bar{F}$, hence $\mathcal{D}_{R,0} \to 0$ if we can show that for each fixed j, $\sum_{n=0}^{N-1}(R + n\lambda + j)^{-1}$ converges to a constant that is independent of j. For simplicity, we let $a_j = \frac{R+j}{\lambda}$ and note that a_j is proportional to $R \to \infty$ for fixed j and fixed λ. Then for each $j = 0, \ldots, \lambda - 1$, we can write

$$\sum_{n=0}^{N-1} \frac{1}{R + n\lambda + j} = \frac{1}{\lambda}\sum_{n=0}^{N-1} \frac{1}{a_j + n}.$$

We can show that the limit of $\lambda^{-1}\sum_{n=0}^{N-1}(a_j+n)^{-1}$ is $\lambda^{-1}\ln(1+\pi)$ (the argument is analogous to West's derivation of the limit of $a_{R,0} = \sum_{t=R}^{T} t^{-1} = \sum_{n=0}^{P-1}(R + n)^{-1}$, which is $\ln(1 + \pi)$, as he shows in p. 1082). In particular, for each j,

$$\frac{1}{\lambda}\int_0^{N-1} \frac{1}{a_j + z}dz \le \frac{1}{\lambda}\sum_{n=0}^{N-1} \frac{1}{a_j + n} \le \frac{1}{\lambda}\int_{-1}^{N-1} \frac{1}{a_j + z}dz$$

$$\Rightarrow \quad \frac{1}{\lambda}\ln\Big(\frac{a_j + N - 1}{a_j}\Big) \le \frac{1}{\lambda}\sum_{i=0}^{N-1} \frac{1}{a_j + n} \le \frac{1}{\lambda}\ln\Big(\frac{a_j + N - 1}{a_j - 1}\Big)$$

$$\Rightarrow \quad \frac{1}{\lambda}\ln\Big(1 + \frac{N}{a_j} - \frac{1}{a_j}\Big) \le \frac{1}{\lambda}\sum_{n=0}^{N-1} \frac{1}{a_j + n} \le \frac{1}{\lambda}\ln\Big(1 + \frac{N}{a_j - 1}\Big),$$

where for fixed j, $\frac{N}{a_j} = \frac{N\lambda}{R+j} = \frac{P}{R+j} = (\frac{R}{P} + \frac{j}{P})^{-1} \to \pi$, $\frac{1}{a_j} = \frac{\lambda}{R+j} \to 0$, $\frac{N}{a_j-1} \to \pi$, and $\pi = \lim P/R$. Hence, $\frac{1}{\lambda}\sum_{n=0}^{N-1}\frac{1}{a_j+n} \to \frac{1}{\lambda}\ln(1 + \pi)$. By Assumption 5, $0 \le \pi < \infty$, implying that $0 \le \ln(1 + \pi) < \infty$, independently of j. Hence,

$$\lim_{R,P\to\infty} \mathcal{D}_{R,0} = \ln(1 + \pi)\lambda^{-1}\sum_{j=0}^{\lambda-1}(F^{(j+1)} - \bar{F}) = 0,$$

since $\lambda^{-1}\ln(1 + \pi)$ is a constant and $\lambda^{-1}\sum_{j=0}^{\lambda-1}(F^{(j+1)} - \bar{F}) = 0$.

For $S_{D.2} \equiv P^{-1/2}\sum_{i=1}^{P-1}\mathcal{D}_{R,i}Bh_{R+i} = o_p(1)$, we use Chebyshev's inequality. By Assumption 2, $Eh_{R+i} = 0$, which implies that $ES_{D.2} = 0$. Hence, it suffices to show that $Var(S_{D.2}) \to 0$. Let

$$d_j = \mathcal{D}_{R,1}\mathcal{D}_{R,j+1} + \ldots + \mathcal{D}_{R,P-j-1}\mathcal{D}_{R,P-1} \quad \text{for} \quad 0 \le j \le P - 2.$$

For $-P + 2 \le j < 0$, $d_j \equiv d_{-j}$. Using the same logic as in the proof of equation (A-1b) of West (1996), we can write

$$Var(S_{D.2}) = P^{-1}\sum_{j=-P+2}^{P-2} d_j\gamma_j \le P^{-1}\sum_{j=-P+2}^{P-2} |d_j||\gamma_j|$$

where $\gamma_j = Cov(Bh_t, Bh_{t+j}) = Cov(Bh_t, Bh_{t-j})$. Note that for any $0 \le j \le P-2$,

$$|d_j| \le \max_{0\le j\le P-2} |d_j| \le \max_{0\le j\le P-2} \{|\mathcal{D}_{R,1}||\mathcal{D}_{R,j+1}| + \ldots + |\mathcal{D}_{R,P-j-1}||\mathcal{D}_{R,P-1}|\} \le P(\max_{0\le j\le P-1} |\mathcal{D}_{R,j}|)^2.$$

Then

$$Var(S_{D.2}) \le (\max_{0\le j\le P-1} |\mathcal{D}_{R,j}|)^2 \sum_{j=-P+2}^{P-2} |\gamma_j|$$

where $\sum_{j=-\infty}^{\infty} |\gamma_j| < \infty$. Hence $\lim Var(S_{D.2}) = o(1)$ if $\max_{0\le j\le P-1} |\mathcal{D}_{R,j}| \to 0$. Because for $0 \le j \le P-1$, $\mathcal{D}_{R,j}$ is deterministic, we can find $j = m$ such that $|\mathcal{D}_{R,m}| = \max_{0\le j\le P-1} |\mathcal{D}_{R,j}|$. Using the triangle inequality, we can write

$$|\mathcal{D}_{R,m}| \le |\mathcal{D}_{R,0}| + |\mathcal{D}_{R,0} - \mathcal{D}_{R,m}|.$$

Hence, for $0 \le m \le P-1$, $|\mathcal{D}_{R,m}| \to 0$ if the RHS of the above inequality goes to zero. When $m = 0$, $|\mathcal{D}_{R,0}| = o(1)$ as shown above. When $1 \le m \le \lambda$,

$$|\mathcal{D}_{R,0} - \mathcal{D}_{R,m}| = \Big|\sum_{t=R}^{R+m-1} t^{-1}(F_{t+\tau} - \bar{F})\Big| \le R^{-1} \sum_{t=R}^{R+m-1} |F_{t+\tau} - \bar{F}| \le O(\lambda/R) \to 0.$$

