Detecting the Predictive Power of Imperfect Predictors with Smoothly Varying Components *

Matei Demetrescu
TU Dortmund University

Mehdi Hosseinkouchak
EBS University†

Preliminary version: February 24, 2022
Please do not quote

Abstract

The typical predictor in predictive regressions for stock returns exhibits high persistence, which leads to nonstandard limiting distributions of the least-squares estimator and the associated t statistic. While there are several methods to deal with the issue of nonstandard distributions, the high predictor persistence also opens the door to spurious regression findings induced by the use of imperfect predictors, i.e. when the predictors do not perfectly span the conditional mean of the stock returns. To deal with such imperfect predictors, we take here a technical approach. Concretely, we robustify IVX predictive regression (Kostakis et al., 2015, Review of Financial Studies 28, 1506–1553) to the presence of *smoothly varying* components of the predictive system. This allows us to deal with situations where the predictors are imperfect without requiring additional knowledge on the predictive system, which is often unavailable in practice. In specific, we employ a filter which effectively employs smoothness to identify the mean component of the stock returns unaccounted for by the imperfect predictors. The limiting distribution of the resulting modified IVX t statistic is derived under sequences of local alternatives, and a wild bootstrap implementation improving the finite-sample behavior is provided. Compared to standard IVX predictive regression, there is a price to pay for such robustness in terms of power; at the same time, the IVX statistic without adjustment consistently rejects the false null of no predictability in the presence of imperfect predictors.

Keywords: Stock return predictability; Persistence; Imperfect predictors; IVX-based inference.
JEL classification: C12 (Hypothesis Testing), C22 (Time-Series Models), G17 (Financial Forecasting and Simulation)

*The authors would like to thank Peter Boswijk, Jörg Breitung, Robin Kruse-Becher, Tassos Magdalinos, Rob Taylor and Dominik Wied for very helpful comments and suggestions.

†Gustav-Stresemann-Ring 3, 65189 Wiesbaden, Germany, email: mehdi.hosseinkouchack@ebs.edu

1 Introduction

If the literature on predictive regressions for stock returns has taught us something the past decades, it is that predictability of stock returns is difficult to confirm empirically. Although many putative predictors such as the dividend yield or the earnings/price ratio are well rooted in economic theory, and are often found to be quite persistent in practice (implying high convergence rates of the slope coefficient estimators), the signal-to-noise ratio of the typical predictive regression is low, reducing chances of establishing predictability. Furthermore, model instabilities cause additional complications in applied work, and, finally, one also faces technical challenges, most prominently so-called predictive regression endogeneity inducing 2nd order bias of the LS estimators and nonstandard limiting null distributions of the associated t statistic. See Campbell (2008) and Phillips (2015) for reviews of predictive regressions for stock returns.

Various ways to deal with the technical aspects of predictive regressions have been put forward in the past years; see, among others, Campbell and Yogo (2006); Jansson and Moreira (2006); Bauer and Maynard (2012); Elliott et al. (2015); Kostakis et al. (2015); Breitung and Demetrescu (2015). In particular, the IV estimation procedures proposed by Kostakis et al. (2015) and Breitung and Demetrescu (2015) have become popular; see e.g. Phillips and Lee (2013), Lee (2016), Pavlidis et al. (2017), Gonzalo and Pitarakis (2017), Yang et al. (2020), Demetrescu et al. (2022b), Demetrescu and Hillmann (2022) or Demetrescu et al. (2022a), where Demetrescu et al. (2022b,a) specifically discuss IV-based inference in the presence of instability of the slope coefficients. In particular, these IV methods do not require additional variables to be observed since they rely on self-generated instrumental variables.

