ARTICLE IN PRESS
Journal of Econometrics xxx (xxxx) xxx

Contents lists available at ScienceDirect

Journal of Econometrics

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

Transformed regression-based long-horizon predictability tests[☆]

Matei Demetrescu[a], Paulo M.M. Rodrigues[b], A.M. Robert Taylor[c,*]

[a] Department of Statistics, TU Dortmund University, Germany
[b] Banco de Portugal and Nova School of Business and Economics, Portugal
[c] Essex Business School, University of Essex, United Kingdom

ARTICLE INFO

Article history:
Received 1 April 2021
Received in revised form 5 April 2022
Accepted 4 June 2022
Available online xxxx

JEL classification:
C12
C22
G17

Keywords:
Long-horizon predictive regression
IVX estimation
(Un)conditional heteroskedasticity
Unknown regressor persistence
Endogeneity
Residual augmentation

ABSTRACT

We propose new tests for long-horizon predictability based on IVX estimation of a transformed regression which explicitly accounts for the over-lapping nature of the dependent variable in the long-horizon regression arising from temporal aggregation. To improve efficiency, we moreover incorporate the residual augmentation approach recently used in the context of short-horizon predictability testing by Demetrescu and Rodrigues (2022). Our proposed tests improve on extant tests in the literature in a number of ways. First, they allow practitioners to remain ambivalent over the strength of the persistence of the predictors. Second, they are valid under much weaker conditions on the innovations than extant long-horizon predictability tests; in particular, we allow for general forms of conditional and unconditional heteroskedasticity in the innovations, neither of which are tied to a parametric model. Third, unlike the popular Bonferroni-based methods in the literature, our proposed tests can handle multiple predictors, and can be easily implemented as either one or two-sided hypotheses tests. Monte Carlo analysis suggests that our preferred tests offer improved finite sample properties compared to the leading tests in the literature. We report results from an empirical application investigating the use of real exchange rates for predicting nominal exchange rates and inflation.

1. Introduction

Since the seminal work of Fama and French (1988) and Campbell and Shiller (1988) there has been substantial interest in testing for long-horizon predictability, most notably in stock returns, exchange rates and the term structure of interest rates; see, *inter alia*, Campbell and Shiller (1987, 1988), Fama (1998); Campbell and Cochrane (1999); Campbell and Viceira (1999); Menzly et al. (2004); Mishkin (1990); Boudoukh and Matthew (1993) and Chang et al. (2018).

Empirical evidence on the short- or long-horizon predictability of returns largely derives from inference obtained from predictive regressions and, as such, the size and power properties of tests from these regressions are of fundamental

☆ The authors thank two anonymous referees, the Co-Editor (Torben Andersen), and Tassos Magdalinos for their helpful and constructive feedback on earlier versions of this paper. Rodrigues gratefully acknowledges financial support from the Portuguese Science Foundation (FCT) through project PTDC/EGE-ECO/28924/2017, and (UID/ECO/00124/2013 and Social Sciences DataLab, Project 22209), POR Lisboa (LISBOA-01-0145-FEDER-007722 and Social Sciences DataLab, Project 22209) and POR Norte (Social Sciences DataLab, Project 22209). Taylor gratefully acknowledges financial support provided by the Economic and Social Research Council of the United Kingdom under research grant ES/R00496X/1.

* Correspondence to: Essex Business School, University of Essex, Wivenhoe Park, Colchester, CO4 3SQ, United Kingdom.
E-mail address: robert.taylor@essex.ac.uk (A.M.R. Taylor).

https://doi.org/10.1016/j.jeconom.2022.06.006
0304-4076/

importance. Many early studies are based on the assumption that the predictor is weakly persistent and are therefore based on the use of standard OLS t and F-type regression statistics, constructed using either Newey–West or Hodrick type standard errors (see, for example, Weigand and Irons, 2007). However, data analysis presented in, among others, Campbell and Yogo (2006a) and Welch and Goyal (2008) suggests that many of the variables used in predictive regressions are strongly persistent with autoregressive roots close to unity, and that a large negative correlation often exists between the series we are attempting to forecast (e.g. returns) and the predictor's innovations, such that the predictive regressor is endogenous. In such cases these methods, developed for use with weakly persistent regressors, are theoretically invalid and this can lead to sizeable finite sample bias in the estimates of the coefficients from the predictive regression (Stambaugh, 1986 and Mankiw and Shapiro, 1986) and, correspondingly, to significant over-rejections of the null hypothesis of no predictability (both in short- and long-horizon contexts), thereby significantly increasing the likelihood that any finding of long-horizon predictability is spurious; see, *inter alia*, Valkanov (2003), Cochrane (2011), and Phillips (2015).[1]

As a result, more recently a number of procedures for testing for short- and long-horizon predictability have been developed in the literature which are designed to be robust as to whether the predictors are weakly or strongly persistent; see, in particular, Gonzalo and Pitarakis (2012), Phillips and Lee (2013), Phillips (2014), Elliott et al. (2015), Lee (2016), Kostakis et al. (2015), Breitung and Demetrescu (2015), Demetrescu et al. (2022b), Demetrescu and Hillmann (2022) and Demetrescu and Rodrigues (2022). Many of these procedures are based on the extended instrumental variable estimation [IVX] method of Phillips and Magdalinos (2009) which has gained widespread popularity in this literature and which will form the basis of the tests which we propose in this paper. The IVX approach consists of filtering putative predictors such that, where these are strongly (weakly) persistent, the filtered series are approximately mildly integrated (weakly dependent) variables. These filtered variables are then used to instrument the predictor in the predictive regression of interest. As a result of the reduced persistence of the instrument when compared to the original variable when the latter is strongly persistent, the resulting predictability test will follow a standard limit distribution (e.g. Gaussian or chi-squared) irrespective of whether the predictors are strongly, moderately, or weakly persistent.

An additional complication, relative to the case of short-horizon predictability testing, arises when looking to develop tests for long-horizon predictability. Specifically, serial correlation is induced into the error term in the long-horizon predictive regression, arising from the temporal aggregation of the dependent variable (which therefore contains overlapping observations). To address this issue, Valkanov (2003) and Hjalmarsson (2011) propose using the conventional OLS t-statistic but scaled by a constant to reflect the inflation of the standard errors as the prediction horizon increases. The methods in Valkanov (2003) and Hjalmarsson (2011) are, however, somewhat restrictive in practice as they are based on the assumption that the predictor is strongly persistent. Tests for multiple-horizon predictability designed to be asymptotically valid regardless of whether the predictors are strongly or weakly persistent and for handling the issues arising from temporal aggregation are also considered by Phillips and Lee (2013) who develop tests from a reversed predictive regression framework, estimated by IVX. Their approach consists of switching from a predictive regression from the h-period returns on a predetermined variable to a predictive regression of single period returns on the same predetermined variable aggregated over h-periods. Xu (2020) proposes an alternative approach, which allows the predictors to be either weakly or strongly persistent, and builds on an implied estimator obtained from the short-horizon predictive regression model. Implied estimation dates back to Campbell and Shiller (1988) and Hodrick (1992), and was used by Cochrane (2008) and Lettau and Van Nieuwerburgh (2008). Xu (2020) derives the asymptotic distribution of the implied test statistic and proposes the use of a Bonferroni-type approach along the lines of Phillips (2014) together with a wild bootstrap for computing critical values.

In this paper we add to the corpus of available tests for long-horizon predictability in the literature. The tests we will develop are designed to be valid under weaker conditions than the leading long horizon predictability tests in the literature, all of which either assume the strength of the persistence of the predictor is known (some assume it is weakly persistent, some that it is strongly persistent) and/or assume that the innovations are conditionally homoskedastic. In particular, our proposed tests can be validly implemented without knowledge of whether the predictors are weakly, moderately, or strongly persistent, and, unlike Bonferroni-based tests, our tests can be easily implemented to test either one or two-sided hypotheses, and can handle the case of multiple predictors. Our test statistics have pivotal limiting null distributions under quite general patterns of unconditional time heteroskedasticity in the innovations, allowing for time-varying innovation variances but also the possibility of time-varying correlations between the innovations, and very general forms of conditional heteroskedasticity. Moreover, the practitioner is not required to assume a parametric model for either the conditional or unconditional time-variation in the innovations. In a detailed Monte Carlo experiment we also compare the finite size and power properties of our proposed tests with the best-performing robust long-horizon predictability tests in the literature, namely the implied test of Xu (2020), the Bonferroni-based approach of Hjalmarsson (2011), and the reversed regression-based test of Phillips and Lee (2013). These results suggest that our proposed tests overall display superior finite sample properties to the extant tests.

The tests we propose are developed within a transformed regression framework which explicitly accounts for the serial correlation induced by temporal aggregation in the error in the original long-horizon regression. We estimate the

1 The standard errors proposed by Hodrick (1992), which exploit the moving-average structure of the temporally aggregated error term under the no predictability null hypothesis, perform slightly better than Newey–West standard errors in finite samples (see Ang and Bekaert, 2007) but are still invalid under endogeneity and strong persistence.

parameters of the transformed regression using the IVX approach of Kostakis et al. (2015). In this sense, our approach is related to the recent work of Kostakis et al. (2018) on IVX long-horizon predictive regression. The use of IVX estimation in our framework has the advantage that it also allows us to implement a feasible form of residual augmentation which cannot be employed where the predictive regression is estimated by OLS. This approach, discussed in Demetrescu and Rodrigues (2022) in the context of the IVX one-step ahead (short-horizon) predictive regression, consists of augmenting the transformed predictive regression with an additional regressor, constructed as the residuals obtained from fitting an autoregression to the predictor. Residual augmentation, at least for the case of a known degree of persistence, can be traced back to at least Phillips (1991), and augmenting regression models with residuals or nonlinear functions thereof is known to be an effective way of increasing efficiency; see, for example, Im and Schmidt (2008). In the context of the short-horizon predictive regression, Demetrescu and Rodrigues (2022) show that this approach is particularly effective for strongly persistent predictors. We will demonstrate that the estimation effect from fitting this autoregression to the predictor is asymptotically negligible in the set-up we consider and leads to more efficient estimation of the transformed predictive regression model on which our long-horizon tests are based, and therefore higher local power. In particular, akin to Amihud and Hurvich (2004), this form of residual augmentation eliminates endogeneity in the limit, such that the finite-sample bias of the IVX slope coefficient estimator is reduced compared to the corresponding IVX estimation from the transformed regression without this additional regressor.[2]

The remainder of the paper is organised as follows. Section 2 introduces the long-horizon predictive regression testing framework and outlines the assumptions on the model under which we work. In Section 3 we briefly review the leading tests in the literature: namely, Bonferroni-based approaches to testing for long-horizon predictability, focusing on the tests of Hjalmarsson (2011), the reversed regression based approach of Phillips and Lee (2013), and the implied testing approach of Xu (2020). In Section 4 we detail our proposed transformed regression based tests for long-horizon predictability testing, and here we also discuss their large sample properties. For expositional purposes, the material in Sections 2–4 assumes the case of a single predictor. The case of multiple predictors is discussed in Section 5. Section 6 analyses the finite sample properties of the procedures in an in-depth Monte Carlo study. In Section 7 we report an empirical application of the methods developed in the paper to exchange rate predictability. Section 8 concludes. An on-line Supplementary Appendix collects all technical proofs of the results stated in the paper together with some additional supporting Monte Carlo results and technical derivations.

2. The long-horizon predictive regression framework

2.1. The DGP and assumptions

We will base our analysis in what follows on the assumption that the data generating process [DGP] for (y_{t+1}, x_{t+1}) is given by the short-run (one period) predictive recursive system,

$$y_{t+1} = \alpha_1 + \beta_1 x_t + u_{t+1}, \qquad t = 1, \ldots, T-1, \tag{2.1}$$

$$x_{t+1} = \mu_x + \xi_{t+1}, \qquad \text{and} \qquad \xi_{t+1} = \rho\xi_t + v_{t+1}, \tag{2.2}$$

where y_{t+1} is, for example, a continuously compounded excess return of an asset or the variation of a nominal exchange rate from t to $t+1$ and x_{t+1} is some (putative) predictor variable. The errors u_t are assumed to form a martingale difference [MD] sequence; precise details will be given below. In our main exposition and technical analysis we will follow the bulk of this literature and focus attention on the case of a single predictor; that is, where x_t in (2.1) is a scalar. Extensions to the case where the predictive regression contains multiple predictors will be discussed in Section 5.

Remark 1. Our assumption that the data on (y_{t+1}, x_{t+1}) are generated by the one period ($h = 1$, where h is the horizon period) predictive system in (2.1)–(2.2) is in common with the extant methods in the long-horizon predictability testing literature discussed in Section 1. It is, however, important to stress that this is an assumption made for the purposes of providing a convenient unified benchmark to allow us to make rigorous statements about the properties of statistics obtained from the implied long-horizon predictive regression models with $h > 1$, defined in Section 2.2. One could alternatively make the assumption that the h-period aggregated model in (2.5) with the error term specified to be a MD sequence constitutes the true DGP. However, this approach seems problematic because the true value of h, such that a well-specified long-horizon model with uncorrelated errors obtains, is unknown; in practice researchers tend to report the outcomes of tests computed for a range of values of h, including $h = 1$. Under this alternative assumption, only at most one of the values of h considered could possibly correspond to the true DGP with the approach rendered invalid for the other values considered. In this regard, assuming $h = 1$ as the true DGP has the advantage that for any $h > 1$ the error term in the long horizon model in (2.5) will be serially correlated (see the discussion in Section 2.2), with the testing methods we develop explicitly designed to account for the maximum degree of serial correlation that could be induced by the data aggregation. ◊

[2] This bias reduction improves the MSE of the forecasts generated using the fitted residual augmented long-horizon regression; see the evidence provided by Demetrescu and Rodrigues (2022) for the one-step ahead case.

Our interest in this paper centres on testing the null hypothesis, H_0, that $(y_{t+1}-\alpha_1)$ is a MD sequence and, hence, that y_{t+1} is not predictable by x_t which entails that $\beta_1=0$ in (2.1).[3] The alternative hypothesis is that y_{t+1} is predictable by x_t, in which case $\beta_1 \neq 0$. As discussed in Section 1, it is important for practical purposes to allow for the possibility of strong persistence in the predictor variable x_t and to allow the shocks driving the predictor, v_t in (2.2), to be contemporaneously correlated with the unpredictable component of y_t; that is, u_t in (2.1). We will allow for both of these through Assumptions 1–4 which follow.

First, with respect to the degree of persistence in x_t, this is controlled via the parameter ρ. We allow x_t to be either weakly, moderately, or strongly persistent through the following assumptions. Second, in line with the literature on predictive regression with financial data (see in particular the arguments of Phillips and Lee, 2013), we focus on parameters β_1 that are small in magnitude, reflecting the fact that the signal-to-noise ratio in the typical predictive regression is low. To capture this in the asymptotics we take β_1 to be local to zero, $\beta_1 = o(1)$, at rates specific to the persistence of the putative predictor. In particular, this will allow us to obtain expressions for the local power of long-horizon predictability test procedures.

Assumption 1. The data are generated according to (2.1) and (2.2) with initial condition ξ_1 which is bounded in probability.

Assumption 2. Exactly one of the three following conditions holds true:

(i) **Strongly persistent predictors:** The autoregressive parameter ρ in (2.2) is local-to-unity with $\rho := 1-c/T$, where c is a fixed constant. Furthermore, $\beta_1 := T^{-1/2-\eta/2}b$, where b is a finite constant and where $\eta \in (0,1)$ is the IVX tuning parameter discussed in Section 3.3.

(ii) **Weakly persistent predictors:** The autoregressive parameter ρ in (2.2) is fixed and bounded away from unity, $|\rho| < 1$. Furthermore, $\beta_1 := T^{-1/2}b$, with b a finite constant.

(iii) **Moderately persistent predictors:** The autoregressive parameter ρ in (2.2) is moderately close to unity with $\rho := 1 - c/T^{\kappa}$, where $c > 0$ is a fixed constant and $\kappa \in (0,1)$. Furthermore, $\beta_1 := T^{-1/2-\min\{\eta,\kappa\}/2}b$, where b is a finite constant and where $\eta \in (0,1)$ is the IVX tuning parameter discussed in Section 3.3.

Remark 2. Many commonly used predictors are strongly persistent, exhibiting sums of sample autoregressive coefficients which are close to or only slightly smaller than unity. Near-integrated asymptotics have been found to provide better approximations for the behaviour of test statistics in such circumstances; see, *inter alia*, Elliott and Stock (1994). However, not all (putative) predictors are strongly persistent and a large part of the literature works with models which take x_t to be generated from a stable autoregression; see, for example, Amihud and Hurvich (2004). While the long-horizon predictability tests developed in Valkanov (2003) and Hjalmarsson (2011) are only valid for the case where x_t is strongly persistent, we allow for either of these possibilities to hold for x_t. Kostakis et al. (2015) extend the range of possible degrees of persistence by allowing x_t to be mildly integrated, and we also allow for this persistence class through Assumption 2(iii). Because it is very difficult to distinguish between these three types of persistence in practice, covering all three within Assumption 2 provides an approach that applied researchers can use with some confidence. It is, however, important to stress that Assumption 2 does not allow for fractionally integrated predictors. In the context of short-horizon ($h=1$) predictability testing, a number of important contributions allow for fractionally integrated predictors; see, *inter alia*, Maynard and Phillips (2001), Maynard and Shimotsu (2009), Bauer and Maynard (2012), and Andersen and Varneskov (2021a, 2021c). Within the framework of Andersen and Varneskov (2021a,b) also allow for "imperfect" predictors whereby a component of the conditional mean of returns exists that is not linearly spanned by the chosen predictor(s); see also Georgiev et al. (2018). So far as we are aware, neither fractionally integrated nor imperfect predictors have been considered in the long-horizon testing literature and, as such, constitute important areas for further research. ♢

Remark 3. The (Pitman) neighbourhoods within which our proposed tests will have non-trivial power can be seen to depend on the persistence of the regressor and, in the case where the predictor is strongly or moderately persistent, additionally on the IVX tuning parameter, η. We note that it is only in the strongly persistent case where the localisation rates on β_1 given in Assumption 2 are less favourable than those which apply in connection with OLS estimation and testing, for which the relevant localisation is given by $\beta_1 := T^{-1}b$. This is common to all IVX approaches, and this power loss is offset by the size control offered by IVX estimation. In related work, Kostakis et al. (2018) deal with the case where the slope coefficient in the long-horizon predictive regression can be of larger magnitude, captured by assuming β_1 is fixed as $T \to \infty$. In such cases, estimators from the long-horizon regression may exhibit bias depending on the persistence of the predictor; see Kostakis et al. (2018) for details. Examining the proofs in the Supplementary Appendix (see e.g. for strong persistence the proof of Theorem 4.1), it can be seen that our methods might be expected to handle the case of fixed β_1 provided one places additional restrictions on the horizon period, h, in particular that h is fixed. However, we will not pursue these issues further here and will work within the relevant localisations on β_1 given in Assumption 2. ♢

[3] All of the tests we discuss in this paper could equally well be used to test the null hypothesis $H_0 : \beta_1 = \beta_0$, say by replacing y_{t+1} by $y_{t+1} - \beta_0 x_t$, but as the focus in equity forecasting is on testing the null hypothesis of a zero coefficient on the predictor we will restrict our discussion to $\beta_0 = 0$.

To complete the specification of our predictive regression model, we make the following assumptions with regard to the error terms, u_t and v_t, which are designed to allow for empirically relevant features frequently found in economic and financial time series.

Assumption 3. The errors u_t and v_t in (2.1) and (2.2), respectively, are characterised as

$$u_t = \gamma \varpi_t + \varepsilon_t, \qquad t \in \mathbb{Z} \tag{2.3}$$

$$v_t = a_1 v_{t-1} + \cdots + a_{p-1} v_{t-p+1} + \varpi_t, \tag{2.4}$$

where $(\varepsilon_t, \varpi_t)'$ is serially uncorrelated, satisfying the conditions of Assumption 4, and the lag polynomial $A(L) := 1 - a_1 L - \cdots - a_{p-1} L^{p-1}$ is invertible. For further reference we define $\omega := \left(1 - \sum_{k=1}^{p-1} a_k\right)^{-1}$ and we denote by ϕ_k the coefficients of the lag polynomial $(1-\rho L)A(L)$; in case of weak persistence, let b_k denote the coefficients of the (infinite-order) MA representation of the process ξ_t, $\sum_{k\geq 0} b_k L^k = ((1-\rho L)\,A(L))^{-1}$.

Assumption 4. Let

$$\begin{pmatrix} \varepsilon_t \\ \varpi_t \end{pmatrix} := \begin{pmatrix} \sigma_{\varepsilon t}\zeta_{\varepsilon t} \\ \sigma_{\varpi t}\zeta_{\varpi t} \end{pmatrix}$$

where $\boldsymbol{\zeta} := (\zeta_{\varepsilon t}, \zeta_{\varpi t})'$ is a uniformly L_4-bounded stationary and ergodic martingale difference [MD] sequence satisfying $\mathrm{E}\left(\boldsymbol{\zeta}_t \boldsymbol{\zeta}_t'\right) = \mathbf{I}_2$ and $\mathrm{E}\left(\left\|\mathrm{E}_0\left(\sum_{t=1}^{T}(\boldsymbol{\zeta}_t\boldsymbol{\zeta}_t' - \mathbf{I}_2)\right)\right\|^2\right) = O(T^{2\epsilon})$ for some $\epsilon < \frac{1}{2}$, with $\mathrm{E}_0(\cdot)$ denoting expectation conditional on $\{\boldsymbol{\zeta}_{-i}\}_{i=0}^{\infty}$ and $\mathbf{I}_k$ the $k \times k$ identity matrix. Furthermore, let $\sigma_{\varepsilon t} := \sigma_\varepsilon\left(\frac{t}{T}\right)$ and $\sigma_{\varpi t} := \sigma_\varpi\left(\frac{t}{T}\right)$, where $\sigma_\cdot(\cdot)$ are piecewise Lipschitz-continuous bounded, non-stochastic functions on $(-\infty, 1]$, which are bounded away from zero.

Remark 4. Assumption 3 imposes, through (2.4), the condition that the errors v_t driving ξ_t in (2.2) follow a finite-order autoregression (AR) such that the predictor x_t is an $AR(p)$ process with $p \geq 1$; Valkanov (2003) makes the same assumption. The finite-order AR assumption is required for the tests developed in Section 4.2 which make use of the residual augmented regression approach of Demetrescu and Rodrigues (2022). Here the transformed long-horizon predictive regression is augmented by the residuals from fitting an $AR(p)$ model to the predictor x_t. We conjecture that these tests would also be asymptotically valid under a linear process type assumption on v_t, provided the truncation lag for the fitted autoregression is allowed to increase at a suitable rate with the sample size, T. It is, however, important to note that the long-horizon predictability tests developed in both Hjalmarsson (2011) and Xu (2020) are based on the considerably more restrictive assumption that $A(L) = 1$, such that v_t is serially uncorrelated and, hence, that x_t follows an $AR(1)$. ◊

Remark 5. Assumption 4 is similar to Assumption 3 of Demetrescu et al. (2022b) and we defer to Demetrescu et al. (2022b) for a detailed discussion of these conditions. Briefly, it allows for unconditional time heteroskedasticity of quite general form in the innovations $(\varepsilon_t, \varpi_t)'$ through the functions $\sigma_\varepsilon(\cdot)$ and $\sigma_\varpi(\cdot)$ which allow both ε_t and ϖ_t to display time-varying unconditional variances and for both contemporaneous and time-varying (unconditional) correlation between ε_t and ϖ_t. The MD structure placed on $\boldsymbol{\zeta}_t$ allows for conditional heteroskedasticity of a general form obviating the need to choose a specific parametric model by instead adopting an explicit assumption of martingale approximability whereby $\mathrm{E}(\|\mathrm{E}_0(\sum_{t=1}^{T}(\boldsymbol{\zeta}_t\boldsymbol{\zeta}_t' - \mathbf{I}_2))\|^2) = O(T^{2\epsilon})$ for some $\epsilon < \frac{1}{2}$, where ϵ controls the degree of persistence permitted in the conditional variances. Stationary vector GARCH processes with finite fourth-order moments satisfy this condition with $\epsilon = 0$, although Assumption 4 is considerably more general as it also allows for asymmetric effects in the conditional variance. Stationary autoregressive stochastic volatility processes as, for example, are assumed in Johannes et al. (2014) are also permitted. ◊

Remark 6. Assumption 4 is considerably weaker than the corresponding conditions imposed by the leading tests for long-horizon predictability in the literature. Valkanov (2003), Phillips and Lee (2013) and Xu (2020) all impose conditional (and, hence, unconditional) homoskedasticity on the innovations. In Remark 12, page 4414, Xu (2020) suggests the possibility that his approach could be modified (but does not actually develop such a modification) to allow for the case where the innovations can be conditionally heteroskedastic satisfying essentially the same conditions as are imposed in Assumption INNOV of Kostakis et al. (2015, p. 1512) for their short-horizon predictability tests. These conditions are, however, still considerably more restrictive than Assumption 4 as, in additional to imposing unconditional homoskedasticity, they also impose the condition that the error term in (2.1) is generated according to a stationary finite-order GARCH(p, q) model with finite fourth moments. Hjalmarsson (2011) allows for conditional heteroskedasticity but again assumes unconditional homoskedasticity; notice, however, that Hjalmarsson (2011) does not allow for the case where x_t is weakly persistent, which as discussed in Remark 12 of Xu (2020), is the case where allowing for conditional heteroskedasticity is most problematic. ◊

Remark 7. The error term u_t in (2.1) is formulated as a linear combination of the uncorrelated innovations ε_t and ϖ_t. The degree of endogeneity present is measured by the correlation between u_t and ϖ_t, defined as $\phi_t := \gamma\sigma_{\varpi t}/\sigma_{ut}$, which can be either constant or time-varying under Assumption 4. Where $\gamma = 0$, $u_t = \varepsilon_t$ and, hence, the error term in (2.1) is uncorrelated with the innovation driving the predictor, so that $\phi_t = \phi = 0$ for all t. The constant correlation case, where $\phi_t = \phi$ for all t, can occur either where $\sigma_{\varpi t}$ and σ_{ut} are both time-invariant, or where any time-variation is common to both $\sigma_{\varpi t}$ and σ_{ut}. Notice that Assumption 3 restricts γ to be time-invariant. This assumption is needed to establish the large sample validity of the residual augmentation method used in Section 4.2. It might be possible to relax the assumption of a constant γ by using local (nonparametric) estimation thereof, but we leave such developments for future research. The restriction that γ is constant is common to all of the existing long-horizon tests discussed above. ◊

2.2. The long-horizon predictive regression specification

The most common long-horizon predictive regression specification used in empirical analysis results from the h-period, $h \geq 1$, temporal aggregation of (2.1) and is given by

$$y_{t+h}^{(h)} = \alpha_h + \beta_h x_t + error_{t+h}, \;\; t = 1, \ldots, T-h \tag{2.5}$$

where $y_{t+h}^{(h)} := \sum_{j=1}^{h} y_{t+j}$ is the h-period cumulative variable to be predicted. Notice that for $h = 1$, (2.5) is simply the short-horizon predictive regression in (2.1). To gain further insight into the specific features of (2.5), let us examine the h-horizon cumulated dependent variable $y_{t+h}^{(h)}$ more closely. From (2.1), the long-horizon predictive model can be written as,

$$y_{t+h}^{(h)} = h\alpha_1 + \beta_1 \sum_{j=0}^{h-1} x_{t+j} + u_{t+h}^{(h)}, \tag{2.6}$$

where, from (2.3), $u_{t+h}^{(h)} := \sum_{j=1}^{h} u_{t+j} = \gamma v_{t+h}^{(h)} + \varepsilon_{t+h}^{(h)}$, with $v_{t+h}^{(h)}$ and $\varepsilon_{t+h}^{(h)}$ defined implicitly.

The properties of the cumulated variable, $\sum_{j=0}^{h-1} x_{t+j}$, and, as a result, the implied relationships between β_h in the long-horizon predictive regression in (2.5) and β_1 in the underlying DGP in (2.1) and between the regression error in (2.5) and the innovation sequences u_t and v_t in the DGP, turn out to depend on the particular persistence class to which x_t belongs. To see why consider first the case where the predictor is either strongly or moderately persistent. Using the autoregressive representation of the predictor in (2.2), which can be written as $x_{t+1} = \mu_x(1-\rho) + \rho x_t + v_{t+1}$, by recursive substitution we then have that,

$$\sum_{j=0}^{h-1} x_{t+j} = I_{h\geq 2} \sum_{j=1}^{h-1} \sum_{i=1}^{j} \rho^{i-1}\mu_x(1-\rho) + \sum_{j=0}^{h-1} \rho^j x_t + I_{h\geq 2} \sum_{j=1}^{h-1} \sum_{i=1}^{h-j} \rho^{i-1} v_{t+j} \tag{2.7}$$

where $I_{h\geq 2}$ is an indicator variable which takes the value 1 when $h \geq 2$ and 0 otherwise. Consequently, replacing $\sum_{j=0}^{h-1} x_{t+j}$ in (2.6) by the expression on the right-hand side of (2.7), the general representation of the long-horizon predictive regression model specification is obtained,

$$y_{t+h}^{(h)} = \alpha_h + \beta_h x_t + w_{t+h}^{(h)} \tag{2.8}$$

where $\alpha_h := h\alpha_1 + \beta_1 I_{h\geq 2} \sum_{j=1}^{h-1} \sum_{i=1}^{j} \rho^{i-1}\mu_x(1-\rho)$, $\beta_h := \beta_1 \sum_{j=0}^{h-1} \rho^j$ and $w_{t+h}^{(h)} := u_{t+h}^{(h)} + \beta_1 I_{h\geq 2} \sum_{j=1}^{h-1} \sum_{i=1}^{h-j} \rho^{i-1} v_{t+j}$. Consequently, in the strongly persistent case where $\rho = 1 - c/T$, it can be seen that $\beta_h \approx h\beta_1$ in (2.8), provided $h/T \to 0$. This approximate relation also holds in the moderately persistent case, where $\rho = 1 - c/T^\kappa$, $\kappa \in (0,1)$, provided $h/T^\kappa \to 0$. Notice that both of these rate conditions are implied by the relevant rate conditions needed to establish the large sample properties of our proposed estimators and test statistics in Section 4.3.

In the weakly persistent case, although (2.8) still holds, the implied expressions for β_h and $w_{t+h}^{(h)}$ change relative to the strongly and moderately persistent cases. In particular, we now write

$$\sum_{j=0}^{h-1} x_{t+j} = I_{h\geq 2} \sum_{j=1}^{h-1} \left(1 - \frac{\theta_j}{\theta_0}\right) \mu_x + \sum_{j=0}^{h-1} \frac{\theta_j}{\theta_0} x_t + I_{h\geq 2} \sum_{j=1}^{h-1} v_{t+j}^{\perp}$$

with $v_{t+j}^{\perp} := \xi_{t+j} - \frac{\theta_j}{\theta_0}\xi_t$, or $x_{t+j} = \mu_x\left(1 - \frac{\theta_j}{\theta_0}\right) + \frac{\theta_j}{\theta_0} x_t + v_{t+j}^{\perp}$ with $\theta_j := \sum_{k\geq 0} b_k b_{k+j}$, $j = 0, \ldots, h-1$, where b_k are the coefficients of the (infinite-order) *MA* representation of the process ξ_t.[4] Consequently, in the weakly persistent case we have that $\beta_h = \beta_1 \sum_{j=0}^{h-1} \frac{\theta_j}{\theta_0}$ in (2.8), together with $w_{t+h}^{(h)} := u_{t+h}^{(h)} + \beta_1 I_{h\geq 2} \sum_{j=1}^{h-1} v_{t+j}^{\perp}$.[5] As in the strongly and moderately persistent cases, β_h can be seen to be proportional to β_1, albeit with a different factor of proportionality.