When $\lambda < m \le P-1$, there must exist $1 \le I_m \le N-1$ and $0 \le J_m \le \lambda - 1$ $s.t.$ $R+m-1 = R + I_m\lambda + J_m$. For example, when $m = \lambda + 1$, $I_m = 1$ and $J_m = 0$; when $m = P-1$ then $I_m = N-1$, and $J_m = \lambda - 1$. Hence, we can write

$$\begin{aligned}
|\mathcal{D}_{R,0} - \mathcal{D}_{R,m}| &= \Big|\sum_{t=R}^{R+m-1} t^{-1}(F_{t+\tau} - \bar{F})\Big| \\
&= \Big|\sum_{n=0}^{I_m-1}\sum_{j=0}^{\lambda-1}(R+n\lambda+j)^{-1}(F_{R+\tau+n\lambda+j} - \bar{F}) + \sum_{j=0}^{J_m}(R+I_m\lambda+j)^{-1}(F_{R+\tau+I_m\lambda+j} - \bar{F})\Big| \\
&\le \Big|\sum_{n=0}^{I_m-1}\sum_{j=0}^{\lambda-1}(R+n\lambda+j)^{-1}(F_{R+\tau+n\lambda+j} - \bar{F})\Big| + \frac{1}{R}\sum_{j=0}^{J_m} |F_{R+\tau+I_m\lambda+j} - \bar{F}|
\end{aligned}$$

where the second term is $O(\lambda/R) \to 0$, and the first term can be written as

$$\Big|\sum_{j=0}^{\lambda-1}(F^{(j+1)} - \bar{F}) \sum_{n=0}^{I_m-1} \frac{1}{R+n\lambda+j}\Big|.$$

Recall that λ is finite and $\sum_{j=0}^{\lambda}(F^{(j+1)} - \bar{F}) = 0$. Then $|\mathcal{D}_{R,0} - \mathcal{D}_{R,m}| = o(1)$ if $\sum_{n=0}^{I_m-1} \frac{1}{R+n\lambda+j} arrowc$ where c is a constant that is independent of j. This can be shown using arguments similar to those used above to show that $\sum_{n=0}^{N-1}(R+n\lambda+j)^{-1} \to \ln(1+\pi)$, so we skip the details.

■

Proof of Theorem 5.1. The proof follows from the continuous mapping theorem if

$$\begin{pmatrix} S_{1P} \\ BS_{2P} \end{pmatrix} \xrightarrow{d} N(0,V) \text{ for a full rank matrix } V = \begin{pmatrix} \Omega_1 & \Omega_{12}B' \\ & \\ B\Omega_{12}' & B\Omega_2 B' \end{pmatrix}.$$

This result follows by showing that the vector $(S_{1P}, S_{2P}'B')'$ satisfies the regularity conditions of the central limit theorem for near-epoch dependent functions of the heterogeneous mixing process in de Jong (1997). In particular, we use the periodic structure of f_t to rewrite S_{1P} as

$$S_{1P} = N^{-1/2}\sum_{n=0}^{N-1}(\lambda^{-1/2}\sum_{j=0}^{\lambda-1}(f_{R+\tau+n\lambda+j}^{(j+1)} - Ef_{R+\tau+n\lambda+j}^{(j+1)})) \equiv N^{-1/2}\sum_{n=0}^{N-1}(\mathcal{Z}_n - E\mathcal{Z}_n),$$

where $\mathcal{Z}_n \equiv \lambda^{-1/2}\sum_{j=0}^{\lambda-1} f_{R+\tau+n\lambda+j}^{(j+1)}$ is the scaled average of the λ versions of $f_{t+\tau}$. Since we assume that $(f_t^{(1)}, f_t^{(2)}, \ldots, f_t^{(\lambda)})$ is covariance stationary and strong mixing, it follows that for finite λ, $\mathcal{Z}_n - E\mathcal{Z}_n$ is strong mixing (hence, near-epoch dependent on a mixing process), and covariance stationary. ■

Proof of Lemma 5.3. Letting $\sum_{t=R}^{T} H(t) = a_{R,0}[\sum_{s=1+\tau+\ddot{r}}^{R} h_s] + \sum_{i=1}^{P-1} a_{R,i}h_{R+i}$, where $a_{R,i} = (R+i)^{-1} + \ldots + T^{-1}$, we can write $S_{2P} \equiv P^{-1/2}\sum_{t=R}^{T} H(t) = S_{2P.1} + S_{2P.2}$, where

$$S_{2P.1} \equiv P^{-1/2}\sum_{s=1+\tau+\ddot{r}}^{R} a_{R,0}h_s \quad \text{and} \quad S_{2P.2} \equiv P^{-1/2}\sum_{i=1}^{P-1} a_{R,i}h_{R+i}$$

It follows that

$$\Omega_{12} \equiv \lim_{R,P\to\infty} Cov(S_{1P}, S_{2P}) = \lim_{R,P\to\infty} Cov(S_{1P}, S_{2P.1}) + \lim_{R,P\to\infty} Cov(S_{1P}, S_{2P.2}).$$

We can show that $\lim_{R,P\to\infty} Cov(S_{1P}, S_{2P.1}) = 0$. To see this, let $T - R + 1 = P = N\lambda$. Then we can write

$$Cov(P^{-1/2}\sum_{t=R}^{T} f_t, P^{-1/2}\sum_{s=1+\tau+\ddot{r}}^{R} a_{R,0}h_s) = \frac{a_{R,0}}{P}\sum_{i=0}^{\lambda-1}\sum_{n=0}^{N-1} Cov(f_{R+n\lambda+i}^{(i+1)}, \sum_{s=1+\tau+\ddot{r}}^{R} h_s).$$

For $i = 0, \ldots, \lambda-1$, let $\gamma_j^{(i+1)} \equiv Cov(f_t^{(i+1)}, h_{t+j}) = Cov(f_{t+j}^{(i+1)},\ h_t) \equiv \gamma_{-j}^{(i+1)}$. Then we can bound $\sum_{n=0}^{N-1} Cov(f_{R+n\lambda+i}^{(i+1)}, \sum_{s=1+\tau+\ddot{r}}^{R} h_s)$ by $\sum_{i=0}^{\lambda-1}(|\gamma_0^{(i+1)}| + \sum_{j=-\infty}^{\infty}|j||\gamma_j^{(i+1)}|) < \infty$ for each i and the result follows since $\frac{a_{R,0}}{P} \leq \frac{1}{R} \to 0$.