However, even if predictability by certain variables is found, the high persistence of the typical putative predictor may prevent a "clean" interpretation of the outcomes of predictive regressions. In particular, Ferson et al. (2003) point out that persistent components in the mean of stock returns are going to be detected by standard predictive regressions solely by virtue of the high persistence of say financial ratios. In other words, the danger of spurious regression with persistent variables looms in predictive regressions for stock returns. Motivated by such concerns, Georgiev et al. (2019) provide a conditional bootstrap test to check whether a predictive regression is spurious or not; in econometric parlance, this is nothing else than a test for omitted (persistent) variables, or for omitted variable bias. A further, related empirical problem involving persistence is the presence of non-predictive components of the putative predictors. Just like the mean component of stock returns may contain a persistent component not spanned by the putative predictors, one may argue that predictors also contain components that are not present in the predictable component of stock returns, if predictability is given. In technical terms, this translates to errors in variables. Under certain circumstances, such errors in variables are not problematic, e.g. when the persitence of the non-informative component is lower than that of the predictive signal. Yet it is rather the case that the mean-reverting component of regressors rather than their highly persistent trend component has predictive power; see Lettau and Nieuwerburgh (2008). And when the non-predictive component is at least as persistent as the predictive signal, this would cause spurious regressions again. Andersen and Varneskov (2021) provide a thorough discussion in a predictive regression setup with fractionally integrated predictors.

Pástor and Stambaugh (2009) coin the term "imperfect predictors" for such predictive systems, where imperfection stems, in a nutshell, from not observing the correct predictors, either due to measurement error or by misssing a predictive component altogether. In such situations, instrumental variable methods would normally be used as a workaround. This luxury is however not available in forecasting practice, where typically all available predictors are employed in the predictive regression and no instrument is available.

Here, the issue of imperfect regressors is tackled by means of a filtering approach. In specific, we argue that the problem of imperfect regressors may be side-stepped if the troublesome components vary smoothly enough. First, we show that, smoothness of the troublesome components given, IVX based tests will by construction remove smooth components of the putative predictors, leaving only the stochastic, possibly mean-reverting component as predictor. Second, we introduce a demeaning scheme tailored for eliminating smoothly varying components of stock returns such that, in conjunction with IVX based predictability testing, we obtain a robust inferential method for stock return predictability in the presence of imperfect persistent predictors with noninformative smooth components.

The remainder of this paper is organized as follows. After introducing the model and discussing the consequences of ignoring time-varying mean components in Section 2, we introduce our method of mean adjustment in Section 3 and establish its asymptotic validity, together with the validity of a bootstrap implementation. The finite-sample properties of the proposed implementation if IVX predictive regression are analyzed via Monte Carlo simulation in Section 4, and the final section concludes. Profs and additional material have been gathered in the appendix.

2 Setup

Beginning in a univariate framework, let

$$y_t = \alpha + \beta x_{t-1} + u_t, \qquad t = 2, \dots, T \tag{1}$$

where $\beta = 0$ under the null of no predictability. The distinctive feature here is that the intercept possibly varies in time, $\alpha = \alpha_t$, thus modelling the situation where the predictors don't perfectly span the conditional mean of the stock returns.

Assumption 1 *Let* $\alpha_t = \mu\left(t/T\right)$ *where* $\mu(\cdot)$ *is smooth with Lipschitz-continuous derivative on* $[0, 1]$.

This time variation, unaccounted for, makes the regression imperfect. As we shall see, the smooth variation assumption as captured by the Lipschitz-continuity condition is essential for identification.

The usual component structure of the regressor is assumed,

$$x_t = \mu_{x,t} + \xi_t, \qquad \xi_t = \rho\xi_{t-1} + \Psi(L)v_t, \tag{2}$$

(with Ψ a lag polynomial) where we allow for persistence (via suitable choice of ρ and short run dynamics as captured by Ψ) and contemporaneous correlation between the errors u_t and v_t (i.e. what is often called endogeneity in the predictive regression literature). More precisely, we take the usual near-to-unity specification (cf. Campbell and Yogo, 2006), but, following Demetrescu and Rodrigues (2022), allow the strength of mean reversion to be time-varying as well:

Assumption 2 *Let* $\rho = \rho_t = 1 - c_t/T$ *where* $c_t = c\left(t/T\right)$ *with* $c(\cdot)$ *a piecewise Lipschitz continuous function.*

Remark 2.1 *A large part of the literature on predictive regressions assumes autoregressive coefficients belonging to the stationarity region, i.e.* $|\rho|$ *bounded below unity; see e.g. Amihud and Hurvich (2004); Amihud et al. (2009). We posit however that standard asymptotics results in that case and we do not deal with this case.*

Assumption 3 *Let* $\mu_{x,t} = \sqrt{T}\mu_x\left(t/T\right)$ *where* $\mu_x(\cdot)$ *is smooth with Lipschitz-continuous derivative on* $[0, 1]$.