[4] We note in passing that, under unconditional homoskedasticity, the quantities $v_{t+j}^{\perp}$ are projection errors from an orthogonal projection of ξ_{t+j} onto ξ_t, while, under time-varying volatility, they can be interpreted as local counterparts thereof.

[5] A different expression is given for $\sum_{j=0}^{h-1} x_{t+j}$ compared to the strongly persistent case because $v_{t+j}^{\perp}$ are, by construction, orthogonal to x_t; indeed, this orthogonality property is a key ingredient needed for the asymptotic analysis of the weakly persistent case; see the proofs of Theorems 4.3 and 4.4 in the Supplementary Appendix.

To distinguish between these expressions for β_h in the three persistence classes considered, we introduce the additional notation $\beta_h^{(i)} = \beta_h^{(iii)} := \beta_1 \sum_{j=0}^{h-1} \rho^j$ for the strongly and moderately persistent cases, respectively, and, for the weakly persistent case, $\beta_h^{(ii)} := \beta_1 \sum_{j=0}^{h-1} \frac{\theta_j}{\theta_0}$. We will use this notation wherever we need to distinguish explicitly between the three cases (e.g. when discussing the limiting behaviour of the estimators from the long-horizon predictive regression (2.8) in Section 4.2). Should a distinction not be essential for the exposition, we will simply refer to β_h without specifying the persistence type.

Irrespective of the persistence type, we note from the foregoing algebra that $\beta_h \neq 0$ for $h > 1$ whenever $\beta_1 \neq 0$. The coefficient β_h in (2.8) is therefore empirically useful, as a finding of statistical significance from an estimate of β_h can still be interpreted as evidence of long-horizon predictability, given that if there is no short-run predictability ($\beta_1 = 0$) then there is also no predictability at other horizons ($h \geq 1$). Consequently, under suitable assumptions, the null hypothesis of no-predictability, H_0, can be tested using statistics computed from (2.5). If x_t is weakly persistent, tests can be based on conventional regression t-statistics, provided h is fixed. However, care is needed because the dynamics of the error term $w_{t+h}^{(h)}$ in (2.8) differ according to whether there is predictability or not. In particular, if $\beta_1 = 0$ (and, hence, $\beta_h = 0$), then this error term is an $MA(h-1)$ process. Where $\beta_1 \neq 0$, any serial correlation in v_t will change the dynamics of $w_{t+h}^{(h)}$; for example, if v_t were an $MA(1)$ process, then $w_{t+h}^{(h)}$ will follow an $MA(h)$ process.[6] To account for these dynamics the t-statistic needs to be based on either HAC (Newey and West, 1987) or Hodrick (1992) standard errors. Although these are asymptotically equivalent, simulation evidence presented in Ang and Bekaert (2007) suggests the latter deliver tests with better finite sample behaviour. Moreover, Nelson and Kim (1993) show that finite sample biases present in the OLS estimate, $\hat{\beta}_1^{OLS}$ say, of β_1 from the short-horizon predictive regression in (2.1) (which are larger, other things equal, the greater the persistence of the predictor and the higher the endogeneity correlation between the innovations) are exacerbated by the long-horizon aggregation. Consequently, several bias correction approaches have been suggested for the case where x_t is weakly persistent; see for instance, Stambaugh (1999), Lewellen (2004), Amihud and Hurvich (2004), Amihud et al. (2009, 2010) and Kim (2014).

The standard t-tests and bias-correction methods discussed above are, however, not valid when x_t is strongly persistent. In particular, the limiting null distribution of the t-statistic is not pivotal because the endogeneity present in the model is not accounted for.

3. Extant tests allowing for strongly persistent predictors

In this section we present a brief overview of test procedures for long-horizon predictability which allow for strongly persistent predictors.

3.1. Bonferroni-based tests

Assuming x_t is a strongly persistent (near-integrated) predictor, Hjalmarsson (2011) builds on the approach of Amihud and Hurvich (2004) to compute a second-order bias corrected estimate of β_h in order to develop a feasible long-horizon predictability test. In the context of (2.8), this is based on the infeasible augmented regression,

$$y_{t+h}^{(h)} = \alpha_h + \beta_h x_t + \gamma \varpi_{t+h}^{(h)} + \varepsilon_{t+h}^{(h)} + r_{t+h}, \ t = 1, \ldots, T-h, \tag{3.1}$$

where $r_{t+h} := w_{t+h}^{(h)} - u_{t+h}^{(h)} = \beta_1 I_{h\geq 2} \sum_{j=1}^{h-1} \sum_{i=1}^{h-j} \rho^{i-1} v_{t+j}$, and, from Assumption 3 and (2.2),

$$\varpi_{t+h}^{(h)} := \sum_{j=1}^{h} \varpi_{t+j} = \sum_{j=1}^{h} \left[(x_{t+j} - \mu_x) - \sum_{k=1}^{p-1} \phi_k (x_{t+j-k} - \mu_x) \right].$$

The inclusion of $\varpi_{t+h}^{(h)}$ in (3.1) serves to remove the endogeneity bias present in standard OLS estimation of (2.8). Assuming $(u_{t+1}, v_{t+1})'$ is an unconditionally homoskedastic MD process, Hjalmarsson (2011) shows that, for fixed h, the infeasible scaled OLS estimator from (3.1), $\hat{\beta}_h^I$ say, when divided by h has a mixed normal null limiting distribution whose variance does not depend on h.

In order to obtain a feasible version of (3.1), Hjalmarsson (2011) adopts an approach based on Bonferroni-bounds. This involves computing a first-stage confidence interval for the local to unity parameter c which is then used to develop a test for long-horizon predictability based on a bias reduced estimate of β_h (see also Campbell and Yogo, 2006a). Denoting this confidence interval, with confidence level $100(1-\lambda_1)\%$, by $[\underline{c}_{\lambda_1}, \bar{c}_{\lambda_1}]$, feasible, yet conservative, versions of tests for $H_0 : \beta_h = 0$ against $H_A : \beta_h > 0$ and $H_0 : \beta_h = 0$ against $H_A : \beta_h < 0$, which we will generically define as t_h^{Bonf}, are, respectively,

$$h^{-1/2} t_{h,\bar{c}^*}^{min} := \min_{\bar{c} \in [\underline{c}_{\lambda_1}, \bar{c}_{\lambda_1}]} h^{-1/2} t_{h,\bar{c}}^{OLS} > z_{\lambda_2} \tag{3.2}$$

[6] Technically, we exclude a finite-order MA structure of the increments v_t; the MA example is still of relevance given that we (quite plausibly) conjecture in Remark 4 that MA processes could be allowed for under suitable conditions.

and

$$h^{-1/2}t^{max}_{h,\tilde{c}^*} := \max_{\tilde{c}\in[\underline{c}_{\lambda_1},\bar{c}_{\lambda_1}]} h^{-1/2}t^{OLS}_{h,\tilde{c}} < z_{\lambda_2}, \tag{3.3}$$

with $t^{OLS}_{h,\tilde{c}}$ being the OLS t-ratio for $\beta_h = 0$ computed from a feasible version of (3.1) where $\hat{\varpi}^{(h)}_{t+h}$ is obtained based on $\hat{\rho} := 1-\tilde{c}/T$ with $\tilde{c} \in [\underline{c}_{\lambda_1}, \bar{c}_{\lambda_1}]$, and z_{λ_2} is the standard normal critical value associated with the significance level λ_2 of the test, such that $\lambda_1 + \lambda_2 = \lambda$, where λ is the desired significance level of the test. In other words, a rejection occurs for the Bonferroni bounds test only if it occurs for every possible value of c in the first stage confidence interval. The requirement that $\lambda_1+\lambda_2 = \lambda$ can lead to overly conservative tests and, in practice, adjustments to λ_1, to shrink the coverage rates of the confidence intervals for c, are typically recommended; see Cavanagh et al. (1995) and Campbell and Yogo (2006b). In the linear predictive regression context, Hjalmarsson (2012) finds that his test has better power properties than the earlier test of Valkanov (2003). It is important to stress that these Bonferroni-based tests are developed under the assumption that x_t is strongly persistent and are not valid if x_t is weakly persistent. As we will see from the simulation results in Section 6, these tests do indeed not perform well when x_t is weakly persistent. Moreover, it is important to note that Hjalmarsson (2011)'s approach is based on the assumption that $A(L) = 1$ in Assumption 3, such that x_t follows an $AR(1)$ model.

3.2. Xu (2020)'s implied test

Xu (2020) develops an alternative approach to testing for long-horizon predictability which allows for the case where the predictor, x_t, is either strongly or weakly persistent based on the computation of the implied long-horizon coefficients from short-horizon regression estimates; see, among others, Campbell and Shiller (1987), Kandel and Stambaugh (1996), Hodrick (1992) and Bekaert and Hodrick (1992). This choice of estimator is motivated by the observation that short-horizon estimation is often more efficient than long-horizon estimation; see, for example, Boudouk and Richardson (1994). Xu (2020) bases his test on the implied estimator of β_h, $\tilde{\beta}_h := \hat{\beta}^{OLS}_1 \sum_{j=0}^{h-1} \hat{\rho}^j$ where $\hat{\beta}^{OLS}_1$ and $\hat{\rho}$ are the OLS estimates obtained from (2.1) and (2.2), respectively.

The implied long-horizon predictability test of Xu (2020) is based on the statistic

$$t^{Xu}_h = \upsilon^{-1}_{IM}\tilde{\beta}_h \tag{3.4}$$

where $\upsilon^2_{IM} := \hat{\boldsymbol{q}}\hat{\boldsymbol{\Omega}}(\sum_{t=1}^{T-1}\bar{x}_t)\hat{\boldsymbol{q}}'$ with $\hat{\boldsymbol{q}} := (\hat{q}_1, \hat{q}_2)$, where $\hat{q}_1 := \sum_{j=0}^{h-1}\hat{\rho}^j$ and $\hat{q}_2 := \hat{\beta}^{OLS}_1 \sum_{j=0}^{h-1} j\hat{\rho}^{j-1}$, and where the vector of OLS residuals, $\hat{\mathbf{e}}_{t+1} := (\hat{u}_{t+1}, \hat{v}_{t+1})'$, computed from (2.1) and (2.2), is used to estimate the covariance matrix of $\mathbf{e}_{t+1}$, $\hat{\boldsymbol{\Omega}} := \sum_{t=1}^{T-1}\hat{\mathbf{e}}_{t+1}\hat{\mathbf{e}}'_{t+1}$.

Under the assumption of conditionally homoskedastic MD innovations, Xu (2020) shows that under $H_0 : \beta_h = 0$: (i) if x_t is strongly persistent, $t^{Xu}_h \xrightarrow{d} \phi\left[(\int_0^1 \bar{J}^2_c(s))^{-1/2}\int_0^1 \bar{J}_c(s)dW(s)\right] + (1-\phi^2)^{1/2}\mathcal{Z}$, where ϕ denotes the (time-invariant) correlation between the innovations u_{t+1} and ϖ_{t+1} in (2.1) and (2.2) (see Assumption 3), J_c an OU process driven by the standard Wiener process W and $\mathcal{Z}$ is a standard normal variate independent of W; and (ii) if x_t is weakly persistent, $t^{Xu}_h \xrightarrow{d} N(0,1)$. These results show that the limiting null distribution of the test statistic changes depending on the persistence of the predictor and the magnitude of ϕ. To account for this, Xu (2020) proposes two alternative ways to compute the necessary critical values. One is based on a Bonferroni procedure and the other, which is the one he recommends, uses a bias-corrected wild bootstrap approach (residual-based with recursive design), although Xu (2020) does not formally establish the asymptotic validity of the latter. It is important to note that the asymptotic validity of Xu (2020)'s test, like that of Hjalmarsson (2011), relies on the assumption that x_t is an $AR(1)$ process, so that $A(L) = 1$ in Assumption 3. The assumption of no serial correlation in v_t is essential for Xu (2020)'s approach under weak persistence, as in this case we have that $\beta_h = \beta^{(ii)}_h = \beta_1\sum_{j=0}^{h-1}\frac{\theta_j}{\theta_0}$ (see Section 2.2), implying that $\beta_1\sum_{j=0}^{h-1}\rho^j$ is not the correct quantity to base a test on.

3.3. Reversed regression-based tests

An alternative to the use of HAC or Hodrick (1992) standard errors to account for the serial correlation in the error term in the long-horizon predictive regression model in (2.8) discussed in Section 2.2 is to use an alternative regression specification that is designed to explicitly account for the overlapping data issue. One such approach is to use so-called *reverse regressions*; see, among others, Jegadeesh (1991) and Cochrane (1991). This approach, instead of being based on the regression from the h-period returns on a predetermined variable, as in (2.5), is based on a regression of single period returns on the same predetermined variable but aggregated over h-periods. Specifically, this *reverse regression* formulation is given by,

$$y_{t+h} = \alpha^{rev}_h + \beta^{rev}_h x^{(h)}_{t+h-1} + u_{t+h}, \ t = 1,\ldots,T-h \tag{3.5}$$

where $x^{(h)}_{t+h-1} := \sum_{j=0}^{h-1} x_{t+j}$. See also Hodrick (1992), Maynard and Ren (2014), Ang and Bekaert (2007), and Wei and Wright (2013), *inter alia*. It is seen from (3.5) that the error term is u_{t+h} which is serially uncorrelated. An implication of this is that the IVX estimation and hypothesis testing methods like in Kostakis et al. (2015) can be directly applied to (3.5), which is not the case for (2.8) because of the induced serial correlation in $w^{(h)}_{t+h}$.

The OLS estimate of β_h^{rev} from (3.5) is given by $\hat{\beta}_h^{rev} := (\sum_{t=1}^{T-h} \bar{x}_{t+h-1}^{(h)} \bar{y}_{t+h})/(\sum_{t=1}^{T-h} (\bar{x}_{t+h-1}^{(h)})^2)$, where for a generic sequence $\{w_t\}_{t=a}^{b}$, $\bar{w}_t := w_t - (b-a+1)^{-1} \sum_{s=a}^{b} w_s$. It is not hard to establish that, regardless of whether x_t is weakly or strongly persistent, $\hat{\beta}_h^{rev} = (\sum_{t=1}^{T-h} \bar{x}_t^2)/(\sum_{t=1}^{T-h} (\bar{x}_{t+h-1}^{(h)})^2)\hat{\beta}_h^{OLS} + o_p(1)$, where $\hat{\beta}_h^{OLS}$ is the OLS estimate of β_h from (2.5). Motivated by this, Phillips and Lee (2013) develop a long-horizon predictability test based on applying IVX estimation to the reverse regression in (3.5). Specifically, they use the IVX instrument z_t used by Kostakis et al. (2015), which is constructed from the predictor as,

$$z_t := (1 - \varrho L)_+^{-1} \Delta x_t = \sum_{j=0}^{t} \varrho^j \Delta x_{t-j}. \tag{3.6}$$

The persistence of z_t is controlled by setting $\varrho := 1 - \frac{a}{T^\eta}$, with $0 < \eta < 1$. If x_t is near integrated, this makes z_t approximately mildly integrated (and thus of lower persistence), while if x_t is weakly persistent then one may decompose $z_t = x_t - \mu_x + r_t$, where the rest term satisfies $r_t \to 0$ as $t \to \infty$ and can be controlled for in the relevant expressions; see e.g. Lemma S.3 in the Supplementary Appendix for details. Because the reversed regression in (3.5) features $x_{t+h-1}^{(h)} := \sum_{j=0}^{h-1} x_{t+j}$, the long-horizon IVX approach is based on instrumenting $x_{t+h-1}^{(h)}$ by $z_{t+h-1}^{(h)} := \sum_{j=0}^{h-1} z_{t+j}$.

Allowing the forecast horizon, h, to grow at rate $T^{1/2}T^{-\eta} + T^\eta h^{-1} + h/T \to 0$, such that it increases at a slower rate than the sample size T, but faster than the (user-controlled) degree of mild integration of the instrument, Phillips and Lee (2013)'s long-horizon predictability statistic is

$$t_{h,ivx}^{rev,PL} := (\mathcal{H}^{-1}\hat{\sigma}_u^2)^{-1/2} \hat{\beta}_{h,ivx}^{rev} \tag{3.7}$$

where $\hat{\beta}_{h,ivx}^{rev} := \left(\sum_{t=1}^{T-h} \bar{x}_{t+h-1}^{(h)} z_{t+h-1}^{(h)}\right)^{-1} \sum_{t=1}^{T-h} z_{t+h-1}^{(h)} \bar{y}_{t+h}$, $\mathcal{H} := \left[\mathcal{H}_{\bar{x}^{(h)}z^{(h)}} (\mathcal{H}_{z^{(h)}z^{(h)}})^{-1} \mathcal{H}'_{x^{(h)}z^{(h)}}\right]^{-1}$, $\mathcal{H}_{x^{(h)}z^{(h)}} := \sum_{t=1}^{T-h} x_{t+h-1}^{(h)} z_{t+h-1}^{(h)}$, $\mathcal{H}_{z^{(h)}z^{(h)}} := \sum_{t=1}^{T-h} (z_{t+h-1}^{(h)})^2$ and $\hat{\sigma}_u^2 := \frac{1}{T-1} \sum_{t=1}^{T-1} \hat{u}_{t+1}^2$. Assuming that the innovations are conditionally homoskedastic, Phillips and Lee (2013) show that $t_{h,ivx}^{rev,PL}$ has a standard normal limiting distribution under H_0. It should be noted that Phillips and Lee (2013) do not formally allow for the possibility that x_t is weakly persistent.

4. Transformed regression-based long-horizon predictability tests

In this section we introduce our new approach to long-run predictability testing which builds on the IVX framework of Kostakis et al. (2015) and the augmented regression approach of Amihud and Hurvich (2004), Hjalmarsson (2011) and Demetrescu and Rodrigues (2022). The tests we develop are asymptotically valid regardless of whether the predictor is weakly, moderately or strongly persistent, without requiring either a Bonferroni or wild bootstrap scheme for implementation, and do not require that the predictor follows an $AR(1)$ process.

4.1. Transformed regression IVX based tests

In a recent paper Britten-Jones et al. (2011) develop a method for conducting inference in linear regression models with overlapping observations and stationary covariates. Before showing how we can apply this approach to the specific setting considered in this paper, we first briefly review the transformed regression approach. To that end, suppose we have a generic linear regression model $\mathbf{A}_h\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \mathbf{u}$, where $\mathbf{y}$ is the $(T-1)$-vector of single period returns, $\mathbf{A}_h$ is the *known* $(T-h)\times(T-1)$ aggregation matrix with entries $a_{ij} = 1$ if $i \le j \le i+h-1$ and zero otherwise, $i = 1, \ldots, T-h$, such that $\mathbf{A}_h\mathbf{y}$ is the vector of (overlapping) h-period returns, $\mathbf{X}$ the regressor matrix with associated vector of coefficients, $\boldsymbol{\beta}$ and $\mathbf{u}$ is the error vector. Britten-Jones et al. (2011) demonstrate that the OLS estimate of $\boldsymbol{\beta}$ from this regression, $\tilde{\boldsymbol{\beta}}$ say, is numerically identical to the OLS estimate from the transformed regression $\mathbf{y} = \tilde{\mathbf{X}}\boldsymbol{\beta} + \tilde{\mathbf{u}}$, where $\tilde{\mathbf{X}} := \mathbf{A}_h'\mathbf{X}(\mathbf{X}'\mathbf{A}_h\mathbf{A}_h'\mathbf{X})^{-1}\mathbf{X}'\mathbf{X}$. The associated estimation error from the transformed regression can then be written as $\tilde{\boldsymbol{\beta}} - \boldsymbol{\beta} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{A}_h\tilde{\mathbf{u}}$, which is seen to depend on the autocorrelation structure of $\tilde{\mathbf{u}}$, the disturbance term in the transformed (non-overlapping) regression, rather than on $\mathbf{u}$, the disturbance in the untransformed (overlapping) regression. The part of the autocorrelation in $\mathbf{u}$ induced by the temporal aggregation (through $\mathbf{A}_h$) is therefore explicitly accounted for and does not need to be estimated from the data when conducting inference on $\boldsymbol{\beta}$ via the transformed regression. In the context of the DGP in (2.1)–(2.2), a key implication of this result is that while the IVX approach of Kostakis et al. (2015) cannot be used to conduct valid inference on β_h in (2.8) under Assumption 3, because of the autocorrelation present in the error term $u_{t+h}^{(h)}$ induced by temporal aggregation, it can when applied to the transformed regression analogue of (2.8).

To that end, consider again (2.8). Using the general result above it can be shown[7] that the OLS estimator of the slope parameter β_h, $\hat{\beta}_h^{OLS} := (\sum_{t=1}^{T-h} \bar{x}_t \bar{y}_{t+h}^{(h)})/(\sum_{t=1}^{T-h} \bar{x}_t^2)$, can be written equivalently as

$$\hat{\beta}_h^{trf} := \frac{\sum_{t=1}^{T-1} \bar{x}_t^{trf,(h)} \bar{y}_{t+1}}{\sum_{t=1}^{T-h} \bar{x}_t^2} \tag{4.1}$$

[7] Derivations for the functional forms of the estimators and statistics from the transformed regression given in this section are provided in section S.2 of the Supplementary Appendix.

where

$$\bar{x}_t^{trf,(h)} := \begin{cases} \sum_{i=1}^{t} \bar{x}_i & \text{for } t = 1, \ldots, h-1 \\ \bar{x}_t^{(h)} := \sum_{i=1}^{h} \bar{x}_{t-h+i} & \text{for } h \le t \le T-h \\ \sum_{i=t-h+1}^{T-h} \bar{x}_i & \text{for } t = T-h+1, \ldots, T-1. \end{cases} \tag{4.2}$$

From (4.1) it can be observed that $\hat{\beta}_h^{trf}$ is computed from the original non-overlapping one period returns. Notice that the transformed estimator in (4.1) can also be obtained from a regression of $\bar{y}_{t+1}$ on $\tilde{\bar{x}}_{t+h-1}^{trf,(h)}$, where

$$\tilde{\bar{x}}_t^{trf,(h)} := \left(\sum_{t=1}^{T-1} \left(\bar{x}_t^{trf,(h)} \right)^2 \right)^{-1} \left(\sum_{t=1}^{T-h} \bar{x}_t^2 \right) \bar{x}_t^{trf,(h)}.$$

Interestingly, it can be shown that the OLS slope estimator from the reverse regression (3.5), $\hat{\beta}_h^{rev}$ say, and $\hat{\beta}_h^{trf}$ are linearly related; specifically,

$$\hat{\beta}_h^{rev} = \frac{\sum_{t=1}^{T-h} \bar{x}_t^2}{\sum_{t=1}^{T-h} (\bar{x}_{t+h-1}^{(h)})^2} \hat{\beta}_h^{trf} + \frac{\sum_{k=1}^{h-1} [(\sum_{t=T-h+1}^{T-k} \bar{x}_i)\bar{y}_{T-k+1} - (\sum_{i=1}^{k} \bar{x}_i)\bar{y}_{k+1}]}{\sum_{t=1}^{T-h} (\bar{x}_{t+h-1}^{(h)})^2}$$

which suggests that when h is small the performance of predictability statistics from the reversed regression and transformed regression should be very similar, but as h increases their performance will likely differ.

If we knew that the predictor, x_t, was weakly persistent then we could base tests on the OLS estimate from the transformed regression discussed above. However, as with the tests of Phillips and Lee (2013) from Section 3.3, we want to allow for strongly persistent predictors. We will therefore apply the IVX framework of Kostakis et al. (2015) to the transformed regression. To that end, recall the IVX instrument z_t defined in (3.6). The transformed regression based IVX estimator is then obtained by regressing $\bar{y}_{t+1}$ on $\tilde{z}_t^{trf,(h)}$, where

$$\tilde{z}_t^{trf,(h)} := \frac{\left(\sum_{t=1}^{T-h} z_t \bar{x}_t \right) z_t^{trf,(h)}}{\sum_{t=1}^{T-1} \left(z_t^{trf,(h)} \right)^2} \tag{4.3}$$

with

$$z_t^{trf,(h)} := \begin{cases} \sum_{i=1}^{t} z_i & \text{for } t = 1, \ldots, h-1 \\ z_t^{(h)} := \sum_{i=1}^{h} z_{t-h+i} & \text{for } h \le t \le T-h \\ \sum_{i=t-h+1}^{T-h} z_i & \text{for } t = T-h+1, \ldots, T-1. \end{cases} \tag{4.4}$$

Hence, we obtain the transformed regression IVX estimator,

$$\hat{\beta}_{h,ivx}^{trf} = \frac{\sum_{t=1}^{T-1} z_t^{trf,(h)} \bar{y}_{t+1}}{\sum_{t=1}^{T-h} z_t \bar{x}_t} \approx \beta_h + \frac{\sum_{t=1}^{T-1} z_t^{trf,(h)} \bar{u}_{t+1}}{\sum_{t=1}^{T-h} z_t \bar{x}_t} \tag{4.5}$$

from which it is seen that the IVX estimate can provide the basis for inference on β_h. In particular, a test for the null hypothesis, $H_0 : \beta_h = 0$, against one or two-sided alternatives, can be obtained using a conventional IVX regression-based t-ratio of the form

$$t_{h,ivx}^{trf} := \frac{\hat{\beta}_{h,ivx}^{trf}}{s.e.\left(\hat{\beta}_{h,ivx}^{trf} \right)}. \tag{4.6}$$

In the context of (4.6), in view of Assumption 4 which allows for both conditional and unconditional heteroskedasticity in the innovations, we implement our IVX-based tests with conventional White heteroskedasticity-robust standard errors; that is,

$$s.e.\left(\hat{\beta}_{h,ivx}^{trf} \right) := \left[\left(\sum_{t=1}^{T-h} z_t \bar{x}_t \right)^{-1} \sum_{t=1}^{T-1} \left(z_t^{trf,(h)} \ddot{u}_{t+1} \right)^2 \left(\sum_{t=1}^{T-h} z_t \bar{x}_t \right)^{-1} \right]^{1/2} \tag{4.7}$$

where $\ddot{u}_{t+1} := \bar{y}_{t+1} - \hat{\beta}_{h,ivx}^{trf} \tilde{z}_t^{trf,(h)}$ are the residuals from the IVX estimation of the transformed regression. When testing the null hypothesis of no predictability, one may alternatively compute the residuals under the null; that is, $\ddot{u}_{t+1} := \bar{y}_{t+1}$.

4.2. Residual augmented tests

Recall the augmented regression in (3.1) where the addition of the infeasible regressor $\varpi_{t+h}^{(h)}$ serves to remove the endogeneity bias present in standard OLS estimation of (2.8). A feasible version of this augmented regression can be implemented if we can find a suitable (residual-based) estimate of $\varpi_{t+h}^{(h)}$. Such an approach is closely related to the concept of fully-modified methods developed by Phillips and Hansen (1990) for estimating equations involving $I(1)$ variables. Indeed, in the context of the short-horizon predictive regression in (2.1), Campbell and Yogo (2006a), use results from Phillips (1991) showing that (error) augmentation of the predictive regression has the potential to deliver efficient inference, in cases where the autoregressive roots are known and the errors are Gaussian, to motivate an infeasible augmented short-horizon predictive regression model (resorting to a Bonferroni-based approach to deal with the estimation error in the near-integrated case). Hjalmarsson (2007) clarifies the relationship between the approach adopted in Campbell and Yogo (2006a) and fully modified estimation. Based on these precedents, feasible implementation of the augmented long-horizon regression seems worth exploring.

At first sight, one might think it is possible to implement a feasible version of (3.1) that can be estimated by OLS simply by replacing the regressor $\varpi_{t+h}^{(h)}$ with an estimate of that quantity constructed from the OLS residuals, $\hat{\varpi}_t$ say, obtained from fitting an $AR(p)$ model to x_t (see (4.8)). However, this will not work. To illustrate why, consider the feasible estimator $\hat{\beta}_h^F := \left(\sum_{t=p}^{T-h} \bar{x}_t^2\right)^{-1} \sum_{t=p}^{T-h} \hat{\bar{y}}_{t+h}^{(h)} \bar{x}_t$, where $\hat{\bar{y}}_{t+h}^{(h)} := \bar{y}_{t+h}^{(h)} - \hat{\gamma}\hat{\varpi}_{t+h}^{(h)}$ and $\hat{\gamma}$ is a consistent estimator of γ, for example the fitted coefficient from an OLS regression of $\bar{y}_t$ on $\hat{\varpi}_t$ when testing the null hypothesis that $\beta_h = 0$.[8] In the simplest possible case where no short-run dynamics are present in the predictor process, it then follows that,

$$\hat{\beta}_h^F = \hat{\beta}_h^I + \gamma\left(\hat{\rho} - \rho\right) \frac{\sum_{t=1}^{T-h} \bar{x}_t \bar{x}_{t+h-1}^{(h)}}{\sum_{t=1}^{T-h} \bar{x}_t^2} + o_p(1)$$

where $\hat{\beta}_h^I$ is the infeasible estimate of β_h from (3.1). This shows that the feasible estimate features an additional term relative to the infeasible estimator, $\hat{\beta}_h^I$, which depends on the estimation error associated with the predictor's autoregressive parameter, $(\hat{\rho} - \rho)$, weighted by $\gamma\left(\sum_{t=1}^{T-h} \bar{x}_t^2\right)^{-1} \sum_{t=1}^{T-h} \bar{x}_t \bar{x}_{t+h-1}^{(h)}$. This term can be shown to be of the same order of magnitude as $\hat{\beta}_h^I$ (see e.g. Cai and Wang, 2014, for the short-horizon case) which renders the limiting null distribution of $\hat{\beta}_h^F$ non-pivotal. In fact, if computing the feasible estimator for $h = 1$ by augmenting the predictive regression with the OLS autoregression residuals $\hat{\varpi}_{t+1}$, it can be shown that $\hat{\varpi}_{t+1}$ are exact orthogonal to the regressor, x_t, and so this version of the feasible estimator will be numerically identical to the standard OLS estimator in the short-horizon case.

In the context of short-horizon predictability testing, Demetrescu and Rodrigues (2022) demonstrate that the problem with implementing a feasible version of (3.1), discussed above, does not arise if we estimate the residual augmented regression by IVX. Following their approach, we can apply residual augmentation to the transformed IVX estimate discussed in Section 4.1 by regressing $\bar{y}_{t+1} - \hat{\gamma}\hat{\varpi}_{t+1}$, rather than $\bar{y}_{t+1}$, on $\tilde{z}_t^{trf,(h)}$, where $\tilde{z}_t^{trf,(h)}$ is as defined in (4.3) and the residuals $\hat{\varpi}_{t+1}$ are computed from an estimated autoregressive model of order p for the predictor x_t, *viz.*,

$$\hat{\varpi}_{t+1} := \bar{x}_{t+1} - \sum_{k=1}^{p} \hat{\phi}_k \bar{x}_{t+1-k} = \varpi_t - \sum_{k=1}^{p} \left(\hat{\phi}_k - \phi_k\right) \bar{x}_{t-k} \tag{4.8}$$

for $t = p, \ldots, T-1$, where $\hat{\phi}_k$, $k = 1, \ldots, p$ are the OLS autoregressive parameter estimates. The dependent variable, $\bar{y}_{t+1} - \hat{\gamma}\hat{\varpi}_{t+1}$, is simply the OLS residual from the regression of $\bar{y}_{t+1}$ on $\hat{\varpi}_{t+1}$. In practice the lag augmentation order, p, in (4.8) can be selected using a standard information criterion, setting the minimum possible lag length allowed to be one. We denote the resulting residual-augmented transformed regression IVX estimator by $\hat{\beta}_{h,ivx}^{trf,res}$.