Next, we derive $\lim_{R,P\to\infty} Cov(S_{1P}, S_{2P.2})$. For simplicity, write $c_R \equiv a_{R,0}$ and for $t = R+1, \ldots, T$, let $c_t = (1/t + \ldots + 1/T) \equiv a_{R,i}$ for $i = 1, \ldots, P-1$. With this notation, we have that

$$
\begin{aligned}
Cov(S_{1P}, S_{2P.2}) &= Cov(P^{-1/2}\sum_{t=R}^{T} f_{t+\tau}, P^{-1/2}\sum_{s=R+1}^{T} c_s h_s) \\
&= P^{-1}\sum_{t=R}^{T}\sum_{s=R+1}^{T} E(z_{t+\tau} c_s h_s), \quad z_{t+\tau} \equiv f_{t+\tau} - Ef_{t+\tau} \\
&= P^{-1}\sum_{t=R}^{T}\sum_{s=R}^{T} E(z_t c_s h_s) + \mathcal{R}_1 + \mathcal{R}_2,
\end{aligned}
$$

where $\mathcal{R}_1 \equiv -c_R/P\sum_{t=R}^{T} E(z_t h_R)$ and

$$
\mathcal{R}_2 \equiv P^{-1}\sum_{s=R}^{T} c_s\{-E[h_s(f_R + \ldots + f_{R+\tau-1}) + E[h_s(f_{T+1} + \ldots + f_{T+\tau})]\}.
$$

Under Assumption 6, only a finite number M of the expectations $E(z_t h_R)$ is non zero. Since $c_R = (1/R + \ldots + 1/T) = O(1)$, we obtain that $\mathcal{R}_1 = O(P^{-1}) = o(1)$. A similar argument implies that $\mathcal{R}_2 = O(P^{-1}) = o(1)$ since τ is finite and $c_s \leq (T - s + 1)/R \leq P/R = O(1)$ for all $s = R, \ldots, T$. Hence,

$$
\Omega_{12} = \lim_{R,P\to\infty} \bar{\Omega}_{12}, \quad \bar{\Omega}_{12} \equiv P^{-1}\sum_{t=R}^{T}\sum_{s=R}^{T} c_s E(z_t h_s).
$$

We can write

$$
\begin{aligned}
\bar{\Omega}_{12} &= \frac{1}{P}\sum_{t=R}^{T} c_t E(z_t h_t) + \frac{1}{P}\sum_{j=1}^{P-1}\sum_{t=R}^{T-j} c_t E(z_{t+j} h_t) + \frac{1}{P}\sum_{j=1}^{P-1}\sum_{t=R+j}^{T} c_t E(z_{t-j} h_t) \\
&= \frac{1}{P}\sum_{t=R}^{T} c_t E(z_t h_t) + \frac{1}{P}\sum_{j=1}^{M}\sum_{t=R}^{T-j} c_t E(z_{t+j} h_t) + \frac{1}{P}\sum_{j=1}^{M}\sum_{t=R+j}^{T} c_t E(z_{t-j} h_t),
\end{aligned}
$$

where we have used Assumption 6 to zero out the covariances with $j > M$ in obtaining the second equality. Adding and subtracting appropriately, it follows that

$$
\bar{\Omega}_{12} = \frac{1}{P}\sum_{t=R}^{T} c_t E(z_t h_t) + \sum_{j=1}^{M}\frac{1}{P}\sum_{t=R}^{T} c_t E(z_{t+j} h_t) + \sum_{j=1}^{M}\frac{1}{P}\sum_{t=R}^{T} c_t E(z_{t-j} h_t) + \mathcal{D}_1 + \mathcal{D}_2
$$

where we can show that the remainder terms $\mathcal{D}_1$ and $\mathcal{D}_2$ are of order $O(M^2/P) = o(1)$. In particular,

$$
\mathcal{D}_1 = -P^{-1}\sum_{j=1}^{M}\sum_{t=T-j+1}^{T} c_t E(z_{t+j} h_t) = O(P^{-1}) \quad \text{and} \quad \mathcal{D}_2 = P^{-1}\sum_{j=1}^{M}\sum_{t=R}^{R+j-1} c_t E(z_{t-j} h_t) = O(P^{-1}),
$$

under the assumption that M is finite. Using this result and the fact that $P = N\lambda$, we can write the first three terms in $\bar{\Omega}_{12}$ as

$$\begin{aligned}
\sum_{j=-M}^{M} \frac{1}{P} \sum_{t=R}^{T} c_t E(z_{t+j} h_t) &= \sum_{j=-M}^{M} \frac{1}{P} \sum_{n=0}^{N-1} \sum_{i=0}^{\lambda-1} c_{R+n\lambda+i} E(z_{R+j+n\lambda+i} h_{R+n\lambda+i}) \\
&= \sum_{j=-M}^{M} \frac{1}{N\lambda} \sum_{n=0}^{N-1} \sum_{i=0}^{\lambda-1} c_{R+n\lambda+i} E(z_{R+j+i} h_{R+i}) \\
&= \sum_{j=-M}^{M} \frac{1}{\lambda} \sum_{i=0}^{\lambda-1} E(z_{R+j+i} h_{R+i}) \Big(\frac{1}{N} \sum_{n=0}^{N-1} c_{R+n\lambda+i} \Big),
\end{aligned}$$

where the second equality follows because $E(z_{R+j+n\lambda+i} h_{R+n\lambda+i}) = E(z_{R+j+i} h_{R+i})$ since $z_t \equiv f_t - Ef_t$ is periodically stationary with periodicity λ. For this reason, and because M and λ are finite, we have that

$$\Omega_{12} = \lim_{R,P\to\infty} \bar{\Omega}_{12} = \sum_{j=-M}^{M} \frac{1}{\lambda} \sum_{i=0}^{\lambda-1} E(z_{R+j+i} h_{R+i}) \Big(\lim_{R,P\to\infty} \frac{1}{N} \sum_{n=0}^{N-1} c_{R+n\lambda+i} \Big).$$