This assumption ensures that the mean component of the putative predictor is not dominated in the limit by the stochastic component.

To allow for a more realistic modelling of the predictor, we allow for short-run dynamics and heterogeneity as follows.

Assumption 4 *The lag polynomial* $\Psi(L)$ *has* 1*-summable coefficients,* $\sum_{j\geq 0} j\,|\psi_j| < \infty$*, such that* $\psi = \Psi(1) \neq 0$.

These are typical assumptions for models involving near-integrated variables. Moreover, the error variances and covariances are allowed to vary in time as well:

Assumption 5 *Let* $(u_t, v_t)' = \mathbf{H}(t/T)(a_t, e_t)'$ *with* $\mathbf{H}$ *a matrix of piecewise Lipschitz continuous functions and* $(a_t, e_t)'$ *white noise as specified below.*

Assumption 6 *Let* $(a_t, e_t)'$ *be a uniformly* $L_{4+\delta}$ *bounded, zero-mean serially independent heterogeneous sequence with unity covariance matrix,* $\operatorname{Cov}\left((a_t, e_t)'\right) = \mathbf{I}_2$.

The assumption allows for time-varying variances and covariances; cf. e.g. Demetrescu et al. (2022b). The iid assumption may be relaxed at the cost of additional technical details, like in Demetrescu et al. (2022a), but we omit the details to save space.

Under the above assumptions, we have (with $\boldsymbol{W}$ a vector of two independent standard Wiener processes) that

$$\frac{1}{\sqrt{T}} \sum_{t=1}^{[sT]} \begin{pmatrix} u_t \\ v_t \end{pmatrix} \Rightarrow \int_0^s \mathbf{H}(s)\,\mathrm{d}\boldsymbol{W}(s) := \begin{pmatrix} M_u(s) \\ M_v(s) \end{pmatrix},$$

(see e.g. Cavaliere et al., 2010) and consequently

$$\frac{1}{\sqrt{T}} x_{[sT]} \Rightarrow \mu_x(s) + \psi \int_0^s e^{-\int_r^s c(t)\mathrm{d}t} \mathrm{d}M_v(r) := \psi J_{c,\mathrm{H}}(s),$$

which is an Ornstein-Uhlenbeck type process driven by $M_v(s)$ with a drift component. We note the presence of the persistent component μ_x in this limit.

For convenience, let $\sigma_u^2(s) = \frac{\mathrm{d}}{\mathrm{d}s}[M_u](s)$ with $[M_u](\cdot)$ denoting the quadratic variation process of M_u. This plays the role of the instantaneous variance of the shocks, since $\operatorname{Var}(u_t) = \sigma_u^2(t/T) + O(T^{-1})$; define $\sigma_v^2(s)$ analogously.

Even without a time-varying intercept α_t, inference in the case of near-integrated regressors is not straightforward to conduct because of the endogeneity issue. To circumvent this difficulty, we follow Kostakis et al. (2015) and confine ourselves to IVX estimation and testing of the predictive regression in Eq. (1); see also Breitung and Demetrescu (2015). IVX amounts to instrumenting the regressor by a filter thereof,

$$z_{t-1} = (1 - \varrho L)_+^{-1} \Delta x_{t-1} = \sum_{j=0}^{t-2} \varrho^j \Delta x_{t-j-1},$$

where $\varrho = 1 - \frac{a}{T^\eta}$ with $\eta \in (0, 1)$. This leads to the IVX-based t statistic

$$t_{vx} = \frac{\sum_{t=2}^T (z_{t-1} - \bar{z})(y_t - \bar{y})}{\sqrt{\sum_{t=2}^T (z_{t-1} - \bar{z})^2 (y_t - \bar{y})^2}}, \tag{3}$$

where the heteroskedasticity-robust standard errors have been computed under the null for simplicity. (This typically does not make a difference asymptotically under local alternatives; see e.g. the discussion in Demetrescu et al., 2022b.) In the constant-intercept case, this has an asymptotically standard normal null distribution and has power against local alternatives in $T^{-1/2-\eta/2}$-neighbourhoods of the null.