The viability of this approach in the IVX framework stems from the fact that the additional term attributable to OLS estimation error in the feasible estimation, discussed above, is asymptotically negligible in the IVX context in the case where the predictor is strongly or moderately persistent. To see why, consider the computational form for $\hat{\beta}_{h,ivx}^{trf,res}$,

$$\hat{\beta}_{h,ivx}^{trf,res} := \frac{\sum_{t=p}^{T-1} z_t^{trf,(h)} (\bar{y}_{t+1} - \hat{\gamma}\hat{\varpi}_{t+1})}{\sum_{t=1}^{T-h} z_t \bar{x}_t} \tag{4.9}$$

which can be written equivalently as

$$\hat{\beta}_{h,ivx}^{trf,res} = \frac{\sum_{t=p}^{T-1} z_t^{trf,(h)} (\beta_1 \bar{x}_t + \bar{u}_{t+1} - \hat{\gamma}\hat{\varpi}_{t+1})}{\sum_{t=1}^{T-h} z_t \bar{x}_t}.$$

Using results from the proofs of Theorems 4.1, 4.3 and 4.5 in the Supplementary Appendix, it can be established straightforwardly that

$$\hat{\beta}_{h,ivx}^{trf,res} = \beta_h + \frac{\sum_{t=p}^{T-1} z_t^{trf,(h)} \bar{\varepsilon}_{t+1}}{\sum_{t=1}^{T-h} z_t \bar{x}_t} + \gamma \sum_{k=1}^{p} (\hat{\phi}_k - \phi_k) \frac{\sum_{t=p}^{T-1} z_t^{trf,(h)} \bar{x}_{t-k}}{\sum_{t=1}^{T-h} z_t \bar{x}_t} + o_p(1).$$

8 If one is testing a null hypothesis other than $\beta_h = 0$, then $\hat{\gamma}$ is correspondingly obtained from the OLS regression of $\hat{u}_t$ (rather than $\bar{y}_t$) on $\hat{\varpi}_t$.

As demonstrated in the formal derivations in the Supplementary Appendix, the usual OLS autoregressive convergence rates on $\hat{\phi}_k$ suffice for the OLS estimation effect to be negligible under strong or moderate persistence. In the short-horizon case, Demetrescu and Rodrigues (2022) show, however, that the variance of $\hat{\beta}^{trf,res}_{h,ivx}$ will be affected by residual augmentation under weak persistence. For this reason they recommend computing the standard errors corresponding to the weak persistence case, and prove that the correction term this entails has an asymptotically negligible effect on the standard errors under strong persistence, such that one may conveniently use the standard errors developed for weak persistence, irrespective of whether the predictor exhibits weak or strong persistence. We will adopt the same approach here in the long-horizon context.

Based on the foregoing arguments, our proposed long-horizon IVX augmented statistic to test the null hypothesis $H_0 : \beta_h = 0$ is then given by,

$$t^{trf,res}_{h,ivx} := \frac{\hat{\beta}^{trf,res}_{h,ivx}}{s.e.\left(\hat{\beta}^{trf,res}_{h,ivx}\right)} \tag{4.10}$$

where

$$s.e.\left(\hat{\beta}^{trf,res}_{h,ivx}\right) := \left(\mathcal{H}_{zx}\right)^{-1}\left[\mathcal{H}_{z^{trf,(h)}\hat{\hat{\varepsilon}}z^{trf,(h)}\hat{\hat{\varepsilon}}} + \hat{\gamma}^2\hat{Q}^{trf,(h)}_T\right]^{1/2}$$

with $\mathcal{H}_{zx} := \left(\sum_{t=1}^{T-h} z_t\bar{x}_t\right)$; $\mathcal{H}_{z^{trf,(h)}\hat{\hat{\varepsilon}}z^{trf,(h)}\hat{\hat{\varepsilon}}} := \left(\sum_{t=p}^{T-1}\left(z^{trf,(h)}_t\right)^2\hat{\hat{\varepsilon}}^2_{t+1}\right)$; and

$$\hat{Q}^{trf,(h)}_T := \mathcal{H}'_{z^{trf,(h)}\bar{\mathbf{x}}}\mathcal{H}^{-1}_{\bar{\mathbf{x}}\bar{\mathbf{x}}}\mathcal{H}_{\bar{\mathbf{x}}\bar{\mathbf{x}}v}\mathcal{H}^{-1}_{\bar{\mathbf{x}}\bar{\mathbf{x}}}\mathcal{H}_{z^{trf,(h)}\bar{\mathbf{x}}};$$

defining $\bar{\mathbf{x}}_t := \left(\bar{x}_t, \ldots, \bar{x}_{t-p+1}\right)'$, we have further $\mathcal{H}_{z^{trf,(h)}\bar{\mathbf{x}}} := \left(\sum_{t=p}^{T-1} z^{trf,(h)}_t\bar{x}_t, \ldots, \sum_{t=p}^{T-1} z^{trf,(h)}_t\bar{x}_{t-p+1}\right)'$, $\mathcal{H}_{\bar{\mathbf{x}}\bar{\mathbf{x}}} := \sum_{t=p}^{T-1}\bar{\mathbf{x}}_t\bar{\mathbf{x}}'_t$; and $\mathcal{H}_{\bar{\mathbf{x}}\bar{\mathbf{x}}v} := \sum_{t=p}^{T-1}\bar{\mathbf{x}}_t\bar{\mathbf{x}}'_t\hat{\varpi}^2_{t+1}$, with $\hat{\hat{\varepsilon}}_{t+1}$ the residuals from regressing y_{t+1} on $\hat{\varpi}_{t+1}$ and an intercept (i.e. computed under the null hypothesis; for null hypotheses other than $\beta_h = 0$, the regression used to obtain $\hat{\hat{\varepsilon}}_{t+1}$ should also contain x_t). These (heteroskedasticity-robust) standard errors are designed to automatically take the estimation variability of $\hat{\phi}_k$ into account whenever needed, such that the standard errors are asymptotically correct without having to specify whether x_t is weakly or strongly persistent; cf. Demetrescu and Rodrigues (2022). As we will subsequently demonstrate, this nice property also extends to the case of moderately persistent predictors, not considered by Demetrescu and Rodrigues (2022).

4.3. Asymptotic theory

In this section we analyse the large sample distributions of the estimators and test statistics proposed in Sections 4.1 and 4.2, when the data generating process is as in (2.1)–(2.2) under Assumptions 1–4. In this setting, it is observed that the partial sums of the innovations v_t and ε_t display joint weak convergence to time-transformed Brownian motions (see Lemma S.1 in the Supplementary Appendix); precisely,

$$\frac{1}{\sqrt{T}}\sum_{t=1}^{\lfloor sT\rfloor}\begin{pmatrix}\varepsilon_t\\ v_t\end{pmatrix} \Rightarrow \begin{pmatrix}\int_0^s\sigma_\varepsilon(r)\,\mathrm{d}W_\varepsilon(r)\\ \int_0^s\sigma_\varpi(r)\,\mathrm{d}W_\varpi(r)\end{pmatrix}$$

were "⇒" denotes weak convergence on the space of càdlàg real functions on $[0, 1]^k$ equipped with the Skorokhod topology, and where W_ε and W_ϖ are independent standard Wiener processes. Moreover, under near integration (Assumption 2(i)), it also follows that the stochastic part of the suitably normalised regressor weakly converges to an Ornstein–Uhlenbeck-type process; that is,

$$T^{-1/2}\xi_{\lfloor sT\rfloor} \Rightarrow \omega\int_0^s e^{-c(s-r)}\sigma_\varpi\mathrm{d}W_\varpi(r) =: \omega J_{c,\sigma}(s)\,. \tag{4.11}$$

In Theorems 4.1 and 4.2, respectively, we first establish the limiting distributions of $\hat{\beta}^{trf,res}_{h,ivx}$ and $\hat{\beta}^{trf}_{h,ivx}$ and their associated standard errors in the case where x_t is strongly persistent.

Theorem 4.1. *Under Assumptions 1, 2(i), 3 and 4 with* $\epsilon < \min\{1-\eta; \eta/2\}$ *and as* $h, T \to \infty$ *such that* $h/(\min\left\{T^{3\eta/2-1/2}; T^{2\eta-1}\right\}) \to 0$*, we have that*

$$\frac{T^{\eta/2+1/2}}{h}\left(\hat{\beta}^{trf,res}_{h,ivx} - \beta^{(i)}_h\right) \Rightarrow \mathcal{MN}\left(0, \frac{a\int_0^1\sigma^2_\varpi(s)\,\sigma^2_\varepsilon(s)\,\mathrm{d}s}{2\omega^2\left(J_{c,\sigma}(1)\bar{J}_{c,\sigma}(1) - \int_0^1 J_{c,\sigma}(s)\mathrm{d}J_{c,\sigma}(s)\right)^2}\right)$$

where a and η are the tuning parameters for the IVX instrument in (3.6) and $\mathcal{MN}$ denotes a mixed normal distribution, with ω defined in Assumption 3, $J_{c,\sigma}(s)$ defined in (4.11) and $\bar{J}_{c,\sigma}(s) := J_{c,\sigma}(s) - \int_0^1 J_{c,\sigma}(s)\mathrm{d}s$, and

$$\frac{T^{\eta/2+1/2}}{h} s.e.\left(\hat{\beta}_{h,ivx}^{trf,res}\right) \Rightarrow \frac{\sqrt{a\int_0^1 \sigma_{\varpi}^2(s)\,\sigma_{\varepsilon}^2(s)\,\mathrm{d}s}}{\sqrt{2\omega^2}\left(J_{c,\sigma}(1)\bar{J}_{c,\sigma}(1) - \int_0^1 J_{c,\sigma}(s)\mathrm{d}J_{c,\sigma}(s)\right)}.$$

Remark 8. The limiting results given in Theorem 4.1 are similar to those given in Theorem 3.2 of Demetrescu and Rodrigues (2022) for the short-horizon, $h = 1$, case, but hold under considerably weaker assumptions on the innovations than are allowed for in Demetrescu and Rodrigues (2022); here, we allow for conditional heteroskedasticity while Demetrescu and Rodrigues (2022) only consider heterogeneous independent error sequences. Compared to the short-horizon case, the results in Theorem 4.1 need to take account of the implied aggregation of various quantities which, although individually asymptotically negligible quantities, arise over h periods. Given that we allow for $h \to \infty$, this entails the need to place additional conditions on the persistence allowed for in the IVX instrument, as controlled by η. In particular, Theorem 4.1 requires that $\eta > 1/3$, in addition to conditions relating the persistence of z_t to the strength of the GARCH effects present in the DGP as controlled by ϵ. The choice of $\eta = 0.95$ for the IVX tuning parameter recommended by Kostakis et al. (2015) is permitted under our rate restrictions, as long as the serial dependence in the conditional variances is not too high. It is important to note that the results require that $h \to \infty$, albeit at a minimal rate which is very mild when η is close to unity. Nevertheless, based on the results in Demetrescu and Rodrigues (2022) who consider the case $h = 1$, it should be possible to also obtain corresponding results for fixed h with some additional technical effort. We do not do so here in the interests of brevity. ♢

Theorem 4.2. *Under the conditions of Theorem 4.1, we have that*

$$\frac{T^{\eta/2+1/2}}{h}\left(\hat{\beta}_{h,ivx}^{trf} - \beta_h^{(i)}\right) \Rightarrow \mathcal{MN}\left(0, \frac{a\int_0^1 \sigma_{\varpi}^2(s)\left(\sigma_{\varepsilon}^2(s) + \gamma^2\sigma_{\varpi}^2(s)\right)\mathrm{d}s}{2\omega^2\left(J_{c,\sigma}(1)\bar{J}_{c,\sigma}(1) - \int_0^1 J_{c,\sigma}(s)\mathrm{d}J_{c,\sigma}(s)\right)^2}\right)$$

and

$$\frac{T^{\eta/2+1/2}}{h} s.e.\left(\hat{\beta}_{h,ivx}^{trf}\right) \Rightarrow \frac{\sqrt{a\int_0^1 \sigma_{\varpi}^2(s)\left(\sigma_{\varepsilon}^2(s) + \gamma^2\sigma_{\varpi}^2(s)\right)\mathrm{d}s}}{\sqrt{2\omega^2}\left(J_{c,\sigma}(1)\bar{J}_{c,\sigma}(1) - \int_0^1 J_{c,\sigma}(s)\mathrm{d}J_{c,\sigma}(s)\right)}.$$

Remark 9. A comparison of the results in Theorems 4.1 and 4.2 highlights the asymptotic efficiency gains which arise from residual augmentation. This can be seen by noting that the asymptotic variance (conditional on $J_{c,\sigma}$) of $\hat{\beta}_{h,ivx}^{trf}$ is strictly larger than that of the residual augmented estimator, $\hat{\beta}_{h,ivx}^{trf,res}$, whenever $\gamma \neq 0$. These asymptotic efficiency gains are reflected by the finite-sample power behaviour of the residual-augmented tests; see Section 6.3. Moreover, the simulation results also indicate an improved size behaviour, which, building on the findings of Demetrescu and Rodrigues (2022) for the case $h = 1$, can be traced back to reductions in the finite-sample bias of the IVX estimator. ♢

In Theorems 4.3 and 4.4 we next establish the limiting distributions of $\hat{\beta}_{h,ivx}^{trf,res}$ and $\hat{\beta}_{h,ivx}^{trf}$ and their associated standard errors in the case where x_t is weakly persistent.

Theorem 4.3. *Under Assumptions 1, 2(ii), 3 and 4, we have as $h, T \to \infty$ such that $h^3/T \to 0$,*

$$\sqrt{\frac{T}{h}}\left(\hat{\beta}_{h,ivx}^{trf,res} - \beta_h^{(ii)}\right) \xrightarrow{d} \mathcal{N}\left(0, \frac{\frac{\omega^2}{(1-\rho)^2}\int_0^1 \sigma_{\varpi}^2(s)\,\sigma_{\varepsilon}^2(s)\,\mathrm{d}s}{\left(\theta_0\int_0^1 \sigma_{\varpi}^2(s)\,\mathrm{d}s\right)^2}\right)$$

where $\theta_0 := \sum_{k\geq 0} b_k^2$ is as defined in Assumption 3, and

$$\sqrt{\frac{T}{h}} s.e.\left(\hat{\beta}_{h,ivx}^{trf,res}\right) \xrightarrow{p} \frac{\omega\sqrt{\int_0^1 \sigma_{\varpi}^2(s)\,\sigma_{\varepsilon}^2(s)\,\mathrm{d}s}}{(1-\rho)\,\theta_0\int_0^1 \sigma_{\varpi}^2(s)\,\mathrm{d}s}.$$

Theorem 4.4. *Under the conditions of Theorem 4.3, we have that*

$$\sqrt{\frac{T}{h}}\left(\hat{\beta}_{h,ivx}^{trf} - \beta_h^{(ii)}\right) \xrightarrow{d} \mathcal{N}\left(0, \frac{\frac{\omega^2}{(1-\rho)^2}\int_0^1 \sigma_{\varpi}^2(s)\left(\sigma_{\varepsilon}^2(s) + \gamma^2\sigma_{\varpi}^2(s)\right)\mathrm{d}s}{\left(\theta_0\int_0^1 \sigma_{\varpi}^2(s)\,\mathrm{d}s\right)^2}\right)$$

and

$$\sqrt{\frac{T}{h}} s.e.\left(\hat{\beta}^{trf}_{h,ivx}\right) \xrightarrow{p} \frac{\omega\sqrt{\int_0^1 \sigma^2_{\varpi}(s)\left(\sigma^2_{\varepsilon}(s)+\gamma^2\sigma^2_{\varpi}(s)\right)\mathrm{d}s}}{(1-\rho)\,\theta_0\int_0^1 \sigma^2_{\varpi}(s)\,\mathrm{d}s}.$$

Finally, we consider the intermediary case of moderate (or mild) persistence.

Theorem 4.5. *Under Assumptions 1, 2(iii), 3 and 4, with $\epsilon < \min\{1-\eta;\eta/2;1-\kappa;\kappa/2\}$ and as $h, T \to \infty$ such that $h/(\min\{T^{3\eta/2-1/2}; T^{2\eta-1}; T^{3\kappa/2-1/2}; T^{2\kappa-1}\}) \to 0$, we have that*

$$\frac{T^{\min\{\eta,\kappa\}/2+1/2}}{h}\left(\hat{\beta}^{trf,res}_{h,ivx}-\beta_h^{(iii)}\right) \xrightarrow{d} \mathcal{N}\left(0, \frac{2g(a,c)}{\omega^2}\frac{\int_0^1 \sigma^2_{\varpi}(s)\,\sigma^2_{\varepsilon}(s)\,\mathrm{d}s}{\left(\int_0^1 \sigma^2_{\varpi}(s)\,\mathrm{d}s\right)^2}\right),$$

where $g(a,c)=a$ if $\eta<\kappa$ and $g(a,c)=c$ if $\kappa<\eta$, and

$$\frac{T^{\min\{\eta,\kappa\}/2+1/2}}{h} s.e.\left(\hat{\beta}^{trf,res}_{h,ivx}\right) \xrightarrow{p} \frac{\sqrt{2g(a,c)\int_0^1 \sigma^2_{\varpi}(s)\,\sigma^2_{\varepsilon}(s)\,\mathrm{d}s}}{\omega\int_0^1 \sigma^2_{\varpi}(s)\,\mathrm{d}s}.$$

Theorem 4.6. *Under the conditions of Theorem 4.5, we have that*

$$\frac{T^{\min\{\eta,\kappa\}/2+1/2}}{h}\left(\hat{\beta}^{trf,res}_{h,ivx}-\beta_h^{(iii)}\right) \xrightarrow{d} \mathcal{N}\left(0, \frac{2g(a,c)}{\omega^2}\frac{\int_0^1 \sigma^2_{\varpi}(s)\left(\sigma^2_{\varepsilon}(s)+\gamma^2\sigma^2_{\varpi}(s)\right)\mathrm{d}s}{\left(\int_0^1 \sigma^2_{\varpi}(s)\,\mathrm{d}s\right)^2}\right),$$

and

$$\frac{T^{\min\{\eta,\kappa\}/2+1/2}}{h} s.e.\left(\hat{\beta}^{trf,res}_{h,ivx}\right) \xrightarrow{p} \frac{\sqrt{2g(a,c)\int_0^1 \sigma^2_{\varpi}(s)\left(\sigma^2_{\varepsilon}(s)+\gamma^2\sigma^2_{\varpi}(s)\right)\mathrm{d}s}}{\omega\int_0^1 \sigma^2_{\varpi}(s)\,\mathrm{d}s}.$$

Remark 10. An implication of Theorems 4.1–4.6 is that the convergence rates of both $\hat{\beta}^{trf,res}_{h,ivx}$ and $\hat{\beta}^{trf}_{h,ivx}$ decrease with the forecast horizon, h. In the strongly and moderately persistent cases, however, $\beta_h=\beta_h^{(i)}=\beta_h^{(iii)}$ increases (approximately) linearly in h which offsets the decreased convergence rate of the estimators. In contrast, in the weakly persistent case, $\beta_h=\beta_h^{(ii)}$ remains bounded leading to power losses as the horizon h increases. We will also see this difference in a comparison of the asymptotic lower power functions of the associated t-statistics in Theorems 4.7 (strongly persistent predictor), 4.8 (weakly persistent predictor) and 4.9 (moderately persistent predictor) which follow next. The Monte Carlo simulation results reported in Section 6.3 clearly bear out this prediction from the asymptotic theory. $\diamond$

Theorem 4.7. *Under the conditions of Theorem 4.1 and local alternatives of the form $\beta_1=bT^{-\eta/2-1/2}$, we have that*

$$t^{trf,res}_{h,ivx} \xrightarrow{d} \mathcal{MN}\left(b\frac{\omega\sqrt{\frac{2}{a}}\left(J_{c,\sigma}(1)\bar{J}_{c,\sigma}(1)-\int_0^1 J_{c,\sigma}(s)\mathrm{d}J_{c,\sigma}(s)\right)}{\sqrt{\int_0^1 \sigma^2_{\varpi}(s)\,\sigma^2_{\varepsilon}(s)\,\mathrm{d}s}}, 1\right)$$

and

$$t^{trf}_{h,ivx} \xrightarrow{d} \mathcal{MN}\left(b\frac{\omega\sqrt{\frac{2}{a}}\left(J_{c,\sigma}(1)\bar{J}_{c,\sigma}(1)-\int_0^1 J_{c,\sigma}(s)\mathrm{d}J_{c,\sigma}(s)\right)}{\sqrt{\int_0^1 \sigma^2_{\varpi}(s)\left(\sigma^2_{\varepsilon}(s)+\gamma^2\sigma^2_{\varpi}(s)\right)\mathrm{d}s}}, 1\right).$$

Theorem 4.8. *Under the conditions of Theorem 4.3 and local alternatives of the form $\beta_1=bh^{1/2}T^{-1/2}$, we have that*

$$t^{trf,res}_{h,ivx} \xrightarrow{d} \mathcal{N}\left(b\frac{(1-\rho)\,\theta_0\int_0^1 \sigma^2_{\varpi}(s)\,\mathrm{d}s}{\omega\sqrt{\int_0^1 \sigma^2_{\varpi}(s)\,\sigma^2_{\varepsilon}\mathrm{d}s}}; 1\right).$$

and

$$t^{trf}_{h,ivx} \xrightarrow{d} \mathcal{N}\left(b\frac{(1-\rho)\,\theta_0\int_0^1 \sigma^2_{\varpi}(s)\,\mathrm{d}s}{\omega\sqrt{\int_0^1 \sigma^2_{\varpi}(s)\left(\sigma^2_{\varepsilon}(s)+\gamma^2\sigma^2_{\varpi}(s)\right)\mathrm{d}s}}; 1\right).$$

Theorem 4.9. *Under the conditions of Theorem 4.5 and local alternatives of the form* $\beta_1 = bT^{-\min\{\eta;\kappa\}/2-1/2}$, *we have that*

$$t_{h,ivx}^{trf,res} \xrightarrow{d} \mathcal{N}\left(b\frac{\omega\int_0^1 \sigma_{\varpi}^2(s)\,ds}{\sqrt{2g(a,c)\int_0^1 \sigma_{\varpi}^2(s)\,\sigma_{\varepsilon}^2(s)\,ds}};1\right).$$

and

$$t_{h,ivx}^{trf} \xrightarrow{d} \mathcal{N}\left(b\frac{\omega\int_0^1 \sigma_{\varpi}^2(s)\,ds}{\sqrt{2g(a,c)\int_0^1 \sigma_{\varpi}^2(s)\left(\sigma_{\varepsilon}^2(s)+\gamma^2\sigma_{\varpi}^2(s)\right)ds}};1\right).$$

Using the results given above in Theorems 4.7–4.9 we are now in a position to establish the limiting null distributions of our proposed transformed regression long-horizon predictability test statistics, $t_{h,ivx}^{trf}$ from Section 4.1 and $t_{h,ivx}^{trf,res}$ from Section 4.2.

Corollary 1. *Under the null hypothesis of no predictability* $H_0 : \beta_h = 0$, *we have that under Assumptions 1–4 with* $\epsilon < \min\{1-\eta;\eta/2\}$ *and* $h/\min\left\{T^{3\eta/2-1/2};T^{2\eta-1};T^{3\kappa/2-1/2};T^{2\kappa-1};T^{\eta/3}\right\} \to 0$ *as* $h,T\to\infty$,

$$t_{h,ivx}^{trf,res} \xrightarrow{d} \mathcal{N}(0,1) \quad \text{and} \quad t_{h,ivx}^{trf} \xrightarrow{d} \mathcal{N}(0,1).$$

Corollary 1 demonstrates the key result for practical implementation of our proposed long-horizon predictability tests, that both $t_{h,ivx}^{trf}$ and $t_{h,ivx}^{trf,res}$ admit standard normal limiting null distributions regardless of whether the predictor is weakly, strongly, or moderately persistent. These results hold under the very general forms of conditional and/or unconditional heteroskedasticity permitted under Assumption 4.

5. Multiple predictors

In empirical work one might wish to consider predictive regression models with several possible predictors. This can help avoid the problem of spurious predictive regression effects in the case where relevant strongly persistent predictors are omitted from the estimated predictive regression; cf. Georgiev et al. (2018) and Andersen and Varneskov (2021b). We now briefly detail how the long-horizon predictability tests developed in Section 4 can be implemented with multiple predictors.

To that end consider replacing (2.8) by its multivariate counterpart

$$y_{t+h}^{(h)} = \alpha_h + \boldsymbol{\beta}_h'\boldsymbol{x}_t^{\dagger} + w_{t+h}^{(h)} \tag{5.1}$$

where $\boldsymbol{x}_t^{\dagger} := (x_{t1},\ldots,x_{tK})'$ follows a K-dimensional vector autoregressive data generating process of order p, VAR(p); that is,

$$\boldsymbol{x}_t^{\dagger} = \boldsymbol{\mu}_x + \mathbf{R}\boldsymbol{x}_{t-1}^{\dagger} + \boldsymbol{v}_t, \text{ and } \boldsymbol{v}_t = \sum_{j=1}^{p-1}\boldsymbol{\Gamma}_j\boldsymbol{v}_{t-j} + \boldsymbol{\varpi}_t \tag{5.2}$$

which is either stable or (near) integrated as before depending on the properties of the (diagonal) autoregressive coefficient matrix $\mathbf{R}$. The process $\boldsymbol{v}_t$ is assumed to follow a stable VAR($p-1$) process.

As with (2.8), the regression coefficients and error term in (5.1) can be related back to those in the corresponding short-horizon regression, $y_{t+1} = \alpha_1 + \boldsymbol{\beta}_1'\boldsymbol{x}_t^{\dagger} + u_{t+1}$, e.g. via the relationships, $\alpha_h := h\alpha_1 + \boldsymbol{\beta}_1'I_{h\geq 2}\sum_{j=1}^{h-1}\sum_{i=1}^{j}\mathbf{R}^{i-1}\boldsymbol{\mu}_x(\mathbf{I}-\mathbf{R})$, $\boldsymbol{\beta}_h' := \boldsymbol{\beta}_1'\sum_{j=0}^{h-1}\mathbf{R}^j$ and $w_{t+h}^{(h)} := u_{t+h}^{(h)} + \boldsymbol{\beta}_1'I_{h\geq 2}\sum_{j=1}^{h-1}\sum_{i=1}^{h-j}\mathbf{R}^{i-1}\boldsymbol{v}_{t+j}$ for the strongly and moderately persistent cases. Again we allow for the possibility of endogeneity in all regressors through the non-zero coefficient vector $\boldsymbol{\gamma}$ in the decomposition

$$u_{t+1} := \boldsymbol{\gamma}'\boldsymbol{\varpi}_{t+1} + \varepsilon_{t+1}, \tag{5.3}$$

where the innovations $\boldsymbol{\varpi}_{t+1}$ and ε_{t+1} are heterogeneous MDs, obeying a multivariate version of Assumption 4.

To implement the transformed bias reduced IVX approach introduced in this paper in the multiple predictive regression case, we first compute the vector of residuals $\hat{\boldsymbol{\varpi}}_t$ from a vector autoregression model of order p of the demeaned predictors; that is, with $\bar{\boldsymbol{x}}_t^{\dagger} := (\bar{x}_{t1},\ldots,\bar{x}_{tK})'$,

$$\hat{\boldsymbol{\varpi}}_{t+1} := \bar{\boldsymbol{x}}_{t+1}^{\dagger} - \sum_{j=1}^{p}\hat{\boldsymbol{\Phi}}_j\bar{\boldsymbol{x}}_{t+1-j}^{\dagger}, \quad t = p,\ldots,T-1, \tag{5.4}$$

with $\hat{\boldsymbol{\Phi}}_j$, $j = 1,\ldots,p$, the OLS coefficient matrix estimates. Again, the lag augmentation order in (5.4) can be selected in practice by using a standard information criterion, setting the minimum possible lag length allowed to be one. The multiple predictor residual augmented IVX estimator vector is then defined as

$$\hat{\boldsymbol{\beta}}_{h,ivx}^{trf,res} := \left(\sum_{t=p}^{T-1} \tilde{\boldsymbol{z}}_t^{trf,(h)} \tilde{\boldsymbol{z}}_t^{trf,(h)\prime}\right)^{-1} \sum_{t=p}^{T-1} \tilde{\boldsymbol{z}}_t^{trf,(h)} \left(\bar{y}_{t+1} - \hat{\boldsymbol{\gamma}}' \hat{\boldsymbol{\varpi}}_{t+1}\right)$$

$$= \left(\sum_{t=1}^{T-h} \boldsymbol{z}_t \bar{\boldsymbol{x}}_t^{\dagger\prime}\right)^{-1} \sum_{t=p}^{T-1} \boldsymbol{z}_t^{trf,(h)} \left(\bar{y}_{t+1} - \hat{\boldsymbol{\gamma}}' \hat{\boldsymbol{\varpi}}_{t+1}\right). \tag{5.5}$$

where $\boldsymbol{z}_t$ is a $K \times 1$ vector of instruments with elements as defined in (3.6) for each predictor in $\boldsymbol{x}_t^{\dagger}$ and

$$\tilde{\boldsymbol{z}}_t^{trf,(h)} := \left(\sum_{t=p}^{T-1} \boldsymbol{z}_t^{trf,(h)} \boldsymbol{z}_t^{trf,(h)\prime}\right)^{-1} \left(\sum_{t=1}^{T-h} \boldsymbol{z}_t \bar{\boldsymbol{x}}_t^{\dagger\prime}\right) \boldsymbol{z}_t^{trf,(h)} \tag{5.6}$$

in which $\boldsymbol{z}_t^{trf,(h)}$ is a $K \times 1$ vector of instruments, whose elements are obtained by applying the definition in (4.4) to each element of $\boldsymbol{z}_t$.

For inference purposes we need to estimate the covariance matrix of $\hat{\boldsymbol{\beta}}_{h,ivx}^{trf,res}$. This can be done by using the familiar "sandwich" formula,

$$\widehat{\mathrm{Cov}\left(\hat{\boldsymbol{\beta}}_{h,ivx}^{trf,res}\right)} := \mathbf{B}_T^{-1} \mathbf{M}_T \left(\mathbf{B}_T^{-1}\right)' \tag{5.7}$$

where $\mathbf{B}_T := \sum_{t=1}^{T-h} \boldsymbol{z}_t \bar{\boldsymbol{x}}_t^{\dagger\prime}$ and

$$\mathbf{M}_T := \sum_{t=p}^{T-1} \boldsymbol{z}_t^{trf,h} \boldsymbol{z}_t^{trf,(h)\prime} \hat{\varepsilon}_{t+1}^2 + \left[\boldsymbol{\gamma}' \otimes \left(\tfrac{1}{T}\sum_{t=p}^{T-1} \boldsymbol{z}_t^{trf,h} \bar{\boldsymbol{\mathcal{X}}}_{t,K}'\right) \left(\sum_{t=p}^{T-1} \bar{\boldsymbol{\mathcal{X}}}_{t,K} \bar{\boldsymbol{\mathcal{X}}}_{t,K}'\right)^{-1}\right] \times$$
$$\times \left(\sum_{t=p}^{T-1} \hat{\boldsymbol{\varpi}}_t \hat{\boldsymbol{\varpi}}_t' \otimes \bar{\boldsymbol{\mathcal{X}}}_{t,K} \bar{\boldsymbol{\mathcal{X}}}_{t,K}'\right) \left[\hat{\boldsymbol{\gamma}} \otimes \left(\sum_{t=p}^{T-1} \bar{\boldsymbol{\mathcal{X}}}_{t,K} \bar{\boldsymbol{\mathcal{X}}}_{t,K}'\right)^{-1} \left(\tfrac{1}{T}\sum_{t=p}^{T-1} \bar{\boldsymbol{\mathcal{X}}}_{t,K} \boldsymbol{z}_t^{trf,(h)\prime}\right)\right]$$

where $\bar{\boldsymbol{\mathcal{X}}}_{t,K}$ is the vector formed from stacking the p lags of each of the K demeaned regressors; that is, $\bar{\boldsymbol{\mathcal{X}}}_{t,K} := \left(\bar{x}_{t,1}, \ldots, \bar{x}_{t,K},\ \bar{x}_{t-1,1}, \ldots, \bar{x}_{t-1,K}, \ldots, \bar{x}_{t-p+1,1}, \ldots, \bar{x}_{t-p+1,K}\right)'$.