For $i = 0, \ldots, \lambda - 1$, let $\gamma_j^{(i+1)} = E(z_{t+j}^{(i+1)} h_t)$ and $\bar{\gamma}_j = \frac{1}{\lambda} \sum_{i=0}^{\lambda-1} \gamma_j^{(i+1)}$. It follows that

$$\Omega_{12} = \Big(\lim_{R,P\to\infty} \frac{1}{N} \sum_{n=0}^{N-1} c_{R+n\lambda+\lambda-1} \Big) \sum_{j=-M}^{M} \bar{\gamma}_j.$$

Note that under Assumption 6,

$$\sum_{j=-M}^{M} \bar{\gamma}_j = \lim_{R,P\to\infty} Cov(P^{-1/2} \sum_{t=R}^{T} f_t, P^{-1/2} \sum_{s=R}^{T} h_s),$$

so the result follows by showing that we can approximate $\frac{1}{N} \sum_{n=0}^{N-1} c_{R+\lambda-1+n\lambda}$ by $\Pi \equiv 1 - \pi^{-1} \ln(1+\pi)$. This can be proven using logic such as that in the proof of (A-1a) of West (1996). We first show for any $n \in \{0, \ldots, N-1\}$, $c_{R+\lambda-1+n\lambda}$ can be approximated by $\ln\big(\frac{T}{R+\lambda-1+n\lambda}\big)$ and then use this result to approximate $\frac{1}{N} \sum_{n=0}^{N-1} c_{R+\lambda-1+n\lambda}$. More specifically, given the definition of $c_{R+\lambda-1+n\lambda}$, we can write $c_{R+\lambda-1+n\lambda} = \sum_{k=R+\lambda-1+n\lambda}^{T} \frac{1}{k}$, where $\frac{1}{k}$ is a positive decreasing function in k. Then for any given $n \in \{0, \ldots, N-1\}$, we have

$$\int_{R+\lambda-1+n\lambda}^{T} x^{-1} dx \leq \sum_{k=R+\lambda-1+n\lambda}^{T} \frac{1}{k} \leq \int_{R+\lambda-2+n\lambda}^{T} x^{-1} dx,$$

where the integral on the left hand side equals $\ln\big(\frac{T}{R+\lambda-1+n\lambda}\big)$ and the integral on the right hand side equals $\ln\big(\frac{T}{R+\lambda-1+n\lambda+1}\big)$. Note that these two values are equal in the limit. This allows us to substitute $c_{R+\lambda-1+n\lambda}$ by $\ln\big(\frac{T}{R+\lambda-1+n\lambda}\big)$ in the limit. We can then approximate

$\frac{1}{N}\sum_{n=0}^{N-1} c_{R+\lambda-1+n\lambda}$ by $N^{-1}\int_{-1}^{N-1}\ln\left(\frac{T}{R+\lambda-1+n\lambda}\right)dn$, whose limit as $R,P\to\infty$ can be shown to be $1-\pi^{-1}\ln(1+\pi)$.

■

Proof of Theorem 5.2. It suffices to show that $\hat{\Omega}_1 \xrightarrow{p} \Omega_1$, $\hat{\Omega}_2 \xrightarrow{p} \Omega_2$, $\hat{\Omega}_{12} \xrightarrow{p} \Omega_{12}$, $\hat{B} \xrightarrow{p} B$, and $\hat{\bar{F}} \xrightarrow{p} \bar{F}$. By Assumption 2, $\hat{B} \xrightarrow{p} B$. For $\hat{\bar{F}}$, we use the periodic structure of f_t and write

$$\hat{\bar{F}} = \frac{1}{P}\sum_{t=R}^{T}\partial f_{t+\tau}(\hat{\beta}_T)/\partial\beta' = \frac{1}{\lambda}\sum_{j=0}^{\lambda-1}\frac{1}{N}\sum_{n=0}^{N-1}\partial f^{(j+1)}_{R+\tau+n\lambda+j}(\hat{\beta}_T)/\partial\beta',$$

where for $j=0,\ldots,\lambda-1<\infty$,

$$\frac{1}{N}\sum_{n=0}^{N-1}\partial f^{(j+1)}_{R+\tau+n\lambda+j}(\hat{\beta}_T)/\partial\beta' \xrightarrow{p} E\partial f^{(j+1)}_{R+\tau+n\lambda+j}/\partial\beta' \equiv F^{(j+1)}$$

by Theorem 3 in Clark and McCracken (2009). Hence, $\hat{\bar{F}} \xrightarrow{p} \bar{F}$. The consistency of $\hat{\Omega}_1$, $\hat{\Omega}_2$, and $\hat{\Omega}_{12}$ follows from Lemma A.3 given Assumptions 1-7 and the fact that H_0 implies that $Ef_t = 0$ for all t. Note in particular the use of Lemma 5.3 in justifying the form Ω_{12} under Assumption 6. ■

B Appendix: Bootstrap Theory

As usual in the bootstrap literature, we use P^* to denote the bootstrap probability measure, conditional on the original sample (defined on a given probability space $(\Omega,\mathcal{F},P)$). For any bootstrap statistic t_T^*, we write $t_T^* = o_p^*(1)$, or $t_T^* \to^{P^*} 0$, when for any $\delta>0$, $P^*(|t_T^*|>\delta) = o_p(1)$. We write $t_T^* = O_p^*(1)$, when for all $\delta>0$ there exists $M_\delta<\infty$ such that $\lim_{T\to\infty} P[P^*(|t_T^*|>M_\delta)>\delta] = 0$. By Markov's inequality, this follows if $E^*|t_T^*|^q = O_p(1)$ for some $q>0$. Finally, we write $t_T^* \to^{d^*} D$, in probability, if conditional on a sample with probability that converges to one, t_T^* weakly converges to the distribution D under P^*, i.e. $E^*(f(t_T^*)) \to_p E(f(D))$ for all bounded and uniformly continuous functions f.