Here, however, the time-varying mean component $\mu(s)$ spuriously correlates with $J_{c,\mathrm{H}}(s)$ (even if the correlation is random), which leads to spurious rejections of the null $\beta = 0$. While this holds more generally (see the discussion in Georgiev et al., 2019), we confirm that IVX statistic will diverge in the presence of non-constant mean components and consistently reject the true null:

Proposition 1 *Under Assumptions 1–6 and the null $\beta = 0$, we have as $T \to \infty$ that*

$$\frac{1}{T^{\eta/2}} t_{vx} \stackrel{p}{\to} \sqrt{\frac{2}{a}} \frac{\int_0^1 (\mu(s) - \bar{\mu}) \, \mathrm{d}M_v(s)}{\sqrt{\int_0^1 \sigma_v^2(s) \sigma_u^2(s) \, \mathrm{d}s + \int_0^1 \sigma_v^2(s) (\mu(s) - \bar{\mu})^2 \, \mathrm{d}s}},$$

where $\bar{\mu} = \int_0^1 \mu(r) \mathrm{d}r$.

Proof: *See Appendix B.*

Remark 2.2 *The smoothly varying regressor component μ_x does not appear in the limit, so IVX is successfully filtering it away, and one of the two issues with imperfect regressors is solved. But the mean component of y_t still causes problems. Clearly, this is because the usual demeaning, $y_t - \bar{y}$ is only appropriate when $\alpha_t = const$. The following section discusses suitable filtering procedures for the smooth component of the stock returns not spanned by the predictors.*

3 A suitable filter for the stock returns

3.1 Preliminaries

To prevent the danger of spurious regressions due to imperfect predictors as captured by a time-varying intercept, one must remove α_t before running the IVX predictive regression.

Should the mean function μ be piecewise constant, i.e. exhibit sudden breaks only, it suggests itself to model the breaks explicitly. While this requires an estimation of the number and the location of the breaks, this is well-understood. We find however that it does not perform well, and therefore resort to adjustment methods that are less parametric in nature.

The difficulty here is to find a procedure that removes the mean component but does not completely eliminate the signal βx_{t-1}, and, at the same time, does not essentially affect the IVX methodology. This is indeed not trivial at all. For instance, take differencing, such that

$$\Delta y_t = \beta \Delta x_{t-1} + \Delta u_t + O\left(\frac{1}{T}\right).$$

In this transformed model, the errors have a MA(1) structure and, due to contemporaneous correlation between u_t and v_t, they will correlate with Δx_{t-1}. While this could be side-stepped by instrumenting Δx_{t-1} by x_{t-2} or Δx_{t-2}, we note that IV estimation using lagged levels or differences of x_{t-1} actually runs into the problem of weak instruments since x_t is near-integrated. And even if not, estimating in differences leads to $\sqrt{T}$-consistency of the slope coefficient estimator since Δx_{t-1} is not near integrated anymore. Furthermore, local demeaning, where we subtract from y_t the average of the previous τ observations, leads to

$$y_t - \frac{1}{\tau}\sum_{j=1}^{\tau} y_{t-j} = \beta\left(x_t - \frac{1}{\tau}\sum_{j=1}^{\tau} x_{t-1-j}\right) + \left(u_t - \frac{1}{\tau}\sum_{j=1}^{\tau} u_{t-1-j}\right) + O\left(\frac{\tau}{T}\right).$$

For suitable choices of $\tau \to \infty$ while $\tau/T \to 0$, the implied errors are approximately white noise, $u_t - \frac{1}{\tau}\sum_{j=1}^{\tau} u_{t-j} \approx u_t$. This does side-step the issue of serial error correlation, but again at the cost of weakening the signal: under near integration, $\frac{1}{\tau}\sum_{j=1}^{\tau} x_{t-1-j} \approx x_t$ and the locally demeaned x_t is effectively overdifferenced again, thus reducing the regression signal again; see the similar discussion in Section 2.3 of Demetrescu and Hosseinkouchack (2021).

We therefore resort to a demeaning procedure that interacts with α_t and x_{t-1} *in a different manner*. This is where the the smoothness of $\mu(\cdot)$ comes into play. Concretely, we exploit it to distinguish between the time-varying mean and the signal βx_{t-1}, which, being near-integrated, is essentially more rough (the Wiener process for instance is