The limiting distribution of $\hat{\boldsymbol{\beta}}_{h,ivx}^{trf,res}$ is (multivariate) normal in the case where the elements of $\boldsymbol{x}_t$ are either all weakly persistent or all moderately persistent, and mixed normal in the case where they are all strongly persistent; the proofs are straightforward multivariate extensions of the results from the single-regressor case given in Section 4.3 and are therefore omitted. An important consequence of these results is that the associated individual and joint significance tests on the elements of $\boldsymbol{\beta}_h$ have standard normal (if one linear restriction is being tested using a t-type ratio) and χ^2 (for multiple restrictions) limiting null distributions irrespective of whether the elements of $\boldsymbol{x}_t$ are weakly, moderately or strongly persistent, and regardless of any heterogeneity present in the DGP, provided the heteroskedasticity-robust covariance matrix estimator in (5.7) is used. Moreover, we conjecture, based on some preliminary examinations given in section S.3 in the Supplementary Appendix, that this result also holds in the case where the predictors have mixed degrees of persistence. Simulation results pertaining to the case of multiple predictors, including mixed persistence cases, are reported in section S.5 of the Supplementary Appendix.

6. Numerical results

6.1. Set-up

We report the results from a Monte Carlo study exploring the finite sample performance of the residual augmented transformed regression based long-horizon predictability test, $t_{h,ivx}^{trf,res}$, from Section 4.2. We will compare the finite sample performance of this test with the Bonferroni-based test, t_h^{Bonf}, of Hjalmarsson (2011) outlined in Section 3.1, the bias-corrected wild bootstrap implementation of the implied test, t_h^{Xu}, of Xu (2020) outlined in Section 3.2, and the reversed predictive regression based test, $t_{h,ivx}^{rev,PL}$, of Phillips and Lee (2013) outlined in Section 3.3. We also considered the nonaugmented transformed regression test, $t_{h,ivx}^{trf}$ defined in (4.6), we found that this did not perform as well as $t_{h,ivx}^{trf,res}$ (its performance was in fact very similar to that of $t_{h,ivx}^{rev,PL}$), and so we only report results for $t_{h,ivx}^{trf,res}$. Empirical size results are reported in Section 6.2 and empirical power properties in Section 6.3. A number of additional Monte Carlo results are presented in the Supplementary Appendix.

For all of the reported experiments, data are generated from (2.1)–(2.2). All of the tests considered are for the null hypothesis of no long-run predictability $H_0 : \beta_h = 0$ in (2.8). We will consider tests directed against both one-sided (left-tailed tests for $H_1 : \beta_h < 0$, and right-tailed tests for $H_1 : \beta_h > 0$), and two-sided alternatives ($H_1 : \beta_h \neq 0$). All tests are run at the 5% nominal (asymptotic) significance level. The simulations were preformed in MATLAB, version R2020a, using the Mersenne Twister random number generator function using 10 000 and 5000 Monte Carlo replications for the empirical size and empirical power simulations, respectively.

In implementing t_h^{Bonf}, we follow the steps outlined in Hjalmarsson (2011), however, we use the GLS detrended ADF approach as suggested in Campbell and Yogo (2006a) to compute the confidence interval for c instead of Chen and Deo (2009) as it gave better results. With the exception of the IVX instrument, z_t, all variables entering the estimated predictive regressions are demeaned. As discussed in Kostakis et al. (2015, p. 1514) the IVX instrument z_t, does not need to be demeaned because the slope estimator in the predictive regression is invariant to whether z_t is demeaned or not. For implementation of $t_{h,ivx}^{trf,res}$ in (4.9) we start by estimating an autoregressive model of order p, where p was chosen applying AIC over $p \in (1, \ldots, \lfloor 4(T/100)^{1/4} \rfloor)$. The resulting residuals, $\hat{\varpi}_{t+1}$ are then used to compute $\bar{y}_{t+1} - \hat{\gamma}\hat{\varpi}_{t+1}$, from a regression of $\bar{y}_{t+1}$ on $\hat{\varpi}_{t+1}$.

6.2. Empirical size

In this section we investigate the finite sample size properties of our proposed $t_{h,ivx}^{trf,res}$ test and contrast them with the results of the t_h^{Bonf}, t_h^{Xu} and $t_{h,ivx}^{rev,PL}$ tests.[9] To that end, we generate data from (2.1)–(2.2) with $\beta_1 = 0$. In generating the simulation data we set the intercepts, α_1 and μ_x, in (2.1)–(2.2), respectively, to zero without loss of generality. The autoregressive process for x_t was generated as in (2.2) with $\rho = 1+c/T$ for $c \in \{0, -5, -10, -20, -50\}$ and was initialised at $\xi_0 = 0$. Results are reported for samples of length $T = \{100, 250, 500\}$ and for forecast horizons $h = \{5, 10, 20, 50\}$; corresponding results for the short horizon case, $h = 1$, are reported in the Supplementary Appendix.[10]

We allow the innovations driving the predictor process in (2.2) to either be serially uncorrelated or to follow an AR(1) process; in particular we set $v_{t+1} = \psi v_t + \varpi_{t+1}$, and consider $\psi \in \{-0.5, 0, 0.5\}$. The innovation vector $(u_{t+1}, \varpi_{t+1})'$ in (2.1)–(2.2) is drawn from an i.i.d. bivariate Gaussian distribution[11] with mean zero and covariance matrix $\mathbf{\Sigma} := \begin{bmatrix} \sigma_u^2 & \phi\sigma_u\sigma_\varpi \\ \phi\sigma_u\sigma_\varpi & \sigma_\varpi^2 \end{bmatrix}$, where ϕ is as defined in Remark 7 and corresponds to the (time-invariant) correlation between the innovations u_{t+1} and v_{t+1}. For all of the simulation DGPs we will consider we set $\sigma_u^2 = \sigma_v^2$ so that it always holds that $\gamma = \phi$, where γ is as defined in Assumption 3; cf. Remark 7. We consider $\phi = \{-0.15, -0.50, -0.95\}$.[12] Tables 1–3 report results for $\psi = 0$, $\psi = 0.5$, and $\psi = -0.5$, respectively, when $\phi = -0.15$ and $\phi = -0.95$, setting $\sigma_u^2 = \sigma_v^2 = 1$ throughout. Results for $\psi = 0$, $\psi = 0.5$, and $\psi = -0.5$ when $\phi = -0.5$ are reported in the Supplementary Appendix in Tables S.1, S.2 and S.3, respectively.

The results in Tables 1–3 highlight the superiority of the IVX-based tests, $t_{h,ivx}^{trf,res}$ and $t_{h,ivx}^{rev,PL}$, over the non-IVX based t_h^{Bonf}, t_h^{Xu} tests in terms of controlling size across both strongly and weakly persistent predictors. Taking the case where $\psi = 0$ to illustrate, it is seen from the results in Table 1, which are for the case where v_{t+1} is serially uncorrelated, that the empirical rejection frequencies of the two-sided $t_{h,ivx}^{trf,res}$ and $t_{h,ivx}^{rev,PL}$ tests when $\phi = -0.15$, for $T = 100$ are in the range $[0.020, 0.055]$ and $[0.022, 0.054]$, respectively; for $T = 250$ the range is $[0.018, 0.047]$ and $[0.018, 0.053]$, respectively, and for $T = 500$ the range is $[0.017, 0.047]$ and $[0.018, 0.049]$, respectively, taken across all of the values of c considered. For $h = 50$ these two tests become slightly conservative when contrasting with the results for $h < 50$. Moreover, when $\phi = -0.95$, for $T = 100$ the empirical rejection frequencies of these tests are in the range $[0.017, 0.058]$ and $[0.044, 0.099]$, respectively; for $T = 250$ in $[0.022, 0.062]$ and $[0.039, 0.065]$, respectively, and for $T = 500$ in $[0.023, 0.058]$ and $[0.046, 0.061]$, respectively, again taken across all of the values of c considered (recall that for $T = 100$ $c = -50$ is not considered). For $\phi = -0.95$ we observe that $t_{h,ivx}^{rev,PL}$ displays some oversizedness for $T = 100$ but improves as the sample size increases.

A comparison of the results in Table 1 with those in Tables 2–3 shows that the results change very little when the innovations v_{t+1} are positively ($\psi = 0.5$) or negatively ($\psi = -0.5$) autocorrelated. While the two-sided $t_{h,ivx}^{trf,res}$ and $t_{h,ivx}^{rev,PL}$ tests both show good finite sample size control it can be seen from the results in Tables 1–3 that for one-sided alternatives ($H_1 : \beta_h < 0$ and $H_1 : \beta_h > 0$) $t_{h,ivx}^{trf,res}$ displays considerably better finite sample size control than $t_{h,ivx}^{rev,PL}$. This is particularly evident in the case of the right-sided tests. To illustrate, the right-sided version of $t_{h,ivx}^{trf,res}$ displays empirical rejection frequencies, taken across all of the results in Tables 1–3, in the range $[0.017, 0.057]$ for $T = 100$, $[0.020, 0.054]$ for $T = 250$ and $[0.022, 0.054]$ for $T = 500$ when $\phi = -0.15$ and in the range $[0.017, 0.078]$ for $T = 100$, $[0.021, 0.074]$ for $T = 250$ and $[0.038, 0.065]$ for $T = 500$ when $\phi = -0.95$. Whereas the right-sided version of $t_{h,ivx}^{rev,PL}$ displays rejection frequencies in the range $[0.028, 0.062]$ for T=100, $[0.026, 0.061]$ for $T = 250$, and $[0.027, 0.058]$ for $T = 500$ when

[9] We are grateful to Ke-Li Xu for making code for computing his test available on his website https://sites.google.com/site/xukeli2015/research.

[10] To ensure a fair comparison for the Bonferroni tests we exclude the case $c = -50$ in the smallest sample size ($T = 100$) where the implied autoregressive root is 0.5 and, hence, very poorly approximated by local-to-unity asymptotics; see also discussion in Phillips (2014).

[11] Additional results are reported in the Supplementary Appendix for the cases where: (i) $(u_{t+1}, \varpi_{t+1})'$ is conditionally heteroskedastic with a $GARCH(1,1)$ formulation characterising the volatility dynamics, and (ii) the unconditional variances of u_{t+1} and ϖ_{t+1} are allowed to display a one-time break at $T/4$, $T/2$, and $3T/4$. The results for (i) (see Tables S.4–S.6) are qualitatively similar to those reported here for i.i.d. innovations for all of the tests reported. For (ii) (see Tables S.7–S.33), for both $t_{h,ivx}^{trf,res}$ and $t_{h,ivx}^{rev,PL}$ the size results are again very similar to those for the i.i.d. case, while for t_h^{Bonf}, t_h^{Xu} larger size distortions are seen relative to results for these tests for the i.i.d. case.

[12] Notice that because we report results for both left-sided and right-sided tests we do not need to report results for the case where $\phi = \{0.15, 0.50, 0.95\}$ because, as noted in Campbell and Yogo (2006a), flipping the sign of ϕ also flips the sign of β. Consequently, the empirical size and power properties for the left-sided and right-sided implementations of any given test in what follows for $\phi = \{-0.15, -0.50, -0.95\}$ will be identical to those for the right-sided and left-sided implementations of those tests, respectively, for $\phi = \{0.15, 0.50, 0.95\}$.

Table 1

Empirical rejection frequencies of one-sided (left and right tail) and two-sided long-horizon predictability tests, for sample sizes $T = 100$, 250 and 500. **DGP (homoskedastic IID innovations):** $y_{t+1} = \beta x_t + u_{t+1}, x_{t+1} = \rho x_t + v_{t+1}$ and $v_{t+1} = \psi v_t + \varpi_{t+1}$, where $\beta = 0$, $\rho = 1 - c/T$, $\psi = 0$ and $(u_{t+1}, \varpi_{t+1})' \sim NIID(\mathbf{0}, \mathbf{\Sigma})$, with $\mathbf{\Sigma} = [1 \quad \phi; \quad \phi \quad 1]$.

h	c	t_h^{Xu}	t_h^{Bonf}	$t_{h,ivx}^{rev,PL}$	$t_{h,ivx}^{trf,res}$	t_h^{Xu}	t_h^{Bonf}	$t_{h,ivx}^{rev,PL}$	$t_{h,ivx}^{trf,res}$	t_h^{Xu}	t_h^{Bonf}	$t_{h,ivx}^{rev,PL}$	$t_{h,ivx}^{trf,res}$
		$T = 100$				$T = 250$				$T = 500$			
		Left-tail tests ($H_0 : \beta_h = 0$ vs $H_a : \beta_h < 0$) and $\phi = -0.15$											
5	0	0.020	0.035	0.015	0.017	0.027	0.036	0.015	0.016	0.034	0.038	0.014	0.015
	−5	0.048	0.034	0.034	0.037	0.048	0.036	0.032	0.035	0.049	0.035	0.029	0.032
	−10	0.051	0.031	0.037	0.042	0.052	0.036	0.038	0.042	0.050	0.034	0.037	0.040
	−20	0.051	0.023	0.037	0.040	0.051	0.034	0.041	0.046	0.050	0.035	0.041	0.042
	−50	–	–	–	–	0.050	0.027	0.039	0.042	0.050	0.033	0.041	0.044
10	0	0.015	0.034	0.016	0.016	0.020	0.035	0.015	0.016	0.028	0.038	0.013	0.016
	−5	0.045	0.031	0.035	0.037	0.046	0.036	0.032	0.034	0.048	0.034	0.030	0.032
	−10	0.053	0.030	0.038	0.041	0.052	0.033	0.039	0.042	0.049	0.034	0.038	0.038
	−20	0.056	0.016	0.036	0.039	0.050	0.030	0.041	0.045	0.052	0.032	0.041	0.041
	−50	–	–	–	–	0.053	0.019	0.039	0.040	0.052	0.026	0.042	0.042
20	0	0.014	0.027	0.016	0.017	0.015	0.035	0.015	0.015	0.020	0.038	0.014	0.015
	−5	0.043	0.027	0.035	0.038	0.045	0.032	0.033	0.035	0.047	0.033	0.031	0.032
	−10	0.059	0.018	0.039	0.042	0.053	0.029	0.040	0.040	0.051	0.033	0.037	0.038
	−20	0.065	0.005	0.035	0.035	0.054	0.022	0.042	0.041	0.052	0.030	0.041	0.043
	−50	–	–	–	–	0.057	0.007	0.039	0.036	0.053	0.015	0.039	0.040
50	0	0.044	0.001	0.017	0.041	0.014	0.027	0.017	0.018	0.013	0.037	0.015	0.015
	−5	0.056	0.000	0.034	0.038	0.044	0.025	0.036	0.033	0.045	0.034	0.033	0.035
	−10	0.079	0.000	0.035	0.031	0.061	0.018	0.039	0.037	0.053	0.027	0.040	0.040
	−20	0.075	0.000	0.026	0.023	0.064	0.007	0.038	0.035	0.058	0.018	0.041	0.041
	−50	–	–	–	–	0.060	0.000	0.029	0.026	0.061	0.003	0.040	0.037

h	c	t_h^{Xu}	t_h^{Bonf}	$t_{h,ivx}^{rev,PL}$	$t_{h,ivx}^{trf,res}$	t_h^{Xu}	t_h^{Bonf}	$t_{h,ivx}^{rev,PL}$	$t_{h,ivx}^{trf,res}$	t_h^{Xu}	t_h^{Bonf}	$t_{h,ivx}^{rev,PL}$	$t_{h,ivx}^{trf,res}$
		$T = 100$				$T = 250$				$T = 500$			
		Left-tail tests ($H_0 : \beta_h = 0$ vs $H_a : \beta_h < 0$) and $\phi = -0.95$											
5	0	0.000	0.005	0.001	0.001	0.000	0.018	0.001	0.001	0.000	0.025	0.001	0.001
	−5	0.096	0.020	0.007	0.007	0.109	0.030	0.008	0.005	0.115	0.038	0.009	0.005
	−10	0.093	0.029	0.013	0.019	0.104	0.040	0.015	0.015	0.104	0.047	0.017	0.016
	−20	0.059	0.047	0.020	0.036	0.067	0.064	0.023	0.030	0.074	0.066	0.025	0.030
	−50	–	–	–	–	0.053	0.169	0.028	0.044	0.059	0.156	0.033	0.044
10	0	0.000	0.002	0.001	0.000	0.000	0.011	0.001	0.001	0.000	0.018	0.001	0.001
	−5	0.080	0.013	0.007	0.006	0.104	0.024	0.007	0.005	0.110	0.033	0.008	0.005
	−10	0.083	0.019	0.013	0.020	0.097	0.034	0.016	0.014	0.103	0.041	0.016	0.016
	−20	0.048	0.022	0.015	0.036	0.061	0.050	0.023	0.030	0.070	0.058	0.027	0.030
	−50	–	–	–	–	0.046	0.119	0.026	0.047	0.055	0.130	0.032	0.047
20	0	0.004	0.001	0.001	0.001	0.000	0.004	0.001	0.000	0.000	0.011	0.001	0.000
	−5	0.055	0.004	0.006	0.009	0.091	0.015	0.007	0.004	0.105	0.028	0.008	0.005
	−10	0.066	0.006	0.010	0.022	0.087	0.023	0.012	0.014	0.097	0.034	0.016	0.016
	−20	0.042	0.004	0.012	0.034	0.051	0.028	0.020	0.031	0.063	0.045	0.025	0.03
	−50	–	–	–	–	0.047	0.044	0.021	0.045	0.051	0.082	0.031	0.046
50	0	0.166	0.000	0.003	0.025	0.005	0.001	0.001	0.001	0.000	0.003	0.000	0.000
	−5	0.028	0.000	0.009	0.033	0.059	0.003	0.006	0.007	0.085	0.016	0.007	0.004
	−10	0.050	0.000	0.007	0.028	0.067	0.006	0.008	0.018	0.083	0.019	0.014	0.017
	−20	0.062	0.000	0.005	0.024	0.041	0.004	0.011	0.034	0.050	0.019	0.020	0.029
	−50	–	–	–	–	0.054	0.003	0.011	0.032	0.051	0.015	0.025	0.043

(*continued on next page*)

Table 1 (*continued*).

h	c	t_h^{Xu}	t_h^{Bonf}	$t_{h,ivx}^{rev,PL}$	$t_{h,ivx}^{trf,res}$	t_h^{Xu}	t_h^{Bonf}	$t_{h,ivx}^{rev,PL}$	$t_{h,ivx}^{trf,res}$	t_h^{Xu}	t_h^{Bonf}	$t_{h,ivx}^{rev,PL}$	$t_{h,ivx}^{trf,res}$	t_h^{Xu}	t_h^{Bonf}	$t_{h,ivx}^{rev,PL}$	$t_{h,ivx}^{trf,res}$	t_h^{Xu}	t_h^{Bonf}	$t_{h,ivx}^{rev,PL}$	$t_{h,ivx}^{trf,res}$	t_h^{Xu}	t_h^{Bonf}	$t_{h,ivx}^{rev,PL}$	$t_{h,ivx}^{trf,res}$
		$T = 100$				$T = 250$				$T = 500$				$T = 100$				$T = 250$				$T = 500$			
		Right-tail tests ($H_0 : \beta_h = 0$ vs $H_a : \beta_h > 0$) and $\phi = -0.15$												***Right-tail tests*** ($H_0 : \beta_h = 0$ vs $H_a : \beta_h > 0$) and $\phi = -0.95$											
5	0	0.044	0.043	0.032	0.028	0.050	0.039	0.027	0.024	0.056	0.038	0.027	0.025	0.120	0.048	0.117	0.067	0.163	0.037	0.103	0.063	0.174	0.032	0.103	0.058
	−5	0.059	0.037	0.053	0.049	0.054	0.036	0.047	0.044	0.051	0.035	0.047	0.044	0.056	0.044	0.108	0.073	0.073	0.039	0.105	0.065	0.077	0.032	0.104	0.059
	−10	0.061	0.031	0.055	0.051	0.055	0.033	0.049	0.048	0.051	0.035	0.049	0.047	0.051	0.036	0.096	0.073	0.059	0.033	0.092	0.067	0.062	0.032	0.093	0.058
	−20	0.062	0.022	0.050	0.047	0.056	0.031	0.050	0.049	0.050	0.033	0.052	0.050	0.050	0.025	0.074	0.073	0.054	0.026	0.082	0.071	0.056	0.025	0.08	0.059
	−50	–	–	–	–	0.056	0.022	0.046	0.043	0.051	0.027	0.05	0.049	–	–	–	–	0.049	0.018	0.061	0.074	0.052	0.016	0.065	0.063
10	0	0.035	0.039	0.034	0.027	0.041	0.037	0.027	0.024	0.049	0.038	0.028	0.024	0.095	0.045	0.119	0.062	0.128	0.036	0.103	0.061	0.152	0.032	0.104	0.057
	−5	0.066	0.032	0.053	0.046	0.055	0.033	0.048	0.043	0.05	0.035	0.046	0.042	0.065	0.032	0.114	0.063	0.060	0.036	0.106	0.065	0.066	0.031	0.104	0.060
	−10	0.071	0.027	0.055	0.047	0.058	0.028	0.051	0.046	0.052	0.032	0.048	0.044	0.061	0.020	0.097	0.061	0.050	0.027	0.095	0.067	0.056	0.030	0.093	0.059
	−20	0.072	0.015	0.051	0.043	0.059	0.022	0.051	0.045	0.052	0.03	0.049	0.046	0.06	0.010	0.076	0.057	0.048	0.018	0.081	0.070	0.051	0.023	0.080	0.060
	−50	–	–	–	–	0.062	0.013	0.047	0.043	0.054	0.022	0.049	0.047	–	–	–	–	0.052	0.007	0.061	0.069	0.048	0.011	0.065	0.065
20	0	0.037	0.027	0.035	0.026	0.035	0.038	0.031	0.024	0.04	0.037	0.029	0.024	0.169	0.027	0.116	0.05	0.092	0.033	0.105	0.059	0.119	0.030	0.100	0.054
	−5	0.079	0.017	0.054	0.041	0.062	0.03	0.049	0.042	0.052	0.033	0.048	0.041	0.120	0.016	0.120	0.047	0.058	0.026	0.108	0.060	0.053	0.028	0.105	0.058
	−10	0.087	0.011	0.055	0.041	0.064	0.023	0.055	0.045	0.054	0.028	0.051	0.045	0.083	0.008	0.111	0.043	0.054	0.016	0.097	0.060	0.047	0.025	0.096	0.056
	−20	0.085	0.004	0.05	0.037	0.067	0.017	0.051	0.044	0.054	0.025	0.054	0.045	0.064	0.003	0.089	0.036	0.054	0.007	0.086	0.057	0.047	0.019	0.084	0.059
	−50	–	–	–	–	0.069	0.005	0.041	0.035	0.06	0.013	0.05	0.044	–	–	–	–	0.057	0.001	0.060	0.046	0.050	0.004	0.069	0.060
50	0	0.064	0.001	0.029	0.041	0.036	0.024	0.034	0.023	0.031	0.035	0.031	0.022	0.347	0.011	0.078	0.049	0.168	0.020	0.109	0.042	0.086	0.027	0.101	0.046
	−5	0.108	0.000	0.050	0.038	0.074	0.018	0.056	0.037	0.061	0.027	0.049	0.040	0.175	0.007	0.132	0.031	0.123	0.012	0.112	0.039	0.061	0.018	0.107	0.047
	−10	0.111	0.000	0.054	0.031	0.082	0.013	0.056	0.037	0.066	0.023	0.053	0.042	0.091	0.003	0.140	0.023	0.082	0.007	0.105	0.037	0.057	0.013	0.102	0.046
	−20	0.094	0.000	0.047	0.021	0.083	0.005	0.054	0.036	0.067	0.014	0.055	0.040	0.065	0.000	0.121	0.017	0.066	0.002	0.095	0.035	0.054	0.006	0.096	0.046
	−50	–	–	–	–	0.070	0.000	0.039	0.027	0.064	0.003	0.047	0.036	–	–	–	–	0.058	0.000	0.067	0.026	0.055	0.001	0.076	0.045
		Two-sided tests ($H_0 : \beta_h = 0$ vs $H_a : \beta_h \neq 0$) and $\phi = -0.15$												***Two-sided tests*** ($H_0 : \beta_h = 0$ vs $H_a : \beta_h \neq 0$) and $\phi = -0.95$											
5	0	0.030	0.041	0.022	0.022	0.039	0.037	0.018	0.018	0.041	0.038	0.019	0.018	0.068	0.040	0.058	0.037	0.096	0.040	0.049	0.034	0.098	0.040	0.048	0.031
	−5	0.055	0.035	0.044	0.045	0.054	0.036	0.04	0.039	0.051	0.037	0.037	0.037	0.075	0.048	0.062	0.043	0.098	0.053	0.056	0.037	0.106	0.054	0.056	0.032
	−10	0.057	0.029	0.046	0.046	0.055	0.034	0.044	0.045	0.054	0.036	0.043	0.044	0.086	0.048	0.058	0.049	0.099	0.058	0.056	0.045	0.105	0.061	0.055	0.036
	−20	0.059	0.020	0.041	0.043	0.057	0.033	0.046	0.046	0.053	0.033	0.047	0.047	0.057	0.051	0.048	0.058	0.064	0.070	0.051	0.052	0.073	0.073	0.054	0.047
	−50	–	–	–	–	0.053	0.022	0.042	0.043	0.051	0.027	0.045	0.046	–	–	–	–	0.049	0.159	0.043	0.062	0.056	0.149	0.048	0.055
10	0	0.024	0.038	0.022	0.02	0.032	0.037	0.018	0.018	0.032	0.038	0.018	0.017	0.067	0.032	0.058	0.035	0.072	0.033	0.052	0.034	0.086	0.035	0.046	0.030
	−5	0.059	0.034	0.045	0.044	0.057	0.035	0.038	0.037	0.053	0.034	0.037	0.036	0.077	0.029	0.061	0.036	0.079	0.041	0.057	0.036	0.095	0.049	0.053	0.032
	−10	0.068	0.026	0.05	0.049	0.057	0.031	0.044	0.043	0.052	0.033	0.042	0.041	0.092	0.023	0.057	0.041	0.087	0.045	0.055	0.043	0.099	0.054	0.056	0.037
	−20	0.074	0.012	0.046	0.043	0.06	0.024	0.046	0.046	0.053	0.029	0.046	0.043	0.061	0.018	0.044	0.048	0.055	0.048	0.052	0.051	0.064	0.063	0.052	0.047
	−50	–	–	–	–	0.061	0.011	0.042	0.042	0.055	0.021	0.044	0.044	–	–	–	–	0.047	0.099	0.043	0.061	0.051	0.115	0.051	0.058
20	0	0.027	0.027	0.025	0.023	0.026	0.037	0.02	0.018	0.027	0.038	0.018	0.017	0.155	0.015	0.063	0.025	0.058	0.026	0.054	0.030	0.062	0.029	0.048	0.028
	−5	0.070	0.021	0.047	0.041	0.058	0.032	0.041	0.035	0.051	0.033	0.038	0.037	0.117	0.008	0.067	0.028	0.073	0.028	0.057	0.032	0.078	0.041	0.055	0.030
	−10	0.090	0.013	0.051	0.043	0.066	0.026	0.046	0.04	0.056	0.029	0.043	0.039	0.094	0.006	0.063	0.031	0.085	0.024	0.054	0.035	0.087	0.043	0.058	0.034
	−20	0.097	0.003	0.045	0.036	0.073	0.016	0.046	0.043	0.058	0.022	0.045	0.042	0.06	0.002	0.053	0.032	0.055	0.022	0.051	0.044	0.058	0.042	0.053	0.045
	−50	–	–	–	–	0.072	0.005	0.036	0.034	0.062	0.010	0.044	0.040	–	–	–	–	0.053	0.028	0.039	0.049	0.050	0.060	0.048	0.053
50	0	0.061	0.000	0.028	0.055	0.024	0.023	0.026	0.021	0.023	0.036	0.021	0.017	0.397	0.004	0.053	0.051	0.159	0.010	0.060	0.022	0.061	0.019	0.051	0.023
	−5	0.113	0.000	0.049	0.046	0.071	0.021	0.048	0.035	0.059	0.031	0.041	0.034	0.158	0.002	0.094	0.037	0.120	0.007	0.065	0.022	0.074	0.021	0.061	0.024
	−10	0.140	0.000	0.054	0.031	0.091	0.013	0.053	0.038	0.069	0.022	0.046	0.040	0.085	0.001	0.099	0.028	0.094	0.005	0.063	0.025	0.088	0.019	0.061	0.029
	−20	0.119	0.000	0.046	0.020	0.100	0.004	0.047	0.034	0.074	0.012	0.049	0.041	0.083	0.000	0.085	0.017	0.058	0.002	0.058	0.030	0.058	0.013	0.060	0.036
	−50	–	–	–	–	0.079	0.000	0.034	0.022	0.075	0.001	0.043	0.036	–	–	–	–	0.060	0.001	0.041	0.027	0.056	0.006	0.052	0.042

Notes: t_h^{Xu} denotes the implied statistic of Xu (2020), t_h^{Bonf} is the Bonferroni based statistic of Hjalmarsson (2011), $t_{h,ivx}^{trf,res}$ is the residual augmented transformed regression based statistic in (4.9) proposed in Section 4.2; and $t_{h,ivx}^{trf,PL}$ is the Phillips and Lee (2013) statistic. h is the forecast horizon considered and c is the local to unity parameter that characterises the persistence of the predictor.