B.1 Auxiliary lemmas

In the following, recall that $f_{t+\tau}^* \equiv f(y^*_{t+\tau|r'}, x_t^*(t), \beta_0)$ and define $f^*_{t+\tau,\beta} \equiv \frac{\partial}{\partial\beta} f^*_{t+\tau}(\beta_0)$. Similarly, let $H^*(t) \equiv t^{-1}\sum_{s=1+\tau+\ddot{r}}^{t} h_s^*$, where $h_s^* \equiv x^*_{s-\tau}(y_s^* - x^{*\prime}_{s-\tau}\beta_0)$, and $B^*(t) \equiv \left(t^{-1}\sum_{s=1+\tau+\ddot{r}}^{t} x^*_{s-\tau}x^{*\prime}_{s-\tau}\right)^{-1}$.

Lemma B.1 *Under Assumptions 1-5 and assuming that $\ell\to\infty$ such that $\ell/\min\{\sqrt{R},\sqrt{P}\}\to$ 0,*

(a) *For any $0\le a<1/2$, $\sup_t |P^a(\hat{\beta}_t^* - \beta_0)| = o_p^*(1)$.*

(b) $P^{-1/2}\sum_{t=R}^{T}\bar{F}(B^*(t)-B)H^*(t) = o_p^*(1)$.

(c) $P^{-1/2}\sum_{t=R}^{T}(f^*_{t+\tau,\beta}-\bar{F})(B^*(t)-B)H^*(t)=o^*_p(1)$.

(d) $P^{-1/2}\sum_{t=R}^{T}(f^*_{t+\tau,\beta}-\bar{F})BH^*(t)=o^*_p(1)$.

Proof of Lemma B.1. The proofs of parts (a) and (b) are identical to the proofs of Lemma A.2 (c) and (e) in GMY(2025), respectively, with the only difference that F there is replaced by $\bar{F}$ here. These results depend only on the final data on $x_{s-\tau}$ and h_s which are assumed to be stationary and strong mixing as in GMY(2025). Parts (c) and (d) depend on $f^*_{t+\tau,\beta}$, the bootstrap version of $f_{t+\tau,\beta}$, which we now assume to be periodically stationary rather than covariance stationary. Nevertheless, we can adapt the proof of Lemma A.2(e) and (f) in GMY(2025) to our context. For instance, the proof of part (c) follows exactly as in the proof of GMY(2025)'s Lemma A.2 (e) by replacing F with $\bar{F}$ and applying their Lemma A.3 (which does not depend on assuming covariance stationarity). Similarly, part (f) follows by applying the same arguments as used in GMY(2025) to prove their Lemma A.2 (d), replacing in particular F with $\bar{F}$ to account for the differences implied by the periodic heterogeneity of $f_{t+\tau}$. Since the arguments are very similar, we omit the details here. ■

The following result is the analogue of Lemma A.4 in GMY(2025), adapted to our annual revisions context. Recall that $S^*_{1P}=P^{-1/2}\sum_{t=R}^{T}(f^*_{t+\tau}-f_{t+\tau})$. Similarly, $S^*_{2P}=a_{R,0}P^{-1/2}\sum_{s=1+\tau+\ddot{r}}^{R}(h^*_s-\bar{h}_R)+P^{-1/2}\sum_{i=1}^{P-1}a_{R,i}(h^*_{R+i}-\bar{h}_P)\equiv S^*_{2P.1}+S^*_{2P.2}$.

Lemma B.2 *Let Assumptions 1-6 hold. Then, if* $\ell\to\infty$ *such that* $\ell/\min\{\sqrt{R},\sqrt{P}\}\to 0$*, under* H_0*, (a)* $Var^*(S^*_{1P})\xrightarrow{p}\Omega_1$*; (b)* $Var^*(S^*_{2P})\xrightarrow{p}\Omega_2$*; and (c)* $Cov^*(S^*_{1P},S^*_{2P})\xrightarrow{p}\Omega_{12}$.

Proof of Lemma B.2. Part (a) follows from Theorem 3.1 in Fitzenberger (1998) or Theorem 2.2 in Gonçalves and White (2002). Both of these results show the consistency of the MBB variance of a sample mean allowing for heterogeneity but imposing a restriction on the mean heterogeneity. In our application, this restriction is satisfied under the null hypothesis $H_0: E(f_{t+\tau})=0$. Part (b) follows from Lemma A.4 (b) in GMY(2025) without modification since we maintain the same assumptions on h_s as theirs. Similarly, the proof of part (c) follows closely the proof of Lemma A.4 (c) in GMY(2025) with some minor modifications due to the heterogeneity of $f_{t+\tau}$. More specifically, by the independence between S^*_{1P} and $S^*_{2P.1}$, $Cov^*(S^*_{1P},S^*_{2P})=Cov^*(S^*_{1P},S^*_{2P.2})$. Let $k_2=P/\ell$ and define $c_{R+i}\equiv a_{R,i}$, where $a_{R,i}\equiv 1/(R+i)+\ldots+1/T$ for $i=1,\ldots,P-1$ as in Appendix A, setting $c_{R+i}=0$ when

$i > P-1$. Similarly, let $b_i = R+1+i\ell$. Then, we can show that

$$\begin{aligned} Cov^*(S^*_{1P}, S^*_{2P.2}) =& P^{-1}\sum_{m=1}^{\ell-1}(\sum_{i=0}^{k_2-1}\sum_{j=0}^{\ell-1-m} c_{b_i+j+m})P^{-1}\sum_{s=R+1}^{T+1} Cov(f_s, h_{s+m}) \\ &+ P^{-1}(\sum_{i=0}^{k_2-1}\sum_{j=0}^{\ell-1} c_{b_i+j})P^{-1}\sum_{s=R+1}^{T+1} Cov(f_s, h_s) \\ &+ P^{-1}\sum_{m=1}^{\ell-1}(\sum_{i=0}^{k_2-1}\sum_{j=0}^{\ell-1-m} c_{b_i+j})P^{-1}\sum_{s=R+1}^{T+1} Cov(f_{s+m}, h_s) + o_p(1). \end{aligned}$$