Table 2
Empirical rejection frequencies of one-sided (left and right tail) and two-sided long-horizon predictability tests, for sample sizes $T = 100$, 250 and 500. **DGP (Positive Autocorrelation):** $y_{t+1} = \beta x_t + u_{t+1}$, $x_{t+1} = \rho x_t + v_{t+1}$ and $v_{t+1} = \psi v_t + \varpi_{t+1}$, where $\beta = 0$, $\rho = 1 - c/T$, $\psi = 0.50$ and $(u_{t+1}, \varpi_{t+1})' \sim NIID(\mathbf{0}, \mathbf{\Sigma})$, with $\mathbf{\Sigma} = [1 \quad \phi; \quad \phi \quad 1]$.

h	c	t_h^{Xu}	t_h^{Bonf}	$t_{h,ivx}^{rev,PL}$	$t_{h,ivx}^{trf,res}$	t_h^{Xu}	t_h^{Bonf}	$t_{h,ivx}^{rev,PL}$	$t_{h,ivx}^{trf,res}$	t_h^{Xu}	t_h^{Bonf}	$t_{h,ivx}^{rev,PL}$	$t_{h,ivx}^{trf,res}$
		$T = 100$				$T = 250$				$T = 500$			
		Left-tail tests ($H_0 : \beta_h = 0$ vs $H_a : \beta_h < 0$) and $\phi = -0.15$											
5	0	0.010	0.043	0.015	0.017	0.016	0.043	0.015	0.017	0.024	0.043	0.013	0.015
	−5	0.028	0.050	0.035	0.039	0.038	0.049	0.032	0.035	0.042	0.048	0.030	0.032
	−10	0.046	0.053	0.039	0.044	0.050	0.053	0.039	0.043	0.053	0.053	0.037	0.038
	−20	0.054	0.051	0.043	0.048	0.056	0.059	0.044	0.046	0.057	0.058	0.041	0.043
	−50	–	–	–	–	0.052	0.060	0.045	0.047	0.050	0.068	0.046	0.046
10	0	0.008	0.048	0.014	0.016	0.010	0.045	0.015	0.017	0.015	0.047	0.014	0.015
	−5	0.017	0.053	0.036	0.038	0.031	0.051	0.034	0.034	0.037	0.053	0.030	0.033
	−10	0.038	0.055	0.041	0.043	0.046	0.058	0.040	0.042	0.051	0.057	0.037	0.039
	−20	0.054	0.044	0.042	0.045	0.056	0.059	0.044	0.047	0.059	0.064	0.042	0.042
	−50	–	–	–	–	0.053	0.049	0.045	0.047	0.052	0.067	0.044	0.046
20	0	0.030	0.034	0.015	0.018	0.009	0.048	0.015	0.014	0.010	0.047	0.014	0.015
	−5	0.012	0.038	0.037	0.040	0.020	0.054	0.034	0.034	0.028	0.054	0.031	0.033
	−10	0.029	0.030	0.043	0.044	0.041	0.055	0.041	0.040	0.046	0.056	0.038	0.038
	−20	0.056	0.017	0.043	0.044	0.055	0.050	0.044	0.043	0.057	0.060	0.044	0.044
	−50	–	–	–	–	0.055	0.024	0.046	0.045	0.051	0.047	0.042	0.043
50	0	0.108	0.002	0.017	0.049	0.032	0.029	0.017	0.019	0.007	0.047	0.015	0.015
	−5	0.014	0.001	0.035	0.046	0.011	0.028	0.037	0.035	0.016	0.053	0.033	0.034
	−10	0.030	0.001	0.039	0.038	0.031	0.019	0.041	0.040	0.036	0.048	0.042	0.041
	−20	0.077	0.000	0.036	0.029	0.057	0.012	0.044	0.039	0.059	0.038	0.044	0.043
	−50	–	–	–	–	0.066	0.003	0.039	0.033	0.057	0.015	0.047	0.044
		Right-tail tests ($H_0 : \beta_h = 0$ vs $H_a : \beta_h > 0$) and $\phi = -0.15$											
5	0	0.021	0.045	0.030	0.027	0.039	0.038	0.026	0.023	0.049	0.037	0.027	0.025
	−5	0.038	0.040	0.054	0.051	0.043	0.034	0.047	0.044	0.044	0.033	0.046	0.042
	−10	0.046	0.036	0.058	0.056	0.044	0.029	0.051	0.047	0.043	0.029	0.050	0.047
	−20	0.054	0.025	0.058	0.057	0.049	0.027	0.054	0.052	0.046	0.027	0.051	0.049
	−50	–	–	–	–	0.050	0.020	0.053	0.053	0.049	0.019	0.054	0.052
10	0	0.015	0.043	0.033	0.029	0.024	0.039	0.027	0.024	0.033	0.036	0.028	0.024
	−5	0.028	0.035	0.053	0.047	0.036	0.032	0.048	0.044	0.039	0.032	0.045	0.041
	−10	0.047	0.030	0.059	0.052	0.043	0.026	0.052	0.047	0.043	0.026	0.050	0.045
	−20	0.060	0.020	0.059	0.051	0.049	0.022	0.054	0.049	0.045	0.023	0.051	0.048
	−50	–	–	–	–	0.055	0.015	0.054	0.050	0.051	0.015	0.051	0.050
20	0	0.028	0.028	0.034	0.028	0.016	0.039	0.030	0.024	0.019	0.037	0.029	0.024
	−5	0.024	0.015	0.056	0.044	0.028	0.032	0.050	0.043	0.033	0.031	0.049	0.042
	−10	0.049	0.009	0.060	0.045	0.040	0.023	0.056	0.046	0.042	0.026	0.053	0.045
	−20	0.073	0.004	0.060	0.047	0.055	0.017	0.056	0.049	0.047	0.019	0.056	0.046
	−50	–	–	–	–	0.064	0.008	0.051	0.044	0.056	0.012	0.054	0.048
50	0	0.077	0.002	0.028	0.052	0.026	0.022	0.034	0.023	0.011	0.038	0.031	0.022
	−5	0.032	0.001	0.052	0.047	0.025	0.014	0.056	0.038	0.025	0.028	0.049	0.041
	−10	0.065	0.000	0.061	0.040	0.047	0.006	0.059	0.039	0.042	0.020	0.055	0.044
	−20	0.107	0.000	0.062	0.029	0.073	0.003	0.061	0.039	0.055	0.011	0.058	0.043
	−50	–	–	–	–	0.085	0.001	0.053	0.035	0.066	0.004	0.056	0.043

h	c	t_h^{Xu}	t_h^{Bonf}	$t_{h,ivx}^{rev,PL}$	$t_{h,ivx}^{trf,res}$	t_h^{Xu}	t_h^{Bonf}	$t_{h,ivx}^{rev,PL}$	$t_{h,ivx}^{trf,res}$	t_h^{Xu}	t_h^{Bonf}	$t_{h,ivx}^{rev,PL}$	$t_{h,ivx}^{trf,res}$
		$T = 100$				$T = 250$				$T = 500$			
		Left-tail tests ($H_0 : \beta_h = 0$ vs $H_a : \beta_h < 0$) and $\phi = -0.95$											
5	0	0.001	0.016	0.001	0.001	0.000	0.035	0.001	0.001	0.000	0.043	0.001	0.001
	−5	0.009	0.052	0.008	0.006	0.017	0.072	0.008	0.005	0.025	0.079	0.009	0.005
	−10	0.094	0.139	0.013	0.013	0.115	0.188	0.016	0.013	0.125	0.200	0.017	0.015
	−20	0.149	0.254	0.020	0.026	0.163	0.355	0.024	0.026	0.170	0.382	0.027	0.028
	−50	–	–	–	–	0.081	0.652	0.030	0.038	0.093	0.735	0.035	0.042
10	0	0.005	0.011	0.001	0.000	0.001	0.041	0.001	0.001	0.000	0.059	0.001	0.001
	−5	0.001	0.076	0.007	0.005	0.010	0.132	0.007	0.005	0.017	0.152	0.008	0.005
	−10	0.067	0.180	0.013	0.012	0.101	0.302	0.016	0.013	0.118	0.341	0.016	0.014
	−20	0.137	0.256	0.017	0.026	0.157	0.508	0.023	0.026	0.170	0.586	0.028	0.028
	−50	–	–	–	–	0.074	0.717	0.028	0.040	0.089	0.893	0.034	0.042
20	0	0.055	0.003	0.001	0.001	0.005	0.031	0.001	0.000	0.001	0.061	0.001	0.000
	−5	0.002	0.029	0.005	0.007	0.002	0.150	0.007	0.004	0.008	0.197	0.008	0.004
	−10	0.031	0.071	0.010	0.014	0.074	0.316	0.012	0.012	0.102	0.408	0.015	0.014
	−20	0.109	0.060	0.013	0.027	0.146	0.454	0.020	0.026	0.164	0.640	0.025	0.029
	−50	–	–	–	–	0.064	0.446	0.024	0.042	0.082	0.851	0.032	0.044
50	0	0.326	0.000	0.004	0.025	0.064	0.003	0.001	0.001	0.006	0.028	0.000	0.001
	−5	0.014	0.000	0.009	0.035	0.003	0.039	0.006	0.007	0.001	0.161	0.007	0.004
	−10	0.010	0.000	0.009	0.029	0.023	0.082	0.008	0.017	0.064	0.331	0.014	0.014
	−20	0.071	0.000	0.007	0.023	0.114	0.066	0.011	0.029	0.149	0.433	0.020	0.028
	−50	–	–	–	–	0.055	0.031	0.013	0.034	0.066	0.372	0.027	0.040
		Right-tail tests ($H_0 : \beta_h = 0$ vs $H_a : \beta_h > 0$) and $\phi = -0.95$											
5	0	0.107	0.007	0.112	0.062	0.167	0.004	0.101	0.060	0.200	0.004	0.104	0.056
	−5	0.046	0.001	0.111	0.065	0.050	0.001	0.104	0.059	0.049	0.001	0.103	0.055
	−10	0.038	0.001	0.100	0.068	0.042	0.000	0.098	0.061	0.041	0.000	0.093	0.056
	−20	0.038	0.000	0.087	0.068	0.041	0.000	0.085	0.061	0.040	0.000	0.084	0.056
	−50	–	–	–	–	0.040	0.000	0.072	0.067	0.042	0.000	0.070	0.057
10	0	0.063	0.003	0.113	0.058	0.112	0.001	0.102	0.060	0.159	0.002	0.101	0.057
	−5	0.054	0.000	0.114	0.064	0.044	0.000	0.106	0.059	0.046	0.000	0.104	0.055
	−10	0.053	0.000	0.105	0.061	0.037	0.000	0.097	0.062	0.038	0.000	0.095	0.055
	−20	0.054	0.000	0.092	0.063	0.036	0.000	0.087	0.064	0.037	0.000	0.083	0.057
	−50	–	–	–	–	0.040	0.000	0.073	0.069	0.039	0.000	0.070	0.060
20	0	0.063	0.003	0.106	0.049	0.071	0.001	0.103	0.057	0.107	0.001	0.097	0.052
	−5	0.096	0.000	0.118	0.050	0.046	0.000	0.105	0.056	0.041	0.000	0.105	0.054
	−10	0.109	0.000	0.117	0.047	0.043	0.000	0.101	0.056	0.035	0.000	0.098	0.052
	−20	0.100	0.000	0.108	0.043	0.046	0.000	0.093	0.058	0.034	0.000	0.088	0.055
	−50	–	–	–	–	0.050	0.000	0.078	0.057	0.038	0.000	0.077	0.058
50	0	0.094	0.011	0.072	0.064	0.063	0.002	0.108	0.043	0.056	0.002	0.101	0.045
	−5	0.187	0.004	0.131	0.046	0.099	0.000	0.114	0.038	0.048	0.000	0.106	0.047
	−10	0.222	0.002	0.150	0.036	0.110	0.000	0.110	0.035	0.050	0.000	0.103	0.045
	−20	0.153	0.000	0.151	0.029	0.106	0.000	0.109	0.038	0.054	0.000	0.100	0.046
	−50	–	–	–	–	0.068	0.000	0.095	0.035	0.054	0.000	0.091	0.046

(*continued on next page*)

Table 2 (*continued*).

h	c	t_h^{Xu}	t_h^{Bonf}	$t_{h,ivx}^{rev,PL}$	$t_{h,ivx}^{trf,res}$	t_h^{Xu}	t_h^{Bonf}	$t_{h,ivx}^{rev,PL}$	$t_{h,ivx}^{trf,res}$	t_h^{Xu}	t_h^{Bonf}	$t_{h,ivx}^{rev,PL}$	$t_{h,ivx}^{trf,res}$	t_h^{Xu}	t_h^{Bonf}	$t_{h,ivx}^{rev,PL}$	$t_{h,ivx}^{trf,res}$	t_h^{Xu}	t_h^{Bonf}	$t_{h,ivx}^{rev,PL}$	$t_{h,ivx}^{trf,res}$	t_h^{Xu}	t_h^{Bonf}	$t_{h,ivx}^{rev,PL}$	$t_{h,ivx}^{trf,res}$
		$T = 100$				$T = 250$				$T = 500$				$T = 100$				$T = 250$				$T = 500$			
		Two-sided tests ($H_0 : \beta_h = 0$ vs $H_a : \beta_h \neq 0$) and $\phi = -0.15$												***Two-sided tests*** ($H_0 : \beta_h = 0$ vs $H_a : \beta_h \neq 0$) and $\phi = -0.95$											
5	0	0.013	0.048	0.022	0.022	0.025	0.039	0.018	0.019	0.034	0.039	0.018	0.018	0.061	0.011	0.055	0.034	0.095	0.026	0.049	0.031	0.114	0.030	0.046	0.029
	−5	0.031	0.047	0.044	0.044	0.039	0.043	0.039	0.039	0.043	0.040	0.037	0.036	0.027	0.035	0.062	0.037	0.032	0.051	0.056	0.031	0.035	0.057	0.053	0.031
	−10	0.047	0.044	0.053	0.052	0.049	0.044	0.043	0.046	0.048	0.042	0.043	0.044	0.060	0.107	0.062	0.042	0.079	0.144	0.055	0.035	0.086	0.154	0.055	0.032
	−20	0.056	0.039	0.049	0.050	0.054	0.044	0.049	0.049	0.051	0.043	0.049	0.049	0.115	0.205	0.057	0.050	0.127	0.295	0.057	0.046	0.140	0.318	0.054	0.041
	−50	–	–	–	–	0.053	0.041	0.050	0.053	0.051	0.045	0.051	0.049	–	–	–	–	0.065	0.587	0.050	0.051	0.077	0.678	0.051	0.048
10	0	0.009	0.051	0.023	0.021	0.015	0.043	0.019	0.019	0.022	0.040	0.017	0.016	0.043	0.006	0.055	0.032	0.067	0.028	0.051	0.032	0.088	0.041	0.047	0.029
	−5	0.021	0.049	0.044	0.044	0.031	0.045	0.039	0.037	0.036	0.042	0.037	0.035	0.038	0.049	0.062	0.032	0.026	0.095	0.057	0.031	0.029	0.112	0.053	0.030
	−10	0.043	0.046	0.055	0.051	0.044	0.044	0.045	0.043	0.046	0.043	0.042	0.041	0.058	0.137	0.056	0.038	0.062	0.247	0.056	0.036	0.074	0.278	0.055	0.033
	−20	0.062	0.032	0.055	0.053	0.056	0.043	0.051	0.049	0.053	0.044	0.047	0.044	0.117	0.198	0.054	0.042	0.119	0.437	0.058	0.046	0.131	0.520	0.055	0.042
	−50	–	–	–	–	0.056	0.031	0.051	0.049	0.051	0.042	0.050	0.050	–	–	–	–	0.060	0.648	0.052	0.057	0.071	0.852	0.055	0.052
20	0	0.026	0.031	0.025	0.025	0.010	0.044	0.019	0.018	0.012	0.043	0.018	0.017	0.068	0.002	0.060	0.025	0.049	0.020	0.053	0.029	0.058	0.042	0.048	0.027
	−5	0.016	0.027	0.049	0.045	0.021	0.045	0.041	0.037	0.028	0.042	0.037	0.036	0.080	0.013	0.068	0.026	0.030	0.106	0.058	0.031	0.024	0.150	0.056	0.029
	−10	0.037	0.017	0.058	0.048	0.042	0.041	0.047	0.042	0.043	0.042	0.043	0.040	0.098	0.041	0.069	0.029	0.051	0.255	0.056	0.033	0.060	0.343	0.061	0.032
	−20	0.073	0.008	0.057	0.046	0.061	0.035	0.051	0.046	0.055	0.041	0.047	0.045	0.136	0.033	0.067	0.035	0.118	0.376	0.057	0.040	0.125	0.565	0.055	0.040
	−50	–	–	–	–	0.065	0.014	0.049	0.044	0.054	0.027	0.050	0.046	–	–	–	–	0.061	0.341	0.051	0.046	0.064	0.792	0.056	0.051
50	0	0.105	0.001	0.028	0.074	0.027	0.025	0.026	0.021	0.008	0.044	0.021	0.018	0.283	0.004	0.050	0.066	0.075	0.001	0.059	0.021	0.039	0.016	0.050	0.022
	−5	0.024	0.000	0.051	0.056	0.017	0.018	0.048	0.036	0.018	0.043	0.042	0.034	0.166	0.001	0.095	0.050	0.084	0.019	0.065	0.020	0.034	0.118	0.060	0.024
	−10	0.057	0.000	0.062	0.042	0.036	0.012	0.055	0.041	0.037	0.035	0.048	0.040	0.204	0.001	0.107	0.036	0.096	0.046	0.065	0.023	0.048	0.264	0.062	0.028
	−20	0.126	0.000	0.061	0.029	0.073	0.005	0.054	0.040	0.060	0.022	0.051	0.042	0.156	0.000	0.107	0.024	0.144	0.033	0.066	0.030	0.122	0.343	0.062	0.034
	−50	–	–	–	–	0.099	0.001	0.047	0.033	0.070	0.006	0.051	0.042	–	–	–	–	0.071	0.012	0.058	0.032	0.067	0.261	0.061	0.042

Notes: See Notes to Table 1.

Table 3
Empirical rejection frequencies of one-sided (left and right tail) and two-sided long-horizon predictability tests, for sample sizes $T = 100$, 250 and 500. **DGP (Negative Autocorrelation):** $y_{t+1} = \beta x_t + u_{t+1}, x_{t+1} = \rho x_t + v_{t+1}$ and $v_{t+1} = \psi v_t + \varpi_{t+1}$, where $\beta = 0$, $\rho = 1 - c/T$, $\psi = -0.50$ and $(u_{t+1}, \varpi_{t+1})' \sim NIID(\mathbf{0}, \mathbf{\Sigma})$, with $\mathbf{\Sigma} = [1 \quad \phi; \quad \phi \quad 1]$.

h	c	t_h^{Xu}	t_h^{Bonf}	$t_{h,ivx}^{rev,PL}$	$t_{h,ivx}^{trf,res}$	t_h^{Xu}	t_h^{Bonf}	$t_{h,ivx}^{rev,PL}$	$t_{h,ivx}^{trf,res}$	t_h^{Xu}	t_h^{Bonf}	$t_{h,ivx}^{rev,PL}$	$t_{h,ivx}^{trf,res}$
		$T = 100$				$T = 250$				$T = 500$			
		Left-tail tests ($H_0 : \beta_h = 0$ vs $H_a : \beta_h < 0$) and $\phi = -0.15$											
5	0	0.032	0.023	0.015	0.018	0.034	0.024	0.015	0.016	0.036	0.027	0.013	0.015
	−5	0.047	0.013	0.032	0.036	0.042	0.015	0.031	0.036	0.043	0.015	0.029	0.031
	−10	0.049	0.007	0.033	0.037	0.046	0.011	0.037	0.039	0.046	0.012	0.037	0.038
	−20	0.054	0.003	0.031	0.034	0.047	0.009	0.037	0.040	0.047	0.009	0.038	0.040
	−50	–	–	–	–	0.050	0.006	0.034	0.035	0.051	0.006	0.037	0.039
10	0	0.031	0.023	0.016	0.016	0.031	0.023	0.015	0.016	0.034	0.024	0.013	0.016
	−5	0.051	0.014	0.034	0.036	0.044	0.013	0.032	0.033	0.042	0.013	0.029	0.032
	−10	0.053	0.009	0.034	0.039	0.047	0.010	0.038	0.039	0.047	0.010	0.036	0.037
	−20	0.058	0.002	0.030	0.032	0.049	0.007	0.038	0.042	0.049	0.007	0.039	0.040
	−50	–	–	–	–	0.052	0.003	0.032	0.035	0.054	0.006	0.037	0.039
20	0	0.032	0.023	0.016	0.016	0.029	0.021	0.015	0.015	0.033	0.025	0.014	0.015
	−5	0.058	0.013	0.035	0.036	0.046	0.014	0.033	0.033	0.044	0.013	0.031	0.032
	−10	0.059	0.006	0.035	0.035	0.053	0.009	0.037	0.037	0.049	0.009	0.036	0.036
	−20	0.059	0.001	0.028	0.028	0.053	0.005	0.038	0.038	0.051	0.006	0.040	0.041
	−50	–	–	–	–	0.052	0.001	0.030	0.029	0.054	0.003	0.034	0.035
50	0	0.042	0.001	0.018	0.038	0.029	0.022	0.016	0.018	0.030	0.025	0.015	0.015
	−5	0.067	0.001	0.031	0.033	0.054	0.015	0.035	0.031	0.048	0.017	0.033	0.035
	−10	0.063	0.000	0.027	0.026	0.058	0.008	0.037	0.035	0.054	0.010	0.039	0.038
	−20	0.059	0.000	0.020	0.016	0.055	0.002	0.035	0.031	0.056	0.005	0.039	0.039
	−50	–	–	–	–	0.051	0.000	0.022	0.020	0.054	0.000	0.035	0.031
		Right-tail tests ($H_0 : \beta_h = 0$ vs $H_a : \beta_h > 0$) and $\phi = -0.15$											
5	0	0.070	0.046	0.033	0.028	0.066	0.049	0.028	0.024	0.067	0.049	0.027	0.025
	−5	0.073	0.036	0.052	0.048	0.064	0.051	0.046	0.044	0.057	0.059	0.047	0.044
	−10	0.074	0.027	0.049	0.047	0.063	0.048	0.048	0.046	0.057	0.063	0.048	0.046
	−20	0.072	0.012	0.043	0.042	0.065	0.039	0.047	0.046	0.055	0.063	0.050	0.048
	−50	–	–	–	–	0.060	0.017	0.040	0.039	0.053	0.048	0.045	0.044
10	0	0.073	0.043	0.035	0.027	0.064	0.045	0.028	0.024	0.066	0.048	0.028	0.024
	−5	0.083	0.031	0.050	0.045	0.067	0.044	0.046	0.043	0.060	0.056	0.045	0.042
	−10	0.081	0.020	0.051	0.043	0.067	0.038	0.049	0.044	0.059	0.055	0.048	0.045
	−20	0.074	0.006	0.042	0.037	0.068	0.027	0.047	0.042	0.059	0.047	0.048	0.045
	−50	–	–	–	–	0.060	0.006	0.040	0.037	0.055	0.027	0.044	0.043
20	0	0.078	0.035	0.036	0.026	0.067	0.039	0.030	0.024	0.065	0.043	0.029	0.024
	−5	0.093	0.025	0.051	0.037	0.076	0.036	0.048	0.040	0.064	0.047	0.047	0.041
	−10	0.086	0.012	0.048	0.037	0.074	0.027	0.051	0.042	0.063	0.041	0.049	0.044
	−20	0.075	0.002	0.040	0.030	0.072	0.013	0.046	0.039	0.062	0.031	0.050	0.043
	−50	–	–	–	–	0.061	0.001	0.033	0.029	0.055	0.009	0.044	0.040
50	0	0.098	0.002	0.029	0.038	0.078	0.033	0.034	0.022	0.069	0.038	0.031	0.022
	−5	0.102	0.001	0.049	0.033	0.089	0.029	0.054	0.037	0.076	0.038	0.048	0.040
	−10	0.088	0.000	0.047	0.026	0.082	0.018	0.052	0.036	0.071	0.029	0.052	0.041
	−20	0.075	0.000	0.039	0.017	0.072	0.004	0.047	0.032	0.065	0.014	0.051	0.038
	−50	–	–	–	–	0.061	0.000	0.028	0.020	0.057	0.001	0.040	0.031

h	c	t_h^{Xu}	t_h^{Bonf}	$t_{h,ivx}^{rev,PL}$	$t_{h,ivx}^{trf,res}$	t_h^{Xu}	t_h^{Bonf}	$t_{h,ivx}^{rev,PL}$	$t_{h,ivx}^{trf,res}$	t_h^{Xu}	t_h^{Bonf}	$t_{h,ivx}^{rev,PL}$	$t_{h,ivx}^{trf,res}$
		$T = 100$				$T = 250$				$T = 500$			
		Left-tail tests ($H_0 : \beta_h = 0$ vs $H_a : \beta_h < 0$) and $\phi = -0.95$											
5	0	0.004	0.000	0.001	0.001	0.005	0.001	0.001	0.001	0.006	0.001	0.001	0.001
	−5	0.030	0.000	0.007	0.007	0.028	0.001	0.008	0.005	0.031	0.001	0.009	0.005
	−10	0.031	0.000	0.012	0.020	0.032	0.000	0.015	0.015	0.032	0.000	0.018	0.016
	−20	0.036	0.003	0.018	0.034	0.034	0.002	0.021	0.030	0.037	0.000	0.024	0.029
	−50	–	–	–	–	0.041	0.038	0.025	0.040	0.043	0.013	0.031	0.043
10	0	0.003	0.000	0.001	0.001	0.004	0.000	0.001	0.001	0.005	0.001	0.001	0.001
	−5	0.025	0.000	0.006	0.007	0.025	0.001	0.007	0.005	0.028	0.000	0.008	0.005
	−10	0.026	0.000	0.011	0.020	0.029	0.000	0.015	0.016	0.031	0.000	0.016	0.016
	−20	0.039	0.000	0.013	0.033	0.033	0.001	0.022	0.031	0.034	0.000	0.026	0.029
	−50	–	–	–	–	0.044	0.018	0.023	0.042	0.043	0.007	0.029	0.043
20	0	0.001	0.001	0.001	0.001	0.002	0.000	0.001	0.000	0.004	0.001	0.001	0.000
	−5	0.019	0.000	0.005	0.010	0.021	0.000	0.007	0.005	0.027	0.000	0.008	0.005
	−10	0.025	0.000	0.010	0.020	0.024	0.000	0.012	0.015	0.028	0.000	0.016	0.016
	−20	0.045	0.000	0.010	0.027	0.032	0.000	0.019	0.031	0.033	0.000	0.024	0.028
	−50	–	–	–	–	0.046	0.006	0.019	0.039	0.044	0.003	0.028	0.044
50	0	0.007	0.000	0.003	0.018	0.001	0.000	0.001	0.001	0.002	0.001	0.000	0.000
	−5	0.019	0.000	0.007	0.023	0.015	0.000	0.006	0.008	0.020	0.000	0.007	0.005
	−10	0.039	0.000	0.006	0.017	0.023	0.000	0.008	0.018	0.022	0.000	0.014	0.016
	−20	0.046	0.000	0.005	0.013	0.039	0.000	0.010	0.030	0.035	0.000	0.019	0.026
	−50	–	–	–	–	0.045	0.000	0.009	0.023	0.046	0.000	0.021	0.038
		Right-tail tests ($H_0 : \beta_h = 0$ vs $H_a : \beta_h > 0$) and $\phi = -0.95$											
5	0	0.212	0.271	0.118	0.073	0.221	0.225	0.105	0.066	0.240	0.216	0.103	0.060
	−5	0.118	0.388	0.104	0.078	0.116	0.419	0.103	0.067	0.123	0.431	0.101	0.059
	−10	0.094	0.413	0.088	0.077	0.091	0.521	0.090	0.070	0.095	0.575	0.089	0.061
	−20	0.078	0.423	0.062	0.077	0.073	0.610	0.074	0.073	0.075	0.698	0.076	0.060
	−50	–	–	–	–	0.065	0.584	0.051	0.070	0.064	0.724	0.058	0.062
10	0	0.242	0.291	0.121	0.066	0.216	0.258	0.104	0.064	0.225	0.251	0.103	0.058
	−5	0.130	0.347	0.110	0.068	0.118	0.464	0.101	0.066	0.117	0.501	0.104	0.061
	−10	0.098	0.309	0.087	0.065	0.096	0.559	0.089	0.071	0.093	0.648	0.089	0.060
	−20	0.079	0.188	0.062	0.055	0.079	0.616	0.072	0.071	0.076	0.765	0.076	0.062
	−50	–	–	–	–	0.066	0.478	0.049	0.065	0.066	0.759	0.058	0.064
20	0	0.309	0.220	0.118	0.053	0.242	0.262	0.106	0.061	0.220	0.267	0.102	0.056
	−5	0.133	0.162	0.114	0.048	0.135	0.430	0.105	0.061	0.121	0.516	0.103	0.058
	−10	0.101	0.073	0.096	0.042	0.102	0.470	0.091	0.061	0.098	0.645	0.093	0.057
	−20	0.078	0.009	0.071	0.029	0.079	0.412	0.076	0.061	0.081	0.736	0.079	0.061
	−50	–	–	–	–	0.066	0.109	0.047	0.045	0.069	0.646	0.059	0.057
50	0	0.385	0.044	0.076	0.052	0.329	0.200	0.111	0.044	0.265	0.255	0.100	0.047
	−5	0.135	0.022	0.126	0.032	0.146	0.208	0.110	0.041	0.143	0.431	0.106	0.046
	−10	0.098	0.003	0.125	0.024	0.103	0.129	0.098	0.038	0.108	0.472	0.098	0.046
	−20	0.079	0.000	0.092	0.017	0.080	0.036	0.081	0.031	0.083	0.388	0.088	0.047
	−50	–	–	–	–	0.066	0.001	0.049	0.021	0.067	0.089	0.063	0.038

(*continued on next page*)

Table 3 (*continued*).