The representation above corresponds to $\mathcal{W}_{1.1}+\mathcal{W}_{2.1}+\mathcal{W}_{3.1}$ in GMY(2025) with the difference that here $\Gamma_{fh}(m) = E(f_t h_{t+m})$ is not time invariant due to the periodicitiy of f_t. We next use Assumption 6 (i.e., $Cov(f_s, h_t) = 0$ for $|s-t| > M$ where M is a positive integer) to write

$$\begin{aligned} Cov^*(S^*_{1P}, S^*_{2P}) = & P^{-1}\sum_{m=1}^{M-1}(\sum_{i=0}^{k_2-1}\sum_{j=0}^{\ell-1-m} c_{b_i+j+m})P^{-1}\sum_{s=R+1}^{T+1} Cov(f_s, h_{s+m}) \\ &+ P^{-1}(\sum_{i=0}^{k_2-1}\sum_{j=0}^{\ell-1} c_{b_i+j})P^{-1}\sum_{s=R+1}^{T+1} Cov(f_s, h_s) \\ &+ P^{-1}\sum_{m=1}^{M-1}(\sum_{i=0}^{k_2-1}\sum_{j=0}^{\ell-1-m} c_{b_i+j})P^{-1}\sum_{s=R+1}^{T+1} Cov(f_{s+m}, h_s) + o_p(1). \end{aligned}$$

This allows us to show that $Cov^*(S^*_{1P}, S^*_{2P}) \xrightarrow{p} \Omega_{12}$ using arguments similar to those used in the proof of Lemma 5.3. For instance,

$$P^{-1}(\sum_{i=0}^{k_2-1}\sum_{j=0}^{\ell-1} c_{b_i+j})P^{-1}\sum_{s=R+1}^{T+1} Cov(f_s, h_s) = (P^{-1}\sum_{t=R+1}^{T+1} c_t)P^{-1}\sum_{s=R+1}^{T+1} Cov(f_s, h_s),$$

where $P^{-1}\sum_{t=R+1}^{T+1} c_t \to \Pi \equiv 1-\pi^{-1}\ln(1+\pi)$, and using the fact that f_s is periodically stationary with periodicity λ,

$$\begin{aligned} P^{-1}\sum_{s=R+1}^{T+1} Cov(f_s, h_s) &= N^{-1}\sum_{n=0}^{N-1}\lambda^{-1}\sum_{i=0}^{\lambda-1} Cov(f^{(i+1)}_{R+1+n\lambda+i}, h_{R+1+n\lambda+i}) \\ &= \lambda^{-1}\sum_{i=0}^{\lambda-1} Cov(f^{(i+1)}_{R+1+i}, h_{R+1+i}) = \lambda^{-1}\sum_{i=0}^{\lambda-1}\gamma_0^{(i+1)} = \bar{\gamma}_0. \end{aligned}$$

Similarly, consider

$$\mathcal{A} \equiv P^{-1}\sum_{m=1}^{M-1}(\sum_{i=0}^{k_2-1}\sum_{j=0}^{\ell-1-m} c_{b_i+j})P^{-1}\sum_{s=R+1}^{T+1} Cov(f_{s+m}, h_s).$$

Adding and subtracting appropriately, $\mathcal{A} = \mathcal{A}_1 + \Delta_1$, where

$$\mathcal{A}_1 \equiv \sum_{m=1}^{M-1} (P^{-1} \sum_{t=R+1}^{T+1} c_t) P^{-1} \sum_{s=R+1}^{T+1} Cov(f_{s+m}, h_s),$$

and

$$\Delta_1 \equiv P^{-1} \sum_{m=1}^{M-1} \underbrace{(\sum_{i=0}^{k_2-1} \sum_{j=0}^{\ell-1-m} c_{b_i+j} - \sum_{t=R+1}^{T+1} c_t)}_{\mathcal{D}_m} P^{-1} \sum_{s=R+1}^{T+1} Cov(f_{s+m}, h_s).$$

Note that for finite M,

$$\lim_{R,P\to\infty} \mathcal{A}_1 = \sum_{m=1}^{M-1} \Big(\lim_{R,P\to\infty} P^{-1} \sum_{t=R+1}^{T+1} c_t \Big) \Big(\lim_{R,P\to\infty} P^{-1} \sum_{s=R+1}^{T+1} Cov(f_{s+m}, h_s) \Big),$$

where $P^{-1} \sum_{t=R+1}^{T+1} c_t \to \Pi$, and $\lim_{R,P\to\infty} P^{-1} \sum_{s=R+1}^{T+1} Cov(f_{s+m}, h_s) = \bar{\gamma}_m \equiv \frac{1}{\lambda} \sum_{i=0}^{\lambda-1} \gamma_m^{(i+1)}$. To show $\Delta_1 \to 0$, we show

$$\mathcal{D}_m = P^{-1} \sum_{i=0}^{k_2-1} (\sum_{j=0}^{\ell-1-m} c_{b_i+j} - \sum_{s=0}^{\ell-1} c_{b_i+s}) \to 0$$

for any $m = 1, \ldots, M$. Recall that for any i and j, $c_{b_i+j} \leq c_R = \frac{1}{R} + \ldots + \frac{1}{T}$. We can write

$$\mathcal{D}_m \leq P^{-1} k_2 M c_R \leq \ell^{-1} M P R^{-1} \to 0,$$

since $P/R \to \pi$, and $M < \infty$. ■

B.2 Proofs of Lemma 6.1 and Theorem 6.1 in the main text

Proof of Lemma 6.1. We follow closely the proof of Lemma 5.1 in GMY(2025). As in that paper, we consider two second-order mean value expansions, one of $f^*_{t+\tau}(\hat{\beta}^*_t)$ around β_0, and another of $f_{t+\tau}(\bar{\beta}_t)$ around β_0. The first of these expansions yields

$$P^{-1/2} \sum_{t=R}^{T} f^*_{t+\tau}(\hat{\beta}^*_t) = P^{-1/2} \sum_{t=R}^{T} f^*_{t+\tau} + \xi^*_1 + \xi^*_2$$