Two-sided tests ($H_0 : \beta_h = 0$ vs $H_a : \beta_h \neq 0$) and $\phi = -0.15$

h	c	t_h^{Xu}	t_h^{Bonf}	$t_{h,ivx}^{rev,PL}$	$t_{h,ivx}^{trf,res}$	t_h^{Xu}	t_h^{Bonf}	$t_{h,ivx}^{rev,PL}$	$t_{h,ivx}^{trf,res}$	t_h^{Xu}	t_h^{Bonf}	$t_{h,ivx}^{rev,PL}$	$t_{h,ivx}^{trf,res}$
		$T = 100$				$T = 250$				$T = 500$			
5	0	0.054	0.033	0.022	0.021	0.052	0.034	0.018	0.018	0.049	0.038	0.018	0.018
	−5	0.066	0.020	0.043	0.043	0.057	0.032	0.039	0.038	0.051	0.038	0.037	0.037
	−10	0.068	0.013	0.041	0.040	0.059	0.027	0.042	0.044	0.052	0.038	0.043	0.042
	−20	0.069	0.004	0.034	0.035	0.058	0.020	0.043	0.042	0.054	0.035	0.044	0.045
	−50	–	–	–	–	0.056	0.008	0.034	0.035	0.056	0.022	0.040	0.042
10	0	0.056	0.033	0.023	0.022	0.050	0.031	0.018	0.019	0.049	0.036	0.017	0.017
	−5	0.080	0.020	0.043	0.041	0.063	0.027	0.037	0.037	0.054	0.033	0.036	0.035
	−10	0.079	0.009	0.044	0.043	0.065	0.021	0.040	0.041	0.054	0.029	0.041	0.041
	−20	0.078	0.002	0.035	0.035	0.066	0.012	0.041	0.041	0.058	0.024	0.043	0.041
	−50	–	–	–	–	0.060	0.002	0.036	0.035	0.060	0.010	0.039	0.037
20	0	0.067	0.027	0.025	0.023	0.053	0.030	0.020	0.017	0.049	0.036	0.018	0.017
	−5	0.098	0.016	0.044	0.038	0.074	0.025	0.039	0.034	0.058	0.028	0.037	0.036
	−10	0.093	0.006	0.046	0.037	0.076	0.016	0.043	0.038	0.062	0.023	0.041	0.038
	−20	0.079	0.000	0.036	0.029	0.073	0.006	0.043	0.038	0.065	0.015	0.043	0.039
	−50	–	–	–	–	0.060	0.000	0.030	0.027	0.063	0.003	0.038	0.035
50	0	0.098	0.001	0.028	0.048	0.065	0.027	0.026	0.021	0.052	0.031	0.021	0.017
	−5	0.120	0.000	0.047	0.039	0.094	0.020	0.046	0.034	0.074	0.023	0.041	0.033
	−10	0.102	0.000	0.046	0.024	0.091	0.010	0.050	0.034	0.076	0.014	0.045	0.038
	−20	0.080	0.000	0.036	0.015	0.077	0.001	0.039	0.028	0.073	0.005	0.045	0.039
	−50	–	–	–	–	0.059	0.000	0.026	0.018	0.062	0.000	0.035	0.028

Two-sided tests ($H_0 : \beta_h = 0$ vs $H_a : \beta_h \neq 0$) and $\phi = -0.95$

h	c	t_h^{Xu}	t_h^{Bonf}	$t_{h,ivx}^{rev,PL}$	$t_{h,ivx}^{trf,res}$	t_h^{Xu}	t_h^{Bonf}	$t_{h,ivx}^{rev,PL}$	$t_{h,ivx}^{trf,res}$	t_h^{Xu}	t_h^{Bonf}	$t_{h,ivx}^{rev,PL}$	$t_{h,ivx}^{trf,res}$
		$T = 100$				$T = 250$				$T = 500$			
5	0	0.138	0.230	0.058	0.042	0.141	0.198	0.049	0.035	0.142	0.190	0.048	0.031
	−5	0.088	0.322	0.060	0.047	0.081	0.378	0.055	0.039	0.085	0.397	0.054	0.033
	−10	0.070	0.344	0.053	0.051	0.067	0.473	0.051	0.046	0.067	0.533	0.053	0.036
	−20	0.060	0.342	0.039	0.059	0.058	0.563	0.047	0.055	0.059	0.668	0.052	0.045
	−50	–	–	–	–	0.052	0.576	0.035	0.061	0.057	0.706	0.043	0.055
10	0	0.184	0.234	0.060	0.037	0.141	0.228	0.053	0.034	0.133	0.227	0.046	0.030
	−5	0.098	0.257	0.057	0.039	0.087	0.417	0.056	0.038	0.082	0.462	0.053	0.032
	−10	0.070	0.203	0.050	0.043	0.070	0.501	0.052	0.045	0.067	0.609	0.054	0.037
	−20	0.062	0.089	0.037	0.042	0.061	0.545	0.048	0.056	0.059	0.730	0.050	0.046
	−50	–	–	–	–	0.055	0.403	0.037	0.058	0.058	0.729	0.045	0.055
20	0	0.255	0.149	0.065	0.028	0.175	0.228	0.054	0.031	0.136	0.240	0.049	0.029
	−5	0.096	0.079	0.064	0.029	0.098	0.358	0.055	0.034	0.089	0.469	0.054	0.029
	−10	0.069	0.024	0.056	0.029	0.072	0.373	0.052	0.038	0.071	0.596	0.056	0.035
	−20	0.069	0.001	0.042	0.026	0.061	0.287	0.044	0.046	0.060	0.680	0.049	0.044
	−50	–	–	–	–	0.056	0.043	0.032	0.040	0.060	0.559	0.042	0.051
50	0	0.334	0.020	0.054	0.047	0.276	0.142	0.059	0.024	0.207	0.221	0.050	0.024
	−5	0.094	0.006	0.089	0.031	0.100	0.119	0.062	0.024	0.105	0.353	0.060	0.025
	−10	0.084	0.001	0.089	0.019	0.072	0.059	0.057	0.024	0.075	0.360	0.060	0.030
	−20	0.071	0.000	0.062	0.010	0.070	0.009	0.049	0.028	0.063	0.257	0.055	0.036
	−50	–	–	–	–	0.058	0.000	0.030	0.021	0.061	0.030	0.042	0.037

Notes: See Notes to Table 1.

$\phi = -0.15$, however when the correlation increases significant over-sizing is observed. For instance, when $\phi = -0.95$ the range of rejection frequencies are $[0.054, 0.151]$ for T=100, $[0.047, 0.114]$ for $T = 250$, and $[0.058, 0.107]$ for $T = 500$. In contrast the left-sided versions of these tests display conservative behaviour, which is a common characteristic of IVX-based predictability tests; see, for example, Demetrescu et al. (2022a). In general, however, the degree of undersizing observed in the left-tailed IVX-based tests is less pronounced, often very significantly so, for $t_{h,ivx}^{trf,res}$ than it is for $t_{h,ivx}^{rev,PL}$.

In contrast to the IVX-based tests, the empirical rejection frequencies of the t_h^{Xu} test are very sensitive to the strength of the persistence of the predictor (and magnitude of the correlation ϕ). For example, in Table 1 it can be seen that when $\phi = -0.15$ t_h^{Xu} displays decent size performance, but when the endogeneity correlation increases (increasing the relevance of the strength of the persistence of the predictor on the performance of the test statistics) the test displays substantial size distortions (eg for $\phi = -0.95$ these occur regardless of the sample size and for both one-sided and two-sided implementations of the test). The finite sample behaviour of the one-sided and two-sided t_h^{Xu} tests become generally more erratic when the innovations v_{t+1} are autocorrelated, and are particularly unreliable in the case of negatively autocorrelated v_{t+1}; see Table 3. We recall from the discussion in Section 3.2 that t_h^{Xu} is not valid when v_{t+1} is autocorrelated and these results illustrate this well.

The t_h^{Bonf} tests display empirical rejection frequencies close to the nominal 5% significance level, for both one-sided and two-sided implementations, in the case where v_{t+1} is serially uncorrelated (Table 1), $c \geq -20$ and $h < 20$ and regardless of the sample size (except for $c = -20$ and $T = 100$ where the test is under-sized). As discussed in Section 3.1, this test is based on the assumption that the predictor is strongly persistent and so the deterioration in the empirical rejection rates for $c = -50$ is to be expected. Perhaps most striking, however, is the highly erratic behaviour of t_h^{Bonf} when the innovations v_{t+1} are autocorrelated (Tables 2 and 3). Here the t_h^{Bonf} tests can be either massively over-sized, with size sometimes in excess of 50%, or massively under-sized. On the basis of these results this approach would appear to be too unreliable to use in empirical applications.

We conclude from the results in Tables 1–3 (see also the additional results in Tables S.1–S.3 in the Supplementary Appendix)[13] that only the IVX-based long-horizon predictability tests, $t_{h,ivx}^{trf,res}$ and $t_{h,ivx}^{rev,PL}$, display reliable enough finite sample size control across predictors whose degree of persistence is unknown and which are not driven by uncorrelated innovations to be empirically useful. The t_h^{Bonf} and t_h^{Xu} tests would appear to be too unreliable to be used in practical applications. Of the $t_{h,ivx}^{trf,res}$ and $t_{h,ivx}^{rev,PL}$ tests our results suggest that the former delivers significantly better finite sample size control.

6.3. *Empirical power*

In this section we compare the finite sample power properties of the $t_{h,ivx}^{trf,res}$ and $t_{h,ivx}^{rev,PL}$ tests. (Again, $t_{h,ivx}^{trf}$ and $t_{h,ivx}^{rev,PL}$ perform very similarly and we only report $t_{h,ivx}^{rev,PL}$.) Because of the unreliable size properties of the t_h^{Bonf} and t_h^{Xu} tests reported in Section 6.2 we will not include these tests in our main discussion, however results for these tests can be found in the Supplementary Appendix (see Figures S.31 - S.55). To investigate the finite sample power properties of the $t_{h,ivx}^{trf,res}$ and $t_{h,ivx}^{rev,PL}$ tests we simulate data from (2.2)–(2.1) under the alternative hypothesis $H_1 := b/T$, across the following values of the drift parameter, $b \in \{-15, -14.5, -14, \ldots, 14, 14.5, 15\}$. The innovations $(u_{t+1}, \varpi_{t+1})'$ were generated as described in Section 6.2 with results reported in Figs. 1–2 only for $\psi = 0$; results for $\psi \in \{-0.5, 0.5\}$ are qualitatively similar and can be found in the Supplementary Appendix (Figures S.1–S.30). Figures S.1–S.30 report left-, right- and two-sided test results for $\phi = \{-0.95, -0.50, -0.15\}$ (cf. footnote 12), for prediction horizons $h = \{1, 5, 10, 20, 50\}$ and for five values of the persistence parameter, c, associated with x_t; specifically, $c = \{0, -5, -10, -20\}$. In the interests of space, Figs. 1–2 only present power curves for one-sided tests (left- and right-sided) for prediction horizons $h = 5, 20$, sample sizes $T = 100$ and $T = 250$, $\phi = \{-0.15, -0.95\}$ and noncentrality parameters $c = -5$ and $c = -20$. Consider first Fig. 1 which plots the power curves of the left-sided $t_{h,ivx}^{trf,res}$ and $t_{h,ivx}^{rev,PL}$ tests against $H_1 : \beta_h < 0$. It is clearly seen from these figures that when ϕ is small in absolute value ($\phi = -0.15$), that the left-sided $t_{h,ivx}^{trf,res}$ and $t_{h,ivx}^{rev,PL}$ tests display similar performance. Moreover this figure also illustrates that when ϕ is large in absolute value ($\phi = -0.95$) the left-sided $t_{h,ivx}^{trf,res}$ test displays significantly superior power performance than the left-sided $t_{h,ivx}^{rev,PL}$ test and that this holds regardless of the prediction horizon or the strength of persistence of the predictor. It can also be seen from Fig. 1 that for both tests power decreases as c decreases (i.e. as the persistence of the predictor weakens), other things being equal. This pattern is to be expected as the signal from the predictor becomes stronger the more persistent is the predictor, x_t. Finally, we observe that the power superiority of $t_{h,ivx}^{trf,res}$ over $t_{h,ivx}^{rev,PL}$ generally becomes more pronounced as h becomes larger, other things equal.

Turning to the right-sided tests in Fig. 2 we observe that also in this case, when ϕ is small ($\phi = -0.15$) both tests display suitable size performance, but as the impact of endogeneity increases the performance of $t_{h,ivx}^{rev,PL}$ deteriorates. For instance, for $\phi = -0.95$ results seem to suggest that the $t_{h,ivx}^{rev,PL}$ test displays somewhat higher empirical rejection

[13] Tables S.1–S.3 present the empirical rejection frequencies for DGPs with homoskedastic IID innovations (Table S.1), DGPs with positively autocorrelated innovations (Table S.2) and DGPs with negatively autocorrelated innovations (Table S.3). Each of the three tables in each case present results for one of the values of ϕ considered, $\phi = \{-0.15, -0.50, -0.95\}$. All tables present results for three sample sizes: $T = 100, 250$ and 500.

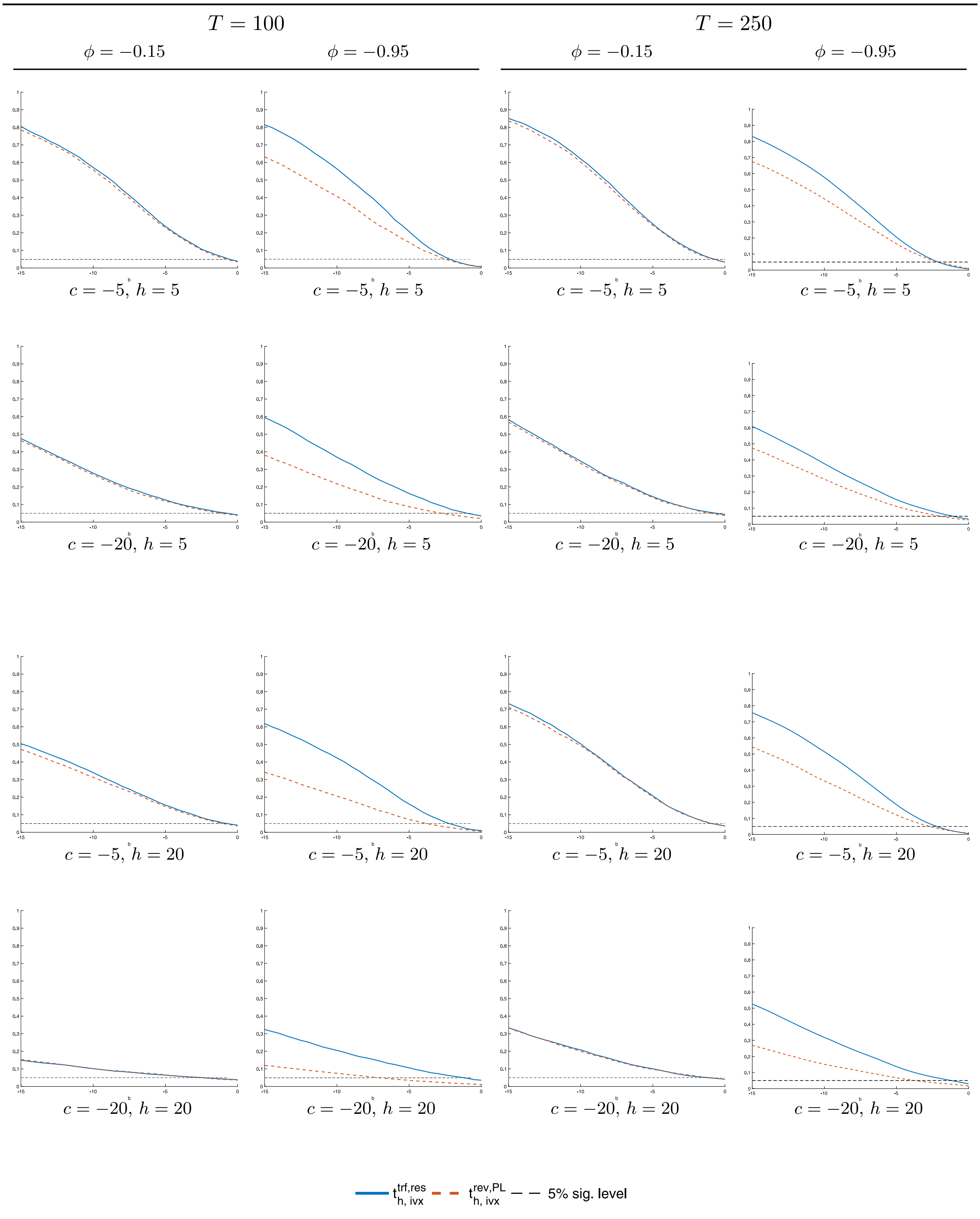

Fig. 1. Power curves of the **LEFT**-sided tests $t_{h,ivx}^{trf,res}$ and $t_{h,ivx}^{rev,PL}$ for prediction horizon $h = \{5, 20\}$ and $T = \{100, 250\}$. **DGP:** $y_{t+1} = \beta x_t + u_{t+1}$, $x_{t+1} = \rho x_t + v_{t+1}$ and $v_{t+1} = \psi v_t + \varpi_{t+1}$, where $\beta = b/T$, $\rho = 1 + c/T$, with $c = \{-5, -20\}$, $\psi = 0.5$ and $(u_{t+1}, \varpi_{t+1})' \sim NIID(\mathbf{0}, \mathbf{\Sigma})$, with $\mathbf{\Sigma} = [1 \quad \phi; \quad \phi \quad 1]$, and $\phi = \{-0.15, -0.95\}$.

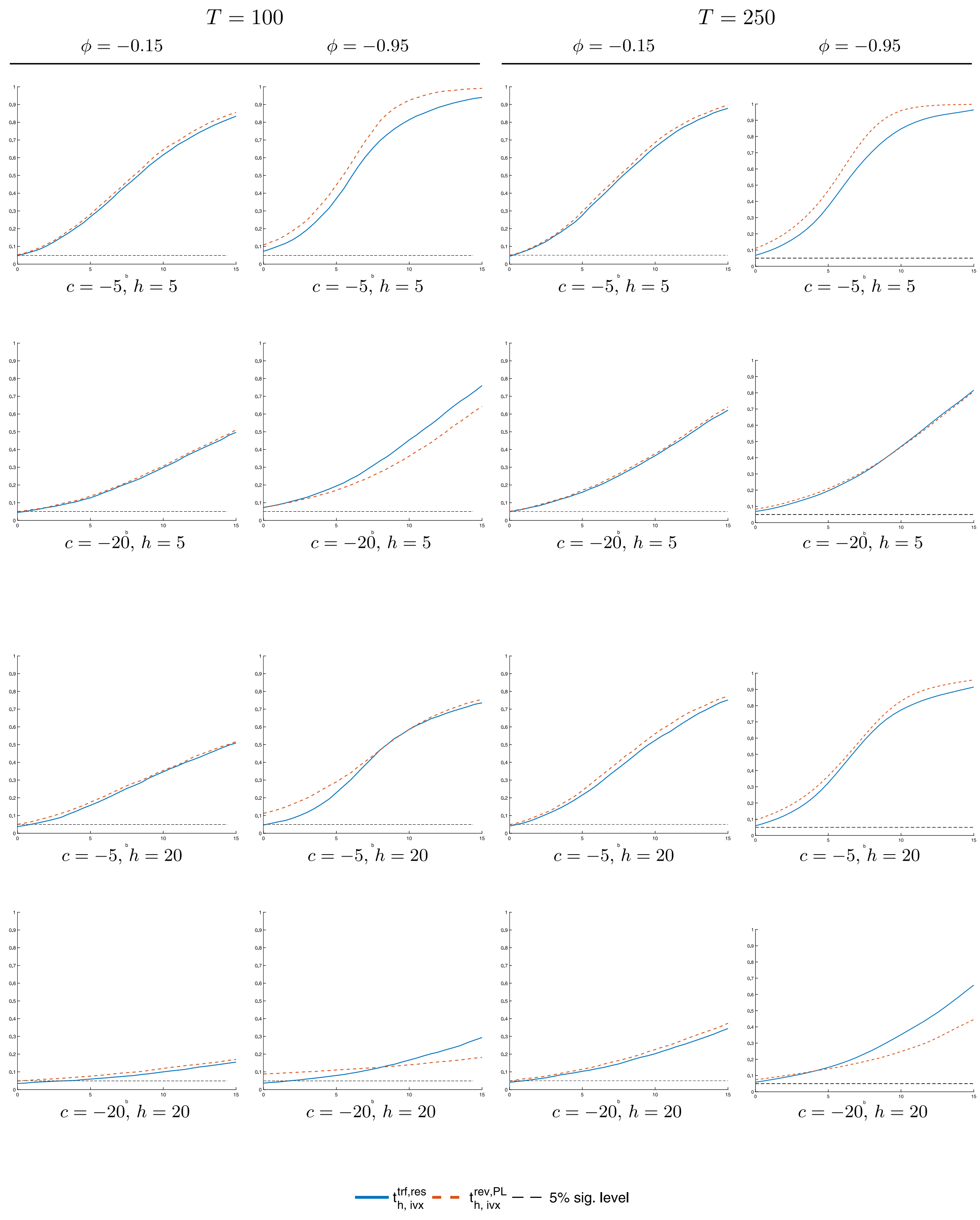

Fig. 2. Power curves of the **RIGHT**-sided tests $t^{trf,res}_{h,ivx}$ and $t^{rev,PL}_{h,ivx}$ for prediction horizon $h = \{5, 20\}$ and $T = \{100, 250\}$. **DGP:** $y_{t+1} = \beta x_t + u_{t+1}$, $x_{t+1} = \rho x_t + v_{t+1}$ and $v_{t+1} = \psi v_t + \varpi_{t+1}$, where $\beta = b/T$, $\rho = 1 + c/T$, with $c = \{-5, -20\}$, $\psi = 0.5$ and $(u_{t+1}, \varpi_{t+1})' \sim NIID(\mathbf{0}, \mathbf{\Sigma})$, with $\mathbf{\Sigma} = [1 \quad \phi; \quad \phi \quad 1]$, and $\phi = \{-0.15, -0.95\}$.

frequencies than $t^{trf,res}_{h,ivx}$. However, this is an artifact of the significant over-sizing seen with the $t^{rev,PL}_{h,ivx}$ test in these scenarios; see Tables 1–3 and which is also visible in these plots. Indeed, when we compare the power properties of the two tests for $c \leq -10$ and $h > 10$ where their empirical sizes are broadly comparable, we observe that $t^{trf,res}_{h,ivx}$ tends to display superior power to $t^{rev,PL}_{h,ivx}$. Again, as h becomes larger $t^{trf,res}_{h,ivx}$ tends to perform better than $t^{rev,PL}_{h,ivx}$; for example for $h = 50$ we see that $t^{trf,res}_{h,ivx}$ is generally more powerful than $t^{rev,PL}_{h,ivx}$ for $c \leq -5$ even though the latter is rather over-sized for $c = -5$, $c = -10$ and $c = -20$ (see Figures S.14 and S.29, in the Supplementary Appendix).

7. Empirical application

Exchange rate predictability has been a topic of considerable interest in the international finance and macroeconomics literatures. We revisit the recent study of Eichenbaum et al. (2020) [henceforth EJR] who document: (i) that current real exchange rates (*RER*) predict nominal exchange rates (*NER*) in the long-run[14]; (ii) that *RER* is a poor predictor of future inflation rates, and (iii) that these regularities depend on the monetary policy regime in effect. EJR further observe that current *RER* is strongly negatively correlated with future changes in *NER*, that this correlation increases with the prediction horizon, and that *RER* is virtually uncorrelated with future inflation rates at all horizons. These empirical observations suggest that *RER* adjusts to shocks in the medium and long run overwhelmingly through changes in *NER*, and not through inflation rate differentials.

EJR base their analysis on a benchmark group of six countries (Australia, Canada, Germany, New Zealand, Sweden, and the UK), which (other than Germany) had adopted inflation targeting before 1997.[15] We revisit the predictive power of *RER* for predicting changes in *NER* and future inflation rates across 45 countries. Our contribution to this literature is to provide further evidence on the stylised features of exchange rate predictability using the new long-horizon predictability tests developed in this paper to evaluate the usefulness of current *RER*s as predictors of future changes in *NER*s and inflation differentials.

7.1. Data

All data used in the empirical analysis is obtained from the International Financial Statistics of the IMF (https://data.imf.org) for the period from 1973:Q1 to 2020:Q1. The analysis will be conducted over four different sample periods: (i) the full sample — 1973:Q1 to 2020:Q1; (ii) from 1973:Q1 to 2008:Q4; (iii) from 1990:Q1 to 2008:Q4; and (iv) from 1999:Q1 to 2020:Q1. The sample includes 45 countries split into developed and emerging markets according to the MCSI classification; see https://www.msci.com/market-classification. The developed markets group comprises Australia, Austria, Belgium, Canada, Denmark, Finland, France, Germany, Hong Kong, Ireland, Israel, Italy, Japan, Luxembourg, Netherlands, New Zealand, Norway, Portugal, Singapore, Spain, Sweden, Switzerland, and United Kingdom. The emerging markets group consists of Brazil, Bulgaria, Chile, China, Colombia, Czech Republic, Egypt, Greece, Hungary, Iceland, India, Indonesia, Korea, Mexico, Peru, Philippines, Poland, Romania, Russian Federation, South Africa, Thailand, and Ukraine. Most of these countries adopted inflation targeting, but at a later stage than the benchmark group considered in EJR (many adopted this policy in 1999 and a few between 1999 and 2005); see Ilzetzki et al. (2017) for details.

Although the overall sample period is from 1973:Q1 to 2020:Q1, the samples for some of the countries are slightly smaller due to lack of available data at the beginning and/or end of the sample. Specifically, for Hungary and Iceland the sample starts in 1976:Q1, for Brazil and Poland in 1980:Q1, for Hong-Kong in 1980:Q4, for China in 1986:Q1, for Romania in 1990:Q4, for Bulgaria in 1991:Q1, for the Czech Republic and the Ukraine in 1993:Q1 and finally for the Russian Federation in 1995:Q2. Moreover, for Egypt and the Ukraine the ending dates are also shorter than for the rest of the countries in the sample (2019:Q3 and 2019:Q4, respectively).

7.2. Empirical results

7.2.1. The nominal exchange rate long-horizon predictive regression

The *NER* long-horizon predictive regression considered by EJR is given by

$$log\left(\frac{NER_{i,t+h}}{NER_{it}}\right) = \alpha^{NER}_{ih} + \beta^{NER}_{ih} log(RER_{it}) + u^{NER}_{i,t+h}, \tag{7.1}$$

where i corresponds to the country under analysis and h to the prediction horizon (in quarters), $h = \{1, 4, 8, 12, 20\}$. The predictor is the real exchange rate of country i relative to the US, i.e., $RER_{it} := NER_{it}(P_{it}/P_t)$, where NER_{it} is the average quarterly nominal exchange rate (domestic currency per US dollar) and P_t and P_{it} denote the consumer price index (CPI) for all items in the US and in country i, respectively.

[14] Mark (1995) and Engel et al. (2007) have also found evidence of predictability of *NER* at medium and long horizons; see Rossi (2013) for a survey.

[15] In EJR the sample periods for Australia, Canada, Germany, New Zealand, Sweden, and the U.K. start in 1993:Q3, 1991:Q2, 1982:Q4, 1990:Q1, 1996:Q1, and 1992:Q4, respectively. All samples end in 2008:Q4, because from 2009 to the present short-term U.S. nominal interest rates were at or near their effective lower bound (however EJR also provide results for the samples ending in 2018 in a supplementary appendix).

To provide an indication of the persistence of RER_{it}, we estimate the augmented Dickey–Fuller regression for each country,

$$RER_{it} = \alpha_i^{RER} + \rho_i^{RER} RER_{i,t-1} + \sum_{k=1}^{p} \delta_k \Delta RER_{i,t-k} + \varpi_{it}^{RER}, \; i = 1, \ldots, 45, \tag{7.2}$$

where for each series the lag order p is determined based on the AIC information criteria with a maximum lag order determined by the so-called Schwert's rule, $\lfloor 4(T/100)^{1/4} \rfloor$. We report the OLS estimates of ρ_i^{RER}, $\hat{\rho}_i^{RER}$, for each country under analysis, as well as estimates of the contemporaneous correlation, ϕ_i, between the innovations (under the assumption that the correlation is constant), specifically,

$$\hat{\phi}_i := \frac{(T-\hat{p}_i)^{-1} \sum_{t=\hat{p}_i}^{T-1} \hat{u}_{i,t+1}^{NER} \hat{\varpi}_{i,t+1}^{RER}}{\sqrt{((T-1)^{-1} \sum_{t=1}^{T-1} (\hat{u}_{i,t+1}^{NER})^2)((T-\hat{p}_i)^{-1} \sum_{t=\hat{p}_i}^{T-1} (\hat{\varpi}_{i,t+1}^{NER})^2)}}, \tag{7.3}$$

where $\hat{u}_{i,t+1}^{NER}$ are the OLS residuals from the predictive regression in (7.1) with $h = 1$, and the OLS residuals $\hat{\varpi}_{i,t+1}^{RER}$ from (7.2). EJR assume that RER is mean reverting (weakly persistent) and highlight a number of features they observe from the estimation of (7.1) by OLS. Their analysis is based on testing for long-horizon predictability by comparing the conventional OLS t-statistic from (7.1) computed with Newey–West standard errors, denoted $t_{h,NW}$, with critical values from the standard normal distribution. As is well known and discussed in Section 2.2 these tests are not theoretically valid and likely to spuriously reject the null hypothesis if RER is strongly persistent.

The estimates of $\hat{\rho}_i^{RER}$ reported in Table 4 (and Tables S.39 and S.40 in the Supplementary Appendix) suggest that for most of the countries considered RER is strongly persistent with an estimated autoregressive root very close to unity. From Panel A of Table 4 we observe that, in general, for all countries $\hat{\rho}^{RER} \geq 0.953$ when considering the sample from 1973:Q1 to 2020:Q1 (except for the Russian Federation, where $\hat{\rho}^{RER} = 0.898$); $\hat{\rho}^{RER} \geq 0.932$ in the sample from 1973:Q1 to 2008:Q4 (except for the Russian Federation and Ukraine, where $\hat{\rho}^{RER} = 0.868$ and $\hat{\rho}^{RER} = 0.853$, respectively; see Table S.39 in the Supplementary Appendix); $\hat{\rho}^{RER} \geq 0.910$ from 1990:Q1 to 2008:Q4 (except for Peru, the Russian Federation and Ukraine, where $\hat{\rho}^{RER} = 0.666$, $\hat{\rho}^{RER} = 0.868$ and $\hat{\rho}^{RER} = 0.853$, respectively; see Table S.40); and finally $\hat{\rho}^{RER} \geq 0.918$ from 1999:Q1 to 2020:Q1 (except for Korea where $\hat{\rho}^{RER} = 0.887$); see Panel B of Table 4.

In Table 4 (and Tables S.39 and S.40 of the Supplementary Appendix) we also report, for the various sample periods discussed above and for each horizon h, the results of the $t_{h,NW}$ test, of our new IVX-based $t_{h,ivx}^{trf,res}$ test and of the $t_{h,ivx}^{rev,PL}$ test of Phillips and Lee (2013). The IVX-based test was implemented exactly as detailed for the simulation study in Section 6. Although we provide results for $t_{h,NW}$, these should be treated with caution given the strong persistence of the predictor highlighted above. As suggested in EJR, the Newey–West standard errors used in $t_{h,NW}$ were computed using the Bartlett kernel setting the number of lags to $h + 8$.[16]

Consider first the results in Panels A and B of Table 4. Here we observe negative outcomes for the IVX-based statistics for almost all countries (the exceptions are a small number of emerging markets) and for all of the values of h considered. This entails that the IVX estimates of the β_{ih}^{NER} slope coefficients are negative, albeit many of these test outcomes are not statistically significant. These findings support EJR's conclusion that current RER and changes in future NERs are negatively correlated. The results in Tables S.39 - S.40 in the Supplementary Appendix suggest that this finding also appears robust to the other sample periods considered. In addition to the observation that the outcomes of the IVX-based statistics are mostly negative, we also observe that the estimated innovation correlations, $\hat{\phi}_i$, are positive for all of the countries and are generally very high. As the Monte Carlo simulation results in Section 6.2 show (see footnote 12), this is precisely the case where the left-sided $t_{h,ivx}^{rev,PL}$ test will be significantly oversized, while our preferred residual-augmented $t_{h,ivx}^{trf,res}$ test is approximately correctly sized. We might therefore expect to see fewer rejections with the $t_{h,ivx}^{trf,res}$ test than with the $t_{h,ivx}^{rev,PL}$ test, and that should be borne in mind in what follows.

Overall, the results in Table 4 provide increasing evidence of predictability as h increases. This is particularly, noticeable in the top part of Panel A which contains the results for the developed markets nations, where an increase in the number of statistically significant cases is observed for larger h. However, we also note that the number of rejections is largest for $t_{h,NW}$ and smallest for $t_{h,ivx}^{trf,res}$. This is unsurprising given that, as discussed above, the former is likely to be invalid for these data and that the latter is the only one of the tests reported which displays reliable size control in this setting. In the case of the emerging markets nations, a similar situation as for the developed markets nations can be observed from the results for the $t_{h,ivx}^{trf,res}$ and $t_{h,ivx}^{rev,PL}$ tests. $t_{h,NW}$ finds that changes in NER in more than 50% of these countries are predictable by RER when $h = 1$, but as h increases the number of statistically significant results decreases slightly. From the results of $t_{h,ivx}^{trf,res}$ and $t_{h,ivx}^{rev,PL}$ we observe that for forecast horizons $h \geq 8$ predictability seems to increase ($h = 20$ displays the largest number of significant cases). The results for $t_{h,ivx}^{trf,res}$ and $t_{h,ivx}^{rev,PL}$ in Panel A of Table 4 suggest that in the full sample (1973:Q1 to 2020:Q1), of the benchmark countries considered by EJR, only Canada seems to become significant when $h \geq 12$, whereas based on $t_{h,NW}$Australia, New Zealand and the UK display statistically significant results.

---

[16] For all but one of the countries considered the fitted lag length, $\hat{p}_i$, from (7.2) was greater than zero in all of the sample periods considered. For that reason, we do not report results for the t_h^{Bonf} test of Hjalmarsson (2011) or the t_h^{Xu} test of Xu (2020) given their likely unreliability in such cases; see Section 6.2.