where $\xi^*_1 \equiv P^{-1/2} \sum_{t=R}^{T} f^*_{t+\tau,\beta}(\hat{\beta}^*_t - \beta_0)$, and $\xi^*_2 \equiv 0.5 P^{-1/2} \sum_{t=R}^{T} \frac{\partial^2}{\partial \beta^2} f^*_{t+\tau}(\tilde{\beta}^*_t)(\hat{\beta}^*_t - \beta_0)^2$, with $\tilde{\beta}^*_t$ lying between $\hat{\beta}^*_t$ and β_0. We can show that $\xi^*_2 = o^*_p(1)$ using the same arguments as in GMY(2025). In particular, we use Lemma B.1(a) to obtain $\sup_t |P^{1/4}(\hat{\beta}^*_t - \beta_0)| = o^*_p(1)$. To show that $P^{-1} \sum_{t=R}^{T} |\frac{\partial^2}{\partial \beta^2} f^*_{t+\tau}(\tilde{\beta}^*_t)| = O^*_p(1)$, we can use Markov's inequality to obtain

$$P^*\Big(P^{-1} \sum_{t=R}^{T} |\frac{\partial^2}{\partial \beta^2} f^*_{t+\tau}(\tilde{\beta}^*_t)| > \Delta\Big) \leq P^*\Big(\sum_{j=0}^{\lambda-1} P^{-1} \sum_{t=R}^{T} m^{(j+1)}_{\eta_{t+\tau}} > \Delta \Big) \leq \Delta^{-1} \sum_{j=0}^{\lambda-1} P^{-1} \sum_{t=R}^{T} E^*(m^{(j+1)}_{\eta_{t+\tau}}) \to_p 0,$$

since for $j=0,\ldots,\lambda-1$, $P^{-1}\sum_{t=R}^{T}E^*(m_{\eta_{t+\tau}}^{(j+1)})=O_p(1)$ under Assumption 1. The proof that $\xi_1^*=\bar{F}BP^{1-}\sum_{t=R}^{T}H^*(t)+o_p^*(1)$ follows the same arguments as in GMY(2025) by replacing F with $\bar{F}$ and using Lemma B.1 (d), (b) and (c) instead of their Lemma A.2 (d), (e), and (f), respectively. Hence,

$$P^{-1/2}\sum_{t=R}^{T}f_{t+\tau}^*(\hat{\beta}_t^*)=P^{-1/2}\sum_{t=R}^{T}f_{t+\tau}^*+\underbrace{\bar{F}BP^{-1/2}\sum_{t=R}^{T}H^*(t)}_{\equiv\xi_{1.1}^*}+o_p^*(1).$$

Similarly, an expansion of $f_{t+\tau|r'}(\bar{\beta}_t)$ around β_0 yields

$$P^{-1/2}\sum_{t=R}^{T}f_{t+\tau}(\bar{\beta}_t)=P^{-1/2}\sum_{t=R}^{T}f_{t+\tau}+\bar{\xi}_1+\bar{\xi}_2,$$

where

$$\bar{\xi}_1=P^{-1/2}\sum_{t=R}^{T}f_{t+\tau,\beta}(\bar{\beta}_t-\beta_0)\text{ and }\bar{\xi}_2=0.5P^{-1/2}\sum_{t=R}^{T}\frac{\partial^2}{\partial\beta^2}f_{t+\tau}(\ddot{\beta}_t)(\bar{\beta}_t-\beta_0)^2.$$

where $\ddot{\beta}_t$ lies between $\bar{\beta}_t$ and β_0, and $f_{t+\tau,\beta}\equiv f_{t+\tau,\beta}(\beta_0)$. We can show that $\bar{\xi}_2=o_p(1)$ using a similar argument to that used to show that $\xi_2^*=o_p^*(1)$. In particular, it suffices to show that $\sup_t|P^{1/4}(\bar{\beta}_t-\beta_0)|=o_p(1)$ (which follows from Lemma A.2) and $P^{-1}\sum_{t=R}^{T}|\frac{\partial^2}{\partial\beta^2}f_{t+\tau}(\ddot{\beta}_t)|=O_p(1)$, which follows under Assumption 1. For $\bar{\xi}_1$, we can follow the same arguments as GMY(2025) to show that this term can be represented as $\bar{\xi}_1=\bar{\xi}_{1.1}+\bar{\xi}_{1.2}+o_p(1)$, where $\bar{\xi}_{1,1}=P^{-1/2}\sum_{t=R}^{T}F_{t+\tau}Rt^{-1}BH(R)$ and $\bar{\xi}_{1.2}=P^{-1/2}\sum_{t=R}^{T}F_{t+\tau}(t-R)t^{-1}BH(P)$. These two terms correspond to $\bar{\xi}_{1.1}$ and $\bar{\xi}_{1.2}$ defined in the proof of Lemma 5.1 of GMY(2025) when $F=F_{t+\tau}$ but not otherwise. The proof that the remainder term is $o_p(1)$ relies on an application of Lemma A.2, which is the analogue of Lemma A.4 of West (1996) under annual revisions. Under our new set of assumptions, we can further decompose $\bar{\xi}_{1.1}=\bar{\xi}_{1.1.1}+o_p(1)$, and $\bar{\xi}_{1.2}=\bar{\xi}_{1.2.1}+o_p(1)$ where

$$\bar{\xi}_{1.1.1}=\bar{F}P^{-1/2}\sum_{t=R}^{T}Rt^{-1}BH(R),\quad\text{and}\quad\bar{\xi}_{1.2.1}=\bar{F}P^{-1/2}\sum_{t=R}^{T}(t-R)t^{-1}BH(P),$$

which correspond to $\bar{\xi}_{1,1}$ and $\bar{\xi}_{1,2}$ in GMY(2025) with F replaced by $\bar{F}$. Combining these results yields

$$\tilde{S}_P^*\equiv P^{-1/2}\sum_{t=R}^{T}(f_{t+\tau}^*(\hat{\beta}_t^*)-f_{t+\tau}(\bar{\beta}_t))=P^{-1/2}\sum_{t=R}^{T}(f_{t+\tau}^*-f_{t+\tau})+(\xi_{1.1}^*-\bar{\xi}_{1.1.1}-\bar{\xi}_{1.2.1})+o_p(1),$$

where $S_{1P}^*=P^{-1/2}\sum_{t=R}^{T}(f_{t+\tau}^*-f_{t+\tau})$, and we can show that $\xi_{1.1}^*-\bar{\xi}_{1.1.1}-\bar{\xi}_{1.2.1}=\bar{F}BS_{2P}^*$ with

$$S_{2P}^*=a_{R,0}P^{-1/2}\sum_{s=1+\tau+\ddot{r}}^{R}(h_s^*-\bar{h}_R)+P^{-1/2}\sum_{i=1}^{P-1}a_{R,i}(h_{R+i}^*-\bar{h}_P),$$

a result that follows from GMY(2025)'s proof when we replace F by $\bar{F}$. ■

Proof of Theorem 6.1. The proof follows the same steps as the proof of Theorem 5.1 of GMY(2025), so we omit the details. The main differences are that (i) we rely on Theorem 5.1 to claim that $\Omega^{-1/2}\tilde{S}_P^{\mu} \to_d N(0,1)$, and (ii) we use Lemma 6.1 and Theorem 5.2 to show that $\tilde{S}_P^{*} \to^{d^*} N(0,\Omega)$ when the null is true. ■

C Additional simulation results

The following table contains results for the equal MSE experiment underlying Table 4 in Section 7, but with $P = 320$.

Table C.1: Size and power results of equal MSE experiment with $P = 320$

Tests	$\lambda = 1$	4	12	$\lambda = 1$	4	12
	size: DL(1)			power: DL(1)		
t_{DM}	0.1276	0.1169	0.1118	0.9548	0.9456	0.9219
t_{CM}	0.0454	0.0392	0.0296	0.9284	0.9124	0.8643
$Bootstrap_R$	0.0463	0.0386	0.0265	0.9218	0.9034	0.8320
Bootstrap	0.0463	0.0386	0.0265	0.9218	0.9034	0.8320
	size: DL(2)			power: DL(2)		
t_{DM}	0.0509	0.0866	0.0996	1.0000	0.5679	0.8240
t_{CM}	0.0505	0.0455	0.0323	1.0000	0.5064	0.7494
$Bootstrap_R$	0.0480	0.3279	0.3555	1.0000	0.8482	0.9642
Bootstrap	0.0480	0.0425	0.0257	1.0000	0.4884	0.7030

References

[1] Barnichon, R., and Nekarda, C. J. (2012), "The Ins and Outs of Forecasting Unemployment: Using Labor Force Flows to Forecast the Labor Market," *Brookings Papers on Economic Activity*, Fall, 83-131.

[2] Birinci, S., Karahan, F., Mercan, Y., and K. See (2024), "Labor Market Shocks and Monetary Policy," *Federal Reserve Bank of St. Louis WP*.

[3] Borup, D., and E.C.M. Schütte (2022), "In Search of a Job: Forecasting Employment Growth Using Google Trends," *Journal of Business & Economic Statistics* 40, 186-200.

[4] Clark, T.E., and M.W. McCracken (2001), "Tests of Equal Forecast Accuracy and Encompassing for Nested Models," *Journal of Econometrics*, 105, 85-110.

[5] Clark, T.E., and M.W. McCracken (2009), "Tests of Equal Predictive Ability with Real-Time Data," *Journal of Business and Economic Statistics* 27, 441-54.

[6] Clements, M.P., and A.B. Galvão (2013), "Real-time Forecasting of Inflation and Output Growth with Autoregressive Models in the Presence of Data Revisions," *Journal of Applied Econometrics*, 28, 458-477.

[7] Clements, M.P., and A.B. Galvão (2019), "Data Revisions and Real-Time Forecasting," in *Oxford Research Encyclopedia of Economics and Finance*, Oxford University Press.

[8] Corradi, V., and N. Swanson (2007), "Nonparametric Bootstrap Procedures for Predictive Inference Based on Recursive Estimation Schemes", *International Economic Review*, 48:67–109, 2007.

[9] De Jong, R. M. (1997), "Central limit theorems for dependent heterogenous random variables " *Econometric Theory*, 13, 353-367.

[10] De Jong, R. M., and Davidson, J. (2000), "Consistency of kernel estimators of heteroscedastic and autocorrelated covariance matrices " *Econometrica*, 68(2), 407-423.

[11] Diebold, F. X., and Mariano, R. S. (1995), "Comparing Predictive Accuracy," *Journal of Business and Economic Statistics*, 20(1), 134-144.

[12] Fitzenberger, B. (1998), "The Moving Blocks Bootstrap and Robust Inference for Linear Least Squares and Quantile Regressions, " *Journal of Econometrics*, 82, 235-287.

[13] Gallant, A.R., and H. White (1988), A Unified Theory of Estimation and Inference for Nonlinear Dynamic Models, (B. Blackwell).

[14] Giacomini, R., and B. Rossi (2010), "Forecast Comparisons in Unstable Environments," *Journal of Applied Econometrics* 25, 595-620.

[15] Gonçalves, S. and H. White (2002), "The Bootstrap of the Mean for Dependent Heterogeneous Arrays", *Econometric Theory*, 18, 1367-1384.

[16] Gonçalves, S., M.W. McCracken, and Y. Yao (2025), "Bootstrapping out-of-sample predictability tests with real-time data," *Journal of Econometrics*, 247, TBD.

[17] Hurd, H.L. and Miamee, A.G. (2007), "Periodically Correlated Random Sequences: Spectral. Theory and Practice, "John Wiley.

[18] Koenig, E.F., Dolmas, S., and J. Piger (2003), "The Use and Abuse of Real-Time Data in Economic Forecasting," *The Review of Economics and Statistics*, 85, 618-628.

[19] Rapach, D.E., and J.K. Strauss (2008), "Forecasting US employment growth using forecast combining methods," *Journal of Forecasting*, 27, 75-93.

[20] Rapach, D.E., and J.K. Strauss (2010), "Bagging or Combining (or Both)? An Analysis Based on Forecasting US Employment Growth," *Econometric Reviews*, 29, 511-533.

[21] West, K.D. (1996), "Asymptotic Inference about Predictive Ability," *Econometrica*, 64, 1067-1084.