Table 4
Nominal exchange rate long-horizon predictive regression results.

	$\hat{\phi}$	$\hat{\rho}^{RER}$	$h=1$			$h=4$			$h=8$			$h=12$			$h=20$		
			$t_{h,NW}$	$t^{trf,res}_{h,ivx}$	$t^{rev,PL}_{h,ivx}$	$t_{h,NW}$	$t^{trf,res}_{h,ivx}$	$t^{rev,PL}_{h,ivx}$	$t_{h,NW}$	$t^{trf,res}_{h,ivx}$	$t^{rev,PL}_{h,ivx}$	$t_{h,NW}$	$t^{trf,res}_{h,ivx}$	$t^{rev,PL}_{h,ivx}$	$t_{h,NW}$	$t^{trf,res}_{h,ivx}$	$t^{rev,PL}_{h,ivx}$
PANEL A: Period from 1973:Q1 to 2020:Q2.																	
Australia	0.933	0.973	−1.653*	−1.301	−0.750	−2.436**	−1.390	−1.053	−2.736***	−1.325	−0.912	−2.943***	−1.273	−0.801	−2.986***	−1.129	−0.705
Austria	0.958	0.975	−0.575	−0.256	−0.795	−0.758	−0.326	−1.191	−0.794	−0.280	−1.145	−0.888	−0.200	−1.258	−0.989	−0.103	−0.859
Belgium	0.970	0.968	−0.081	−1.528	−1.523	−0.199	−1.915*	−2.375**	−0.217	−1.832*	−2.767***	−0.268	−1.637	−3.089***	−0.228	−1.268	−2.642***
Canada	0.951	0.973	−0.178	−0.544	−0.920	−0.400	−0.644	−1.434	−0.465	−0.675	−1.456	−0.572	−0.739	−1.808*	−0.715	−0.783	−1.790*
Denmark	0.952	0.965	0.044	−0.936	−1.103	−0.086	−1.161	−1.609	−0.111	−1.180	−1.926*	−0.147	−1.139	−1.964**	−0.128	−1.038	−1.952*
Finland	0.934	0.953	0.288	−0.304	−0.798	0.201	−0.444	−1.274	0.180	−0.462	−1.573	0.186	−0.415	−1.436	0.297	−0.319	−1.094
France	0.962	0.970	0.197	−0.718	−0.805	0.036	−0.952	−1.195	0.006	−0.985	−1.552	−0.012	−0.963	−1.576	0.006	−0.847	−1.516
Germany	0.974	0.978	−0.714	−0.358	−0.590	−1.031	−0.432	−0.963	−1.148	−0.383	−0.968	−1.340	−0.306	−1.089	−1.587	−0.220	−0.720
Hong Kong	0.334	0.969	1.222	−0.875	−1.920*	1.145	−1.345	−1.849*	1.120	−1.765*	−1.575	1.100	−1.914*	0.544	1.060	−1.789*	0.196
Ireland	0.910	0.964	−1.975**	−1.032	−0.690	−2.410**	−1.145	−0.968	−2.887***	−1.142	−1.084	−2.988***	−1.108	−0.843	−2.991***	−0.937	−0.723
Israel	0.862	0.996	−3.068***	−1.205	−0.882	−2.897***	−1.567	−1.480	−2.773***	−1.990**	−2.012**	−2.862***	−2.350**	−2.621***	−3.755***	−2.937***	−3.563***
Italy	0.862	0.977	1.134	−0.235	−0.667	1.105	−0.267	−0.735	1.095	−0.266	−0.893	1.079	−0.242	−0.557	1.110	−0.170	−0.596
Japan	0.961	0.992	−1.195	−0.495	−0.748	−1.377	−0.579	−1.051	−1.654*	−0.620	−1.221	−1.932*	−0.596	−1.297	−2.189**	−0.498	−0.645
Luxembourg	0.966	0.967	−0.081	−1.206	−1.461	−0.201	−1.505	−2.299**	−0.219	−1.442	−2.661***	−0.271	−1.292	−2.962***	−0.230	−1.004	−2.532**
Netherlands	0.972	0.979	−0.666	−0.669	−0.819	−0.889	−0.770	−1.213	−0.951	−0.690	−1.189	−1.082	−0.572	−1.334	−1.237	−0.426	−0.959
New Zealand	0.896	0.974	−1.492	−0.977	−0.683	−2.071**	−1.069	−1.013	−2.316**	−1.051	−0.918	−2.624***	−1.033	−0.585	−2.702***	−0.950	−0.632
Norway	0.952	0.975	0.623	−0.477	−1.086	0.504	−0.610	−1.551	0.517	−0.574	−1.722*	0.500	−0.553	−1.649*	0.590	−0.429	−1.604
Portugal	0.650	0.987	1.430	−0.536	−0.832	1.387	−0.629	−0.949	1.335	−0.679	−1.197	1.280	−0.687	−1.064	1.229	−0.636	−0.809
Singapore	0.866	0.992	−1.541	−0.302	−0.033	−1.958*	−0.293	−0.215	−1.963**	−0.278	−0.084	−2.271**	−0.266	−0.466	−2.661***	−0.255	−0.469
Spain	0.864	0.980	0.883	−0.099	−0.585	0.864	−0.144	−0.827	0.863	−0.143	−1.027	0.861	−0.121	−0.822	0.911	−0.071	−0.644
Sweden	0.941	0.977	0.782	−0.523	−0.880	0.721	−0.488	−1.210	0.751	−0.435	−1.424	0.777	−0.371	−1.252	0.916	−0.287	−1.091
Switzerland	0.977	0.985	−1.767*	−1.089	−0.766	−2.279**	−1.059	−0.986	−2.441**	−0.879	−0.791	−2.768***	−0.680	−0.774	−3.519***	−0.386	−0.397
UK	0.928	0.955	−1.522	−0.407	−0.688	−1.724*	−0.383	−0.864	−2.098**	−0.311	−0.917	−2.419**	−0.236	−0.644	−2.494**	−0.109	−0.445
Brazil	0.888	0.997	−3.492***	−1.276	−0.884	−3.300***	−1.566	−1.323	−3.169***	−1.871*	−1.863*	−3.173***	−2.138**	−2.371**	−3.626***	−2.575**	−3.198***
Bulgaria	0.947	0.973	−2.069**	−0.902	−0.664	−1.971**	−1.459	−1.202	−2.563**	−1.925*	−1.679*	−3.839***	−2.197**	−1.628	−19.090***	−2.743***	−2.340**
Chile	0.277	0.977	−0.798	1.088	−0.520	−0.665	1.559	−0.542	−0.593	1.811*	−0.145	−0.487	1.147	0.229	−0.297	7.092***	0.346
China	0.741	0.966	0.890	−0.165	−0.589	0.790	−0.095	−0.377	0.652	−0.041	−0.584	0.559	−0.007	−0.774	0.379	−0.005	−0.541
Colombia	0.468	0.996	2.616***	−0.432	−0.850	2.380**	−0.500	−0.972	2.136**	−0.567	−0.988	1.978**	−0.587	−1.072	1.764*	−0.573	−1.402
Czech Rep.	0.946	0.973	−0.473	−1.398	−0.751	−0.527	−0.813	−1.032	−0.495	−0.752	−1.239	−0.491	−0.752	−1.511	−0.589	−0.964	−2.404**
Egypt	0.782	1.000	−0.069	−0.017	0.000	−0.168	−0.125	−0.165	−0.311	−0.158	−0.365	−0.436	−0.205	−0.497	−1.010	−0.320	−0.695
Greece	0.646	0.991	1.902*	−0.028	−0.657	1.856*	−0.036	−0.827	1.740*	−0.059	−1.019	1.638	−0.071	−0.908	1.514	−0.059	−1.212
Hungary	0.803	0.998	1.855*	0.485	−0.015	1.742*	0.473	−0.172	1.578	0.418	−0.398	1.481	0.368	−0.699	1.451	0.323	−1.125
Iceland	0.522	0.988	0.729	−0.575	−1.024	0.472	−0.755	−1.352	0.174	−0.887	−1.474	−0.010	−0.957	−1.442	−0.137	−0.903	−1.216
India	0.754	1.000	3.394***	0.225	−0.392	3.218***	0.184	−0.495	3.040***	0.177	−0.600	2.833***	0.115	−0.556	2.637***	0.049	−1.046
Indonesia	0.919	0.997	2.319**	0.037	−0.483	2.507**	0.063	−0.671	2.558**	0.145	−0.694	2.579***	0.140	−0.732	2.597***	0.104	−0.912
Korea	0.888	0.978	1.322	0.076	−0.428	1.438	0.064	−0.621	1.449	0.037	−0.426	1.473	0.037	−0.422	1.623	0.028	−0.466
Mexico	0.719	0.997	−2.184**	−1.392	−0.962	−2.254**	−1.611	−1.362	−2.307**	−1.780*	−1.821*	−2.384**	−1.877*	−2.224**	−2.586***	−1.983**	−2.332**
Peru	0.933	0.997	−2.704***	−0.724	−0.342	−2.526**	−1.025	−0.776	−2.414**	−1.347	−1.397	−2.359**	−1.659*	−2.070**	−2.417**	−2.027**	−2.893***
Philippines	0.726	0.996	1.880*	0.010	−0.546	1.814*	−0.062	−0.817	1.749*	−0.068	−0.925	1.647*	−0.102	−0.982	1.495	−0.140	−1.264
Poland	0.900	0.992	−2.463**	−1.275	−0.899	−2.327**	−1.629	−1.586	−2.340**	−1.864*	−1.682*	−2.453**	−2.029**	−2.189**	−3.408***	−2.205**	−2.773***
Romania	0.633	0.961	−8.930***	−2.761***	−2.216**	−11.361***	−2.895***	−1.211	−14.737***	−2.763***	−0.685	−14.191***	−2.489**	−0.343	−12.696***	−1.586	−0.409
Russian Fed.	0.048	0.898	0.113	−0.454	−0.493	0.531	−0.270	−0.117	0.669	−0.060	−0.146	0.605	−0.102	−0.167	0.121	−0.797	−0.256
South Africa	0.808	0.997	0.350	−0.014	−0.352	0.175	−0.080	−0.473	0.037	−0.119	−0.553	0.017	−0.103	−0.350	−0.043	−0.045	−0.479
Thailand	0.921	0.988	0.651	−0.316	−0.318	0.654	−0.241	−0.662	0.707	−0.112	−0.668	0.743	−0.109	−0.705	0.844	−0.196	−1.019
Ukraine	0.414	0.980	−1.123	1.651*	−0.801	−1.091	1.942*	−0.622	−0.983	1.067	−0.209	−0.815	8.111***	−0.075	−0.732	34.635***	−0.208

(*continued on next page*)

Table 4 (*continued*).

	$\hat{\phi}$	$\hat{\rho}^{RER}$	$h=1$			$h=4$			$h=8$			$h=12$			$h=20$		
			$t_{h,NW}$	$t_{h,ivx}^{trf,res}$	$t_{h,ivx}^{rev,PL}$	$t_{h,NW}$	$t_{h,ivx}^{trf,res}$	$t_{h,ivx}^{rev,PL}$	$t_{h,NW}$	$t_{h,ivx}^{trf,res}$	$t_{h,ivx}^{rev,PL}$	$t_{h,NW}$	$t_{h,ivx}^{trf,res}$	$t_{h,ivx}^{rev,PL}$	$t_{h,NW}$	$t_{h,ivx}^{trf,res}$	$t_{h,ivx}^{rev,PL}$
PANEL B: Period from 1999:Q1 to 2020:Q1.																	
Australia	0.973	0.958	−0.609	−1.775*	−0.692	−1.049	−2.550**	−1.086	−1.946*	−2.316**	−1.388	−3.393***	−2.126**	−1.437	−8.985***	−1.991**	−0.373
Austria	0.960	0.959	−0.169	−0.355	−0.532	−0.308	−0.543	−0.986	−0.603	−0.539	−1.155	−0.713	−0.447	−1.288	−0.533	−0.403	−0.267
Belgium	0.963	0.956	−0.148	−0.774	−0.554	−0.276	−0.495	−1.026	−0.563	−0.401	−1.175	−0.671	−0.268	−1.292	−0.491	−0.239	−0.257
Canada	0.984	0.964	−1.080	−2.216**	−0.457	−1.344	−2.551**	−0.623	−1.812*	−2.331**	−0.726	−2.913***	−2.303**	−0.959	−8.884***	−2.227**	−0.684
Denmark	0.961	0.956	−0.185	−0.815	−0.636	−0.339	−0.515	−1.108	−0.646	−0.423	−1.263	−0.755	−0.280	−1.371	−0.572	−0.274	−0.282
Finland	0.959	0.960	−0.211	−0.322	−0.537	−0.371	−0.474	−0.986	−0.682	−0.422	−1.155	−0.796	−0.317	−1.293	−0.618	−0.328	−0.261
France	0.963	0.957	−0.212	−0.723	−0.601	−0.368	−0.481	−1.044	−0.678	−0.427	−1.209	−0.790	−0.326	−1.345	−0.606	−0.299	−0.299
Germany	0.967	0.959	−0.421	−1.028	−0.545	−0.699	−0.740	−0.978	−1.199	−0.659	−1.151	−1.417	−0.519	−1.291	−1.278	−0.469	−0.274
Hong Kong	0.001	0.943	0.210	0.268	0.423	0.665	0.421	0.176	0.710	0.477	0.130	0.607	0.465	0.110	0.256	0.359	0.042
Ireland	0.951	0.952	−0.584	−1.217	−1.325	−0.830	−1.313	−1.905*	−0.583	−1.235	−2.054**	−0.295	−1.111	−2.178**	−0.457	−1.132	−0.608
Israel	0.947	0.967	−0.691	−0.993	−0.676	−0.928	−0.221	−1.051	−0.811	−0.020	−0.795	−1.019	0.156	−0.622	−1.876*	−0.178	−0.964
Italy	0.961	0.954	−0.125	−0.837	−0.694	−0.239	−0.523	−1.162	−0.518	−0.449	−1.304	−0.628	−0.316	−1.422	−0.448	−0.302	−0.319
Japan	0.939	0.962	−0.284	−0.559	−0.595	−0.113	−0.337	−0.575	−0.096	−0.250	−0.899	−0.235	−0.179	−1.189	−0.251	−0.090	−0.685
Luxembourg	0.959	0.955	−0.149	−0.396	−0.632	−0.276	−0.571	−1.096	−0.560	−0.564	−1.241	−0.668	−0.481	−1.359	−0.488	−0.472	−0.279
Netherlands	0.966	0.959	−0.388	−1.341	−0.717	−0.634	−1.002	−1.173	−1.030	−0.876	−1.312	−1.174	−0.682	−1.429	−1.011	−0.553	−0.346
New Zealand	0.972	0.953	−0.697	−1.566	−1.020	−1.176	−1.853*	−1.583	−2.082**	−1.488	−1.738*	−3.178***	−1.240	−1.746*	−4.637***	−0.960	−0.259
Norway	0.969	0.977	0.317	−0.288	−0.124	0.148	−0.538	−0.470	−0.094	−0.461	−0.548	−0.192	−0.380	−0.522	−0.046	−0.397	−0.088
Portugal	0.941	0.951	−0.142	−0.908	−0.959	−0.260	−1.021	−1.433	−0.539	−0.933	−1.543	−0.645	−0.771	−1.621	−0.464	−0.755	−0.407
Singapore	0.880	0.970	−1.132	0.149	0.027	−1.135	0.404	0.013	−1.387	0.365	−0.190	−1.847*	0.387	−0.460	−2.638***	0.290	−0.443
Spain	0.916	0.945	−0.137	−0.821	−0.884	−0.255	−0.884	−1.401	−0.533	−0.889	−1.520	−0.640	−0.755	−1.623	−0.459	−0.759	−0.367
Sweden	0.958	0.953	0.172	−1.205	−0.643	0.133	−0.753	−1.068	−0.088	−0.593	−1.216	−0.319	−0.329	−1.122	−0.263	−0.295	−0.232
Switzerland	0.938	0.985	−1.118	−3.237***	−0.505	−1.692*	−3.033***	−0.856	−3.822***	−2.779***	−1.033	−8.865***	−2.462**	−1.168	−7.524***	−1.855*	−0.325
UK	0.975	0.969	−0.722	−2.481**	−0.619	−0.935	−1.111	−0.868	−0.933	−0.632	−0.769	−0.922	−0.427	−0.649	−1.306	−0.226	−0.433

(*continued on next page*)

Table 4 (*continued*).

	$\hat{\phi}$	$\hat{\rho}^{RER}$	$h=1$			$h=4$			$h=8$			$h=12$			$h=20$		
			$t_{h,NW}$	$t_{h,ivx}^{trf,res}$	$t_{h,ivx}^{rev,PL}$	$t_{h,NW}$	$t_{h,ivx}^{trf,res}$	$t_{h,ivx}^{rev,PL}$	$t_{h,NW}$	$t_{h,ivx}^{trf,res}$	$t_{h,ivx}^{rev,PL}$	$t_{h,NW}$	$t_{h,ivx}^{trf,res}$	$t_{h,ivx}^{rev,PL}$	$t_{h,NW}$	$t_{h,ivx}^{trf,res}$	$t_{h,ivx}^{rev,PL}$
Brazil	0.962	0.994	1.124	−0.065	−0.340	0.853	0.130	−0.546	0.736	0.080	−0.482	0.504	0.013	−0.336	0.353	−0.271	−0.128
Bulgaria	0.832	0.918	−0.370	−0.443	−0.891	−0.626	−0.128	−1.277	−0.723	−0.117	−1.025	−0.674	−0.001	−0.733	−0.469	−0.052	0.010
Chile	0.977	0.979	0.934	−0.441	−0.459	0.757	−0.215	−0.698	0.641	−0.232	−0.784	0.462	−0.225	−0.655	0.292	−0.196	−0.119
China	0.530	0.962	−0.958	0.439	0.490	−0.959	0.585	0.286	−1.071	0.610	0.086	−1.179	0.550	−0.099	−1.603	0.436	−0.559
Colombia	0.965	0.977	1.224	−0.443	−0.622	1.028	−0.431	−0.621	0.847	−0.305	−0.511	0.754	−0.072	−0.506	0.577	−1.630	−0.288
Czech Rep.	0.946	0.960	−0.943	−0.695	−0.431	−1.024	−0.308	−0.641	−1.188	−0.284	−0.773	−1.186	−0.241	−0.792	−0.965	−0.311	−0.458
Egypt	0.943	1.014	1.681*	0.397	0.152	2.299**	0.324	−0.065	2.935***	0.208	−0.063	2.821***	0.016	0.068	2.878***	−0.009	0.337
Greece	0.865	0.932	−0.021	−1.219	−1.072	−0.144	−1.158	−1.690*	−0.463	−1.090	−1.854*	−0.597	−0.897	−1.941*	−0.426	−0.958	−0.418
Hungary	0.968	0.974	0.516	−0.661	−0.719	0.429	0.082	−0.727	0.169	0.157	−0.322	0.063	0.185	0.058	0.251	0.097	0.117
Iceland	0.971	0.981	0.733	0.233	−0.750	0.732	−0.192	−1.746*	0.638	−0.394	−1.552	0.478	−0.374	−1.049	0.891	−0.284	−1.643
India	0.850	1.007	2.069**	0.217	0.425	2.121**	0.193	0.319	2.055**	0.288	0.593	1.847*	0.243	0.725	1.877*	0.081	0.293
Indonesia	0.912	0.988	1.410	0.415	−0.486	1.779*	0.373	−0.831	1.989**	0.338	0.033	1.936*	0.234	0.319	2.285**	0.092	−0.163
Korea	0.930	0.887	−0.027	−1.324	−2.170**	−0.018	−2.807***	−2.519**	−0.120	−1.856*	−1.910*	−0.390	−1.201	−1.440	−0.498	−0.698	−1.010
Mexico	0.951	0.997	2.124**	0.038	−0.175	2.484**	0.212	−0.213	3.264***	0.165	−0.150	4.315***	0.093	−0.290	4.764***	0.120	0.079
Peru	0.931	0.984	0.073	0.077	0.418	−0.121	0.491	−0.044	−0.230	0.593	−0.229	−0.312	0.454	−0.421	−0.377	0.173	−0.864
Philippines	0.852	0.937	0.737	−0.491	−1.004	0.682	−0.546	−1.162	0.576	−0.204	−0.652	0.414	−0.203	−0.539	0.084	−0.829	−0.360
Poland	0.950	0.931	−0.302	−1.846*	−0.716	−0.466	−0.394	−0.834	−0.534	−0.176	−0.639	−0.501	−0.062	−0.573	−0.432	−0.144	−0.795
Romania	0.496	0.919	0.524	−1.725*	−0.917	0.312	−1.276	−0.801	0.246	−0.701	−0.436	0.347	−0.173	−0.243	0.822	−0.759	−0.124
Russian Fed.	0.942	0.990	1.479	0.585	−0.062	1.547	0.508	0.044	1.745*	0.369	−0.035	1.813*	0.206	0.074	1.962**	0.425	0.051
South Africa	0.981	0.997	1.214	−0.018	−0.430	1.212	−0.033	−0.688	1.260	−0.033	−0.512	1.290	−0.009	0.040	1.620	0.011	0.002
Thailand	0.954	0.980	−0.601	−0.425	−0.335	−0.799	−0.054	−0.815	−1.052	−0.059	−1.527	−1.348	−0.078	−1.602	−1.406	−0.502	−1.362
Ukraine	0.848	1.002	1.443	0.434	0.000	1.649	0.360	0.130	2.128**	0.299	−0.114	3.090***	0.207	−0.216	6.550***	0.858	−0.181

Notes: h is the prediction horizon, $t_{h,NW}$ is the OLS t-statistic with Newey–West standard errors, $t_{h,ivx}^{trf,res}$ is the t-statistic computed from a residual augmented transformed regression and $t_{h,ivx}^{rev,PL}$ is the t-statistic computed from a reversed regression as suggested by Phillips and Lee (2013). $\hat{\phi}$ is an estimate of the contemporaneous correlation computed as indicated in (7.3), and $\hat{\rho}^{RER}$ is an estimate of ρ_i^{RER} computed as indicated in (7.2).

*Statistically significant at the 10% nominal level.

**Statistically significant at the 5% nominal level.

***Statistically significant at the 1% nominal level.

Because the results may be affected by the period where short-term US nominal interest rates were at or near their effective lower bound (see Amador et al., 2020, for a discussion) the analysis is also conducted for the period from 1973:Q1 to 2008:Q4 (see Table S.39 in the Supplementary Appendix). Even with the exclusion of the information from 2009:Q1 to 2020:Q1 the conclusions are essentially in line with what we have observed in Panel A of Table 4 for the full sample (1973:Q1 to 2020:Q1). The smaller number of significant results obtained for $t_{h,ivx}^{trf,res}$ and $t_{h,ivx}^{rev,PL}$, suggest that there is less evidence of predictability in this period (particularly for the emerging markets), potentially highlighting the importance of inflation targeting policies suggested by EJR.

If we consider the period where most countries adopted inflation targeting policies for most of the time (recall that after 1999 most countries considered had already adopted inflation targeting) we clearly observe the general conclusion of EJR that the *RER*'s predictive power seems to increase as h increases particularly for $h \geq 8$ (see Table S.40 in the Supplementary Appendix). This pattern is most clearly seen for the developed markets group. Finally, if we focus on the period from 1999:Q1 to 2020:Q1 (see Panel B of Table 4), which roughly corresponds to a period where most countries adopted inflation targeting policies, we observe that the number of significant cases reduces considerably, indicating a reduction in predictability of changes in *NER* by the *RER*.

7.2.2. The relative price predictive regression

Table 5 reports the tests results computed from the relative-price long-horizon predictive regression,

$$log\left(\frac{P_{i,t+h}/P_{t+h}}{P_{it}/P_t}\right) = \alpha_{ih}^{\pi} + \beta_{ih}^{\pi} log(RER_{it}) + u_{i,t+h}^{\pi} \tag{7.4}$$

along with estimates of the contemporaneous correlation $\hat{\phi}_i$ in (7.3) where, in this case, for estimation we replace $\hat{u}_{i,t+1}^{NER}$ by $\hat{u}_{i,t+1}^{\pi}$. The full period of analysis, from 1973:Q1 to 2020:Q1, corresponds to a period during which inflation dynamics changed considerably (see Rogoff, 2003). Inflation in industrial economies started to decline in the early 1980s while inflation in emerging economies only began declining in the 1990s. Average inflation was the highest in the seventies, it decreased at the beginning of the eighties and it has been even lower since the beginning of the 1990s.

From the results in Panel A of Table 5 for the period from 1973:Q1 to 2020:Q1 we observe a large number of rejections of the null hypothesis of no predictability regardless of the test considered. This is also the case in Table S.41 in the Supplementary Appendix, corresponding to the 1973:Q1–2008:Q4 period, with very similar conclusions to those just described for Panel A of Table 5.

The impact of the changes in exchange rate policy in emerging markets is observable on comparing the results in Panel A of Tables 5 and S.41 with those in Panel B of Table 5. The latter, computed in the sample from 1999:Q1 onward, a period where most of these countries adopted inflation targeting policies, show that inflation differentials are less predictive. Note that for the developed markets group $t_{h,ivx}^{trf,res}$ and $t_{h,ivx}^{rev,PL}$ suggest rejection of the null hypothesis of no predictability only for Ireland and Israel. Similarly, and in contrast to the results in Panel A of Tables 5 and S.41, these statistics also suggest a relevant decrease in significant results in emerging markets.

7.3. Summary of empirical results

- Our results suggest that for most of the countries considered *RER* is strongly persistent with an estimated autoregressive root very close to unity (for instance, when considering the sample from 1973:Q1 to 2020:Q1, for 44 out of the 45 countries considered the estimated autoregressive root is greater than or equal to 0.953), which may be at odds with the weakly persistent assumption of EJR. This persistence may impact the $t_{h,NW}$ test used in their analysis, as discussed in Section 2.2, and lead to spurious rejections of the null hypothesis of no predictability.
- Based on the outcomes of $t_{h,NW}$, EJR strongly support the conclusion that current *RER* is highly negatively correlated with changes in future *NER*s at horizons of three or more years. We also observe negative outcomes for the IVX-based statistics for almost all countries and for all values of h considered, albeit many of these test outcomes are not statistically significant. This finding also appears robust to the other sample periods we considered.
- In line with ERJ, our results also provide evidence of predictability as h increases. However, we note that the number of rejections is largest for $t_{h,NW}$ and smallest for $t_{h,ivx}^{trf,res}$.
- According to EJR, in countries with inflation-targeting policies, *RER* reverts towards the mean through changes in the *NER*. Hence, current *RER* should predict future nominal exchange rates, but not changes in relative rates of inflation. For the period after 1999 where most countries adopted inflation targeting policies, we observe the general conclusion of EJR that the *RER*'s predictive power appears to increase as h increases, particularly for $h \geq 8$. The results for the sample from 1999:Q1 onward, a period were most emerging markets' countries adopted inflation targeting policies, show that inflation differentials are less predictive.
- The large number of statistically significant results in Panel A of Tables 5 for the period from 1973:Q1 to 2020:Q1 (regardless of the tests considered, $t_{h,NW}$, $t_{h,ivx}^{trf,res}$ or $t_{h,ivx}^{rev,PL}$), would appear to suggest that a large number of countries adjust *RER* through predictable inflation differentials rather than through changes in *NER*. This is consistent with EJR's findings for countries with fixed and quasi-fixed exchange rates (e.g. China and Hong Kong, and France, Ireland, Italy, Portugal, and Spain starting in 1999). Potential justifications for the large number of significant results observed may

Table 5
Relative price long-horizon predictive regression results.

	$\hat{\phi}$	$h=1$			$h=4$			$h=8$			$h=12$			$h=20$		
		$t_{h,NW}$	$t^{trf,res}_{h,ivx}$	$t^{rev,PL}_{h,ivx}$	$t_{h,NW}$	$t^{trf,res}_{h,ivx}$	$t^{rev,PL}_{h,ivx}$	$t_{h,NW}$	$t^{trf,res}_{h,ivx}$	$t^{rev,PL}_{h,ivx}$	$t_{h,NW}$	$t^{trf,res}_{h,ivx}$	$t^{rev,PL}_{h,ivx}$	$t_{h,NW}$	$t^{trf,res}_{h,ivx}$	$t^{rev,PL}_{h,ivx}$
PANEL A: Period from 1973:Q1 to 2020:Q2.																
Australia	0.124	−0.910	−0.564	−0.336	−1.100	−0.814	−0.583	−1.256	−0.996	−0.448	−1.454	−1.090	−0.314	−1.677*	−1.142	−0.140
Austria	0.088	−2.351**	0.173	0.007	−2.358**	−0.093	−0.295	−2.388**	−0.470	−0.684	−2.578***	−0.833	−1.327	−2.871***	−1.288	−1.691*
Belgium	0.139	−0.912	2.952***	2.750***	−0.977	2.120**	1.826*	−1.250	0.685	0.261	−1.753*	−0.705	−1.471	−3.008***	−2.158**	−3.074***
Canada	0.143	−0.355	0.197	0.381	−0.559	−0.018	−0.103	−0.661	−0.062	−0.194	−0.704	−0.051	−0.039	−0.809	−0.195	0.060
Denmark	0.096	0.460	1.938*	1.892*	0.319	1.465	1.412	0.235	1.056	1.086	0.205	0.698	0.800	0.052	0.142	0.434
Finland	0.093	0.415	−0.153	−0.130	0.337	−0.567	−0.431	0.224	−1.087	−0.420	0.092	−1.412	−0.219	−0.084	−1.344	0.267
France	0.151	0.144	2.077**	2.154**	0.061	1.617	1.761*	−0.031	0.996	1.315	−0.126	0.287	0.850	−0.276	−0.812	0.084
Germany	0.084	−4.688***	−1.362	−1.253	−4.753***	−1.655*	−1.555	−4.868***	−2.019**	−1.813*	−5.246***	−2.316**	−2.211**	−5.608***	−2.623***	−2.257**
Hong Kong	0.403	1.297	0.229	0.100	1.144	0.125	0.002	0.972	−0.052	−0.113	0.855	−0.185	−0.160	0.740	−0.357	−0.605
Ireland	0.122	−2.356**	0.265	0.443	−2.428**	−0.196	0.216	−2.468**	−0.715	0.360	−2.452**	−1.106	0.377	−2.158**	−1.401	0.206
Israel	0.851	−3.040***	−1.022	−0.769	−2.749***	−1.409	−1.209	−2.612***	−1.881*	−1.885*	−2.700***	−2.290**	−2.687***	−3.502***	−3.004***	−4.400***
Italy	0.119	2.486**	−0.457	−0.539	2.236**	−0.772	−0.715	2.004**	−1.048	−0.689	1.844*	−1.290	−0.616	1.610	−1.474	−0.494
Japan	0.032	−2.941***	1.157	2.120**	−3.552***	0.805	0.782	−4.460***	0.409	−0.167	−5.816***	0.149	−1.086	−10.475***	−0.027	−1.799*
Luxembourg	0.133	−1.203	2.144**	2.387**	−1.244	1.050	1.196	−1.439	−0.454	−0.425	−1.923*	−1.676*	−2.256**	−3.006***	−2.617***	−3.288***
Netherlands	0.076	−2.245**	0.076	0.345	−2.329**	−0.393	−0.297	−2.656***	−1.012	−1.173	−3.467***	−1.555	−2.409**	−5.320***	−2.327**	−3.286***
New Zealand	0.110	−0.548	0.178	0.382	−0.861	−0.081	−0.023	−1.054	−0.301	−0.137	−1.248	−0.567	−0.005	−1.542	−1.033	−0.129
Norway	0.159	1.025	0.979	0.825	1.011	0.759	0.424	0.986	0.391	0.061	0.904	0.111	−0.065	0.765	−0.377	−0.554
Portugal	0.202	2.558**	−0.342	−0.659	2.300**	−0.519	−0.592	2.069**	−0.677	−1.017	1.898*	−0.794	−1.442	1.643	−0.908	−0.485
Singapore	0.257	−4.348***	−1.498	−2.109**	−9.081***	−1.455	−2.906***	−9.297***	−1.385	−2.483**	−9.040***	−1.301	−1.928*	−8.512***	−1.137	−1.640
Spain	0.130	2.650**	−0.625	−0.876	2.445**	−0.838	−1.141	2.248**	−1.070	−1.340	2.101**	−1.241	−1.130	1.879*	−1.307	−0.130
Sweden	0.143	0.418	0.340	0.223	0.411	0.282	−0.043	0.376	0.039	−0.321	0.302	−0.234	−0.328	0.179	−0.422	−0.097
Switzerland	0.102	−3.416***	−0.538	−0.502	−3.747***	−0.684	−0.779	−3.947***	−0.918	−1.159	−4.270***	−1.070	−1.251	−4.627***	−0.959	−0.996
United Kingdom	0.068	−2.548**	−0.886	−0.885	−2.458**	−1.108	−0.976	−2.405**	−1.220	−0.519	−2.448**	−1.166	−0.191	−2.562**	−0.993	0.088
Brazil	0.889	−3.245***	−1.185	−0.840	−3.025***	−1.481	−1.263	−2.920***	−1.794*	−1.862*	−2.959***	−2.062**	−2.480**	−3.530***	−2.525**	−3.694***
Bulgaria	0.942	−3.094***	−1.078	−0.709	−2.748***	−1.796*	−1.366	−3.270***	−2.351**	−1.706*	−4.784***	−2.629***	−1.779*	−26.919***	−3.189***	−2.478**
Chile	0.226	−0.738	137.640***	−0.559	−0.627	−11.940***	−0.625	−0.560	−1.928*	−0.266	−0.474	−2.529**	0.312	−0.282	−8.628***	0.641
China	0.288	1.393	−0.246	−0.573	1.240	−0.453	−0.765	1.141	−0.638	−0.885	1.063	−0.733	−0.470	1.020	−0.839	−1.301
Colombia	0.189	3.542***	−1.056	−1.582	3.113***	−1.086	−1.747*	2.731***	−1.086	−1.894*	2.469**	−1.065	−2.260**	2.126**	−1.016	−2.280**
Czech Rep.	0.209	1.823*	0.762	0.864	1.578	0.427	0.472	1.372	0.248	0.186	1.276	0.170	0.028	1.230	0.089	−0.418
Egypt	0.141	0.283	0.461	0.542	0.125	0.377	0.334	−0.116	0.264	−0.028	−0.325	0.124	−0.183	−0.658	−0.227	−0.446
Greece	0.231	2.663***	−0.295	−0.842	2.357**	−0.368	−0.527	2.085**	−0.424	−0.852	1.897*	−0.472	−1.032	1.665*	−0.540	−1.548
Hungary	0.436	2.997***	1.020	0.762	2.644***	0.817	0.324	2.343**	0.635	−0.374	2.154**	0.480	−1.289	1.952*	0.212	−3.532***
Iceland	0.320	1.014	−1.702*	−1.964**	0.712	−2.010**	−2.598***	0.465	−2.268**	−2.775***	0.284	−2.469**	−3.181***	0.068	−2.584***	−2.788***
India	0.386	5.730***	1.957*	1.510	5.488***	1.872*	1.697*	5.020***	1.728*	1.590	4.936***	1.645*	0.835	5.413***	1.457	0.382
Indonesia	0.377	4.289***	0.386	−0.008	4.517***	0.156	0.026	4.514***	0.152	−0.109	4.434***	0.229	0.061	4.286***	0.305	−0.075
Korea	0.281	2.810***	−1.850*	−2.312**	2.649***	−2.084**	−2.139**	2.535**	−2.078**	−1.811*	2.457**	−1.916*	−1.407	2.448**	−1.447	−1.189
Mexico	0.600	−2.586***	−1.048	−0.709	−2.406**	−1.352	−1.163	−2.304**	−1.646*	−1.902*	−2.302**	−1.857*	−2.669***	−2.554**	−2.162**	−3.967***
Peru	0.928	−2.573**	−0.666	−0.310	−2.339**	−0.955	−0.800	−2.186**	−1.271	−1.508	−2.130**	−1.553	−2.285**	−2.215**	−1.895*	−3.521***
Philippines	0.390	3.839***	−0.693	−0.924	3.773***	−0.966	−0.950	3.568**	−1.063	−1.621	3.299***	−1.030	−1.909*	2.917***	−0.927	−2.582***
Poland	0.872	−2.295**	−0.954	−0.535	−2.161**	−1.337	−1.303	−2.107**	−1.651*	−1.441	−2.164**	−1.869*	−2.092**	−2.888***	−2.131**	−3.886***
Romania	0.393	−8.395***	−2.305**	−2.162**	−11.743***	−2.657***	−1.617	−14.099***	−2.845***	−1.196	−13.423***	−2.716***	−0.090	−12.107***	−2.029**	−0.426
Russian Federation	0.016	−0.700	−1.839*	−1.327	−0.576	−1.329	−0.820	−0.289	−0.465	0.264	−0.081	−0.248	0.271	−0.088	−1.652*	0.360
South Africa	0.077	0.972	−0.563	−0.495	0.606	−0.726	−1.009	0.295	−0.818	−1.416	0.098	−0.840	−1.675*	−0.165	−0.743	−2.296*
Thailand	0.139	1.152	−0.502	−0.977	0.947	−0.778	−0.546	0.920	−0.645	−1.051	1.062	−0.425	−1.188	1.434	0.140	−0.778
Ukraine	0.382	−1.499	1.809*	−0.836	−1.419	1.876*	−0.599	−1.381	0.904	−0.422	−1.355	10.102***	−0.143	−1.425	51.700***	−0.097

(*continued on next page*)

Table 5 (*continued*).

	$\hat{\phi}$	$h=1$			$h=4$			$h=8$			$h=12$			$h=20$		
		$t_{h,NW}$	$t_{h,ivx}^{trf,res}$	$t_{h,ivx}^{rev,PL}$	$t_{h,NW}$	$t_{h,ivx}^{trf,res}$	$t_{h,ivx}^{rev,PL}$	$t_{h,NW}$	$t_{h,ivx}^{trf,res}$	$t_{h,ivx}^{rev,PL}$	$t_{h,NW}$	$t_{h,ivx}^{trf,res}$	$t_{h,ivx}^{rev,PL}$	$t_{h,NW}$	$t_{h,ivx}^{trf,res}$	$t_{h,ivx}^{rev,PL}$
PANEL B: Period from 1999:Q1 to 2020:Q2.																
Australia	0.202	2.016**	−0.166	0.088	1.754*	−0.169	−0.184	1.646*	−0.432	−0.630	1.756*	−0.673	−1.235	2.335**	−0.395	−1.143
Austria	0.173	−1.303	−0.094	−0.178	−1.276	−0.149	−0.266	−1.105	−0.212	−0.287	−1.058	−0.454	−0.627	−1.011	−0.869	−0.748
Belgium	0.273	−1.226	−0.493	−0.498	−1.302	−0.077	−0.420	−1.130	−0.179	−0.297	−1.133	−0.324	−0.721	−1.142	−0.598	−0.663
Canada	0.335	−0.945	0.360	0.604	−1.147	0.164	0.289	−1.300	−0.064	0.244	−1.656*	−0.392	−0.181	−4.238***	−0.742	−0.936
Denmark	0.226	−2.232**	−0.359	−0.250	−2.345**	−0.259	−0.465	−2.339**	−0.346	−0.466	−2.632***	−0.531	−0.973	−3.541***	−0.633	−1.180
Finland	0.239	−2.635***	−0.294	−0.434	−2.564***	−0.377	−0.532	−2.599***	−0.637	−0.697	−2.766***	−1.002	−1.240	−2.926***	−1.229	−1.218
France	0.203	−4.645***	−0.154	−0.254	−5.228***	0.157	−0.153	−5.886***	0.236	0.176	−7.381***	0.189	−0.151	−16.822***	−0.073	−0.626
Germany	0.194	−4.906***	−0.491	−0.292	−5.647***	−0.342	−0.391	−6.223***	−0.369	−0.268	−7.568***	−0.489	−0.654	−12.625***	−0.815	−0.662
Hong Kong	0.891	−0.797	−0.213	−0.028	−0.584	−0.252	0.338	−0.438	−0.278	0.227	−0.332	−0.295	0.078	−0.114	−0.329	−0.093
Ireland	0.219	1.882*	3.091***	2.241**	2.098**	2.927***	2.068**	2.403**	2.735***	1.998**	2.594***	2.417**	1.127	4.149***	3.059***	0.707
Israel	0.343	−1.479	−1.253	−0.834	−1.606	−1.674*	−1.581	−1.450	−1.568	−1.413	−1.445	−1.274	−1.330	−1.739*	−0.655	−0.409
Italy	0.212	−2.108**	−0.254	−0.184	−2.106**	0.056	−0.275	−2.011**	0.129	−0.149	−2.211**	−0.004	−0.573	−3.513***	−0.157	−1.168
Japan	0.005	−5.346***	−0.420	−0.161	−5.301***	−0.697	−0.352	−5.190***	−1.034	−0.571	−4.988***	−1.244	−0.822	−4.740***	−1.479	−1.013
Luxembourg	0.179	−1.300	−0.128	−0.264	−1.259	−0.249	−0.425	−1.193	−0.230	−0.338	−1.264	−0.344	−0.673	−1.562	−0.418	−0.844
Netherlands	0.233	−0.853	0.531	0.641	−0.805	0.527	0.354	−0.927	0.449	0.167	−1.422	0.267	−0.495	−3.300***	−0.302	−1.027
New Zealand	0.159	0.167	0.046	0.255	−0.125	0.021	−0.074	0.103	0.072	0.137	0.003	−0.116	−0.390	0.143	−0.170	−0.204
Norway	0.306	−0.312	0.389	0.247	−0.251	0.335	0.146	−0.277	0.034	−0.165	−0.355	−0.254	−0.348	−0.478	−0.774	−0.543
Portugal	0.171	−1.163	1.278	0.711	−0.949	1.144	0.724	−0.904	1.026	0.576	−1.053	0.762	0.032	−1.890*	0.829	−0.486
Singapore	0.256	−2.618***	−2.019**	−1.726*	−2.187**	−1.567	−1.458	−1.714*	−1.403	−1.256	−1.407	−1.296	−1.062	−1.015	−1.093	−0.734
Spain	0.120	−0.518	0.735	−0.108	−0.340	0.576	0.073	−0.220	0.719	0.218	−0.219	0.582	0.017	−0.311	0.762	−0.159
Sweden	0.149	−3.382***	0.689	0.251	−3.540***	0.782	0.414	−3.944***	0.501	0.363	−5.158***	−0.006	−0.242	−9.531***	−0.339	−0.642
Switzerland	0.149	−2.055**	−0.056	0.095	−2.269**	−0.090	0.045	−2.373**	−0.112	0.164	−2.715***	−0.122	−0.039	−4.089***	−0.134	−0.106
United Kingdom	0.274	0.519	0.150	0.228	0.334	0.186	0.032	0.103	−0.166	−0.216	−0.049	−0.563	−0.487	−0.303	−1.470	−0.819
Brazil	0.223	3.304***	0.018	0.421	3.231***	−0.262	−0.187	3.501***	−0.323	−0.861	4.143***	−0.319	−1.214	7.797***	−0.464	−0.990
Bulgaria	0.205	1.713*	−1.215	−1.355	1.495	−0.900	−1.180	1.386	−0.630	−0.994	1.284	−0.673	−1.351	1.349	−0.271	−1.022
Chile	0.447	2.311**	−1.333	−0.835	2.461**	−0.790	−1.479	2.745***	−0.708	−1.899*	3.200***	−0.517	−1.947*	3.630***	0.060	−0.910
China	0.401	0.270	−1.040	−1.003	0.165	−0.669	−0.816	0.261	−0.627	−0.669	0.413	−0.523	−0.438	0.863	−0.407	−0.224
Colombia	0.180	5.137***	−0.023	0.307	5.232***	−0.216	−0.021	5.721***	−0.253	−0.243	6.633***	−0.059	−0.273	8.088***	−0.950	−0.210
Czech Rep.	0.188	0.485	−0.195	−0.051	0.252	0.007	−0.245	0.120	0.116	−0.276	−0.189	0.229	−0.542	−0.387	0.320	−0.183
Egypt	0.334	6.238***	1.971**	1.853*	8.487***	1.690*	1.487	11.541***	1.220	1.115	13.180***	0.107	0.892	34.599***	1.736*	0.649
Greece	0.188	−0.780	1.128	−0.441	−0.535	0.768	−0.047	−0.434	0.717	0.042	−0.407	0.540	−0.277	−0.423	0.862	−0.369
Hungary	0.094	3.151***	−1.909*	−1.402	2.868***	−2.268**	−2.086**	2.656***	−2.282**	−2.167**	2.572**	−1.961**	−1.995**	2.651***	−0.966	−1.916*
Iceland	0.502	2.758***	−0.251	−0.287	2.678***	−0.905	−1.679*	2.621***	−1.428	−3.072***	2.603***	−1.555	−2.836***	3.065***	−1.368	−4.410***
India	0.435	4.646***	0.427	0.521	4.384***	0.329	0.356	4.077***	0.072	−0.128	4.026***	0.043	−0.243	4.889***	0.070	−0.690
Indonesia	0.242	5.040***	−0.144	−0.580	5.477***	−0.166	−1.223	5.552***	−0.301	−1.494	5.855***	−0.215	−1.000	6.542***	−0.025	−0.974
Korea	0.293	0.742	1.203	0.960	0.822	0.828	−0.023	1.143	−0.272	−0.523	1.420	−1.195	−1.150	2.006**	−1.775*	−2.110**

(*continued on next page*)

Table 5 (*continued*).

	$\hat{\phi}$	$h=1$			$h=4$			$h=8$			$h=12$			$h=20$		
		$t_{h,NW}$	$t_{h,ivx}^{trf,res}$	$t_{h,ivx}^{rev,PL}$	$t_{h,NW}$	$t_{h,ivx}^{trf,res}$	$t_{h,ivx}^{rev,PL}$	$t_{h,NW}$	$t_{h,ivx}^{trf,res}$	$t_{h,ivx}^{rev,PL}$	$t_{h,NW}$	$t_{h,ivx}^{trf,res}$	$t_{h,ivx}^{rev,PL}$	$t_{h,NW}$	$t_{h,ivx}^{trf,res}$	$t_{h,ivx}^{rev,PL}$
Mexico	0.356	5.977***	−0.402	−0.491	7.380***	−0.186	−0.109	9.263***	−0.167	0.018	12.564***	−0.133	0.133	28.579***	−0.293	0.281
Peru	0.470	1.063	−2.185**	−1.232	1.047	−1.956*	−1.644	1.041	−2.299**	−2.051**	1.158	−2.470**	−2.161**	1.310	−2.589***	−2.581***
Philippines	0.149	5.120***	−0.099	−0.123	5.606***	0.013	−0.322	5.816***	0.074	−0.135	5.495***	0.115	0.095	6.083***	0.622	0.197
Poland	0.232	0.826	−0.712	−0.092	0.519	−0.935	−0.852	0.218	−1.108	−1.229	0.009	−1.105	−1.591	−0.021	−0.620	−1.199
Romania	0.173	1.092	−2.745***	−1.931*	0.978	−2.568**	−1.156	0.932	−1.700*	−0.315	0.907	−0.507	0.013	0.969	−3.271	0.082
Russian Federation	0.403	3.987***	−1.876*	−2.101**	3.843***	−1.668	−2.214**	3.957***	−1.253	−1.958*	4.351***	−0.722	−1.527	5.711***	−1.705*	−1.041
South Africa	0.223	5.009***	1.578	1.811*	5.067***	0.751	0.816	5.477***	0.112	−0.054	6.204***	−0.159	−0.575	6.638***	−0.029	−0.659
Thailand	0.027	−0.650	−1.087	−1.271	−0.464	−1.030	−1.117	−0.246	−0.981	−0.817	−0.061	−0.800	−0.362	0.464	0.470	0.955
Ukraine	0.543	3.210***	0.267	0.299	3.580	0.148	0.185	4.520***	0.065	0.112	6.458***	0.051	−0.049	10.855***	0.272	−0.171

Notes: h is the prediction horizon, $t_{h,NW}$ is the OLS t-statistic with Newey–West standard errors, $t_{h,ivx}^{trf,res}$ is the t-statistic computed from a residual augmented transformed regression and $t_{h,ivx}^{rev,PL}$ is the t-statistic computed from a reversed regression as suggested by Phillips and Lee (2013). $\hat{\phi}$ is an estimate of the contemporaneous correlation computed as indicated in (7.3).
*Statistically significant at the 10% nominal level.
**Statistically significant at the 5% nominal level.
***Statistically significant at the 1% nominal level.

be related to uncontrolled changes in exchange rate policy, as many countries, particular in the emerging markets group adopted several exchange rate regimes between 1973 and 2020 (Ilzetzki et al., 2017), and to the persistence changes of inflation dynamics observed over this period.

8. Conclusions

In this paper, we have contributed to the long-horizon predictability literature by proposing new tests developed within a transformed regression framework using the IVX estimation approach of Kostakis et al. (2015). We have demonstrated that our proposed tests are (asymptotically) robust to whether the predictors are weakly or strongly persistent and to the induced serial correlation in the errors arising from the temporal aggregation of the dependent variable used in the long-horizon predictive regression. Within a residual augmentation framework we have shown that the estimation effect from fitting an autoregression to the predictor to obtain the necessary residuals to augment the predictive regression is asymptotically negligible in the set-up we consider and leads to more efficient estimation of the transformed predictive regression model on which our long-horizon tests are based. Specifically, the residual augmentation approach eliminates endogeneity in the limit, such that the bias of the IVX slope coefficient estimator is reduced compared to the corresponding IVX estimation from the transformed regression without this additional regressor. We have formally established the conditions required for the asymptotic validity of our proposed tests, such that the statistics on which they are based have standard limiting null distributions, free of nuisance parameters arising from the innovations. These conditions allow for quite general patterns of unconditional and conditional time variation in the innovations with no need for the practitioner to specify a parametric model for either the conditional or unconditional time-variation.

Our Monte Carlo results contrast the finite size and power properties of our proposed tests with the leading long-horizon predictability tests in the literature. The results obtained suggest that our proposed tests overall display superior finite sample properties to the extant tests displaying robustness against features which are frequently found in time series, making them a useful addition to the literature. We have also provided an empirical application investigating the predictive power of real exchange rates for changes in nominal exchange rates and future inflation rates of a large number of developed and emerging countries, extending the analysis in Eichenbaum et al. (2020) to a wider range of countries and providing conclusions based on the robust statistics developed in this paper. Overall we find somewhat less evidence of predictability than Eichenbaum et al. (2020). This is perhaps expected as their analysis is based on standard regression t-tests which would appear to be inappropriate given that the predictors they consider appear to be strongly persistent.

Appendix A. Supplementary data

Supplementary material related to this article can be found online at https://doi.org/10.1016/j.jeconom.2022.06.006.

References

Amador, M., Bianchi, J., Bocola, L., Perri, F., 2020. Exchange rate policies at the zero lower bound. Rev. Econom. Stud. 87 (4), 1605–1645.
Amihud, Y., Hurvich, C.M., 2004. Predictive regressions: A reduced-bias estimation method. J. Financ. Quant. Anal. 39 (4), 813–841.
Amihud, Y., Hurvich, C.M., Wang, Y., 2009. Multiple-predictor regressions: Hypothesis testing. Rev. Financ. Stud. 22 (1), 413–434.
Amihud, Y., Hurvich, C.M., Wang, Y., 2010. Predictive regression with order-p autoregressive predictors. J. Empir. Financ. 17 (3), 513–525.
Andersen, T.G., Varneskov, R.T., 2021a. Consistent inference for predictive regressions in persistent economic systems. J. Econometrics 224 (1), 215–244.
Andersen, T.G., Varneskov, R.T., 2021b. Consistent Local Spectrum (LCM) Inference for Predictive Return Regressions. NBER Working Paper 28569, National Bureau of Economic Research.
Andersen, T.G., Varneskov, R.T., 2021c. Testing for parameter instability and structural change in persistent predictive regressions. J. Econometrics (forthcoming).
Ang, A., Bekaert, G., 2007. Return predictability: Is it there? Rev. Financ. Stud. 20 (3), 651–707.
Bauer, D., Maynard, A., 2012. Persistence-robust surplus-lag Granger causality testing. J. Econometrics 169 (2), 293–300.
Bekaert, G., Hodrick, R.J., 1992. Characterizing predictable components in excess returns on equity and foreign exchange markets. J. Finance 47 (2), 467–509.
Boudouk, J., Richardson, M., 1994. The statistics of long-horizon regressions revisited. Math. Finance 4 (2), 103–119.
Boudoukh, J., Matthew, R., 1993. Stock returns and inflation: A long-horizon perspective. Amer. Econ. Rev. 83 (5), 1346–1355.
Breitung, J., Demetrescu, M., 2015. Instrumental variable and variable addition based inference in predictive regressions. J. Econometrics 187 (1), 358–375.
Britten-Jones, M., Neuberger, A., Nolte, I., 2011. Improved inference in regression with overlapping observations. J. Bus. Finance Account. 38 (5–6), 657–683.
Cai, Z., Wang, Y., 2014. Corrigendum to "testing predictive regression models with nonstationary regressors" [J. Econometrics 178 (2014) 4–14]. J. Econometrics 181 (2), 194.
Campbell, J.Y., Cochrane, J.H., 1999. By force of habit: A consumption-based explanation of aggregate stock market behavior. J. Polit. Econ. 107 (1), 205–251.
Campbell, J.Y., Shiller, R.J., 1987. Cointegration and tests of present value models. J. Polit. Econ. 95 (5), 1062–1088.
Campbell, J.Y., Shiller, R.J., 1988. Stock prices, earnings, and expected dividends. J. Finance 43 (3), 661–676.
Campbell, J.Y., Viceira, L.M., 1999. Consumption and portfolio decisions when expected returns are time varying. Q. J. Econ. 114 (2), 433–495.
Campbell, J.Y., Yogo, M., 2006a. Efficient tests of stock return predictability. J. Financ. Econ. 81 (1), 27–60.
Campbell, J.Y., Yogo, M., 2006b. Predicting excess stock returns out of sample: Can anything beat the historical average? J. Financ. Econ. 81 (1), 27–60.
Cavanagh, C.L., Elliott, G., Stock, J.H., 1995. Inference in models with nearly integrated regressors. Econom. Theory 11 (5), 1131–1147.

Chang, C., J., I., Laurila, H., McAleer, M., 2018. Long run returns predictability and volatility with moving averages. Risk 6 (4), 105.
Chen, W.W., Deo, R.S., 2009. Bias reduction and likelyhood-based almost exactly sized hypothesis testing in predictive regressions using the restricted likelihood. Econom. Theory 25 (5), 1143–1179.
Cochrane, J.H., 1991. Production-based asset pricing and the link between stock returns and economic fluctuations. J. Finance 46 (1), 209–237.
Cochrane, J.H., 2008. The dog that did not bark: A defense of return predictability. Rev. Financ. Stud. 21 (4), 1533–1575.
Cochrane, J.H., 2011. Presidential address: Discount rates. J. Finance 66 (4), 1047–1108.
Demetrescu, M., Georgiev, I., Rodrigues, P., Taylor, A.M.R., 2022a. Extensions to IVX methods of inference for return predictability. J. Econometrics (forthcoming).
Demetrescu, M., Georgiev, I., Rodrigues, P., Taylor, A.M.R., 2022b. Testing for episodic predictability in stock returns. J. Econometrics 227 (1), 85–113.
Demetrescu, M., Hillmann, B., 2022. Nonlinear predictability of stock returns? Parametric vs. nonparametric inference in predictive regressions. J. Bus. Econom. Statist. 40, 382–397.
Demetrescu, M., Rodrigues, P.M.M., 2022. Residual-augmented IVX predictive regression. J. Econometrics 227 (2), 429–460.
Eichenbaum, M.S., Johannsen, B.K., Rebelo, S.T., 2020. Monetary policy and the predictability of nominal exchange rates. Rev. Econom. Stud. 88 (1), 192–228.
Elliott, G., Müller, U.K., Watson, M.W., 2015. Nearly optimal tests when a nuisance parameter is present under the null hypothesis. Econometrica 83 (2), 771–811.
Elliott, G., Stock, J.H., 1994. Inference in time series regression when the order of integration of a regressor is unknown. Econom. Theory 10 (3–4), 672–700.
Engel, C., Mark, N.C., West, K.D., Rogoff, K., Rossi, B., 2007. Exchange rate models are not as bad as you think [with comments and discussion]. NBER Macroecon. Annu. 22, 381–473.
Fama, E.F., 1998. Market efficiency, long-term returns, and behavioral finance. J. Financ. Econ. 49 (3), 283–306.
Fama, E.F., French, K.R., 1988. Dividend yields and expected stock returns. J. Financ. Econ. 22 (1), 3–25.
Georgiev, I., Harvey, D.I., Leybourne, S.J., Taylor, A.M.R., 2018. A bootstrap stationarity test for predictive regression invalidity. J. Bus. Econom. Statist. 37 (3), 528–541.
Gonzalo, J., Pitarakis, J.-Y., 2012. Regime-specific predictability in predictive regressions. J. Bus. Econom. Statist. 30 (2), 229–241.
Hjalmarsson, E., 2007. Fully modified estimation with nearly integrated regressors. Finance Res. Lett. 4 (2), 92–94.
Hjalmarsson, E., 2011. New methods for inference in long-horizon regressions. J. Financ. Quant. Anal. 46 (3), 815–839.
Hjalmarsson, E., 2012. Some curious power properties of long-horizon tests. Finance Res. Lett. 9 (2), 81–91.
Hodrick, R.J., 1992. Dividend yields and expected stock returns: Alternative procedures for inference and measurement. Rev. Financ. Stud. 5 (3), 357–386.
Ilzetzki, E., Reinhart, C.M., Rogoff, K.S., 2017. The Country Chronologies to Exchange Rate Arrangements into the 21st Century: Will the Anchor Currency Hold?. NBER Working Papers 23135, National Bureau of Economic Research, Inc.
Im, K.S., Schmidt, P., 2008. More efficient estimation under non-normality when higher moments do not depend on the regressors, using residual augmented least squares. J. Econometrics 144 (1), 219–233.
Jegadeesh, N., 1991. Seasonality in stock price mean reversion: Evidence from the U.S. and the U.K.. J. Finance 46 (4), 1427–1444.
Johannes, M., Korteweg, A., Polson, N., 2014. Sequential learning, predictability, and optimal portfolio returns. J. Finance 69 (2), 611–644.
Kandel, S., Stambaugh, R.F., 1996. On the predictability of stock returns: An asset-allocation perspective. J. Finance 51 (2), 385–424.
Kim, J.H., 2014. Predictive regression: An improved augmented regression method. J. Empir. Financ. 26, 13–25.
Kostakis, A., Magdalinos, T., Stamatogiannis, M.P., 2015. Robust econometric inference for stock return predictability. Rev. Financ. Stud. 28 (5), 1506–1553.
Kostakis, A., Magdalinos, T., Stamatogiannis, M.P., 2018. Taking stock of long-horizon predictability tests: Are factor returns predictable? https://Ssrn.Com/Abstract=3284149.
Lee, J.H., 2016. Predictive quantile regression with persistent covariates: IVX-QR approach. J. Econometrics 192 (1), 105–118.
Lettau, M., Van Nieuwerburgh, S., 2008. Reconciling the return predictability evidence. Rev. Financ. Stud. 21 (4), 1607–1652.
Lewellen, J., 2004. Predicting returns with financial ratios. J. Financ. Econ. 74 (2), 209–235.
Mankiw, G.N., Shapiro, M.D., 1986. Do we reject too often?: Small sample properties of tests of rational expectations models. Econom. Lett. 20 (2), 139–145.
Mark, N., 1995. Exchange rates and fundamentals: Evidence on long-horizon predictability. Amer. Econ. Rev. 85 (1), 201–218.
Maynard, A., Phillips, P.C.B., 2001. Rethinking an old empirical puzzle: econometric evidence on the forward discount anomaly. J. Appl. Econometrics 16 (6), 671–708.
Maynard, A., Ren, D., 2014. Assessing the Power of Long-Horizon Predictive Tests in Models of Bull and Bear Markets. Essays in Honor of Peter C. B. Phillips (Advances in Econometrics, Vol. 33), Emerald Group Publishing Limited, pp. 673–711.
Maynard, A., Shimotsu, K., 2009. Covariance-based orthogonality tests for regressors with unknown persistence. Econom. Theory 25 (1), 63–116.
Menzly, L., Santos, T., Veronesi, P., 2004. Understanding predictability. J. Polit. Econ. 112 (1), 1–47.
Mishkin, F.S., 1990. What does the term structure of interest rates tell us about future inflation? J. Monetary Econ. 25 (1), 77–90.
Nelson, C.R., Kim, M.J., 1993. Predictable stock returns: The role of small sample bias. J. Finance 48 (2), 641–661.
Newey, W.K., West, K.D., 1987. A simple, positive semi-definite, heteroskedasticity and autocorrelation consistent covariance matrix. Econometrica 55 (3), 703–708.
Phillips, P., 1991. Optimal inference in cointegrated systems. Econometrica 59 (2), 283–306.
Phillips, P.C.B., 2014. On confidence intervals for autoregressive roots and predictive regression. Econometrica 82 (3), 1177–1195.
Phillips, P.C.B., 2015. Pitfalls and possibilities in predictive regression. J. Financ. Econom. 13 (3), 521–555.
Phillips, P.C.B., Hansen, B.E., 1990. Statistical inference in instrumental variables regression with I(1) processes. Rev. Econom. Stud. 57 (1), 99–125.
Phillips, P.C.B., Lee, J.H., 2013. Predictive regression under various degrees of persistence and robust long-horizon regression. J. Econometrics 177 (2), 250–264.
Phillips, P.C.B., Magdalinos, T., 2009. Econometric inference in the vicinity of unity. CoFie Working Paper 7, Singapore Management University.
Rogoff, K., 2003. Globalization and global disinflation. Econ. Rev. 88 (Q IV), 45–78.
Rossi, B., 2013. Exchange rate predictability. J. Econ. Lit. 51 (4), 1063–1119.
Stambaugh, R.F., 1986. Bias in Regressions with Lagged Stochastic Regressors. CRSP Working Paper No 156, University of Chicago.
Stambaugh, R.F., 1999. Predictive regressions. J. Financ. Econ. 54 (3), 375–421.
Valkanov, R., 2003. Long-horizon regressions: Theoretical results and applications. J. Financ. Econ. 68 (2), 201–232.
Wei, M., Wright, J.H., 2013. Reverse regressions and long-horizon forecasting. J. Appl. Econometrics 28 (3), 353–371.
Weigand, R.A., Irons, R., 2007. The market p/e ratio, earnings trends, and stock return forecasts. J. Portf. Manag. 33 (4), 87–101.
Welch, I., Goyal, A., 2008. A comprehensive look at the empirical performance of equity premium prediction. Rev. Financ. Stud. 21 (4), 1455–1508.
Xu, K.-L., 2020. Testing for multiple-horizon predictability: Direct regression based versus implication based. Rev. Financ. Stud. 33 (9), 4403–4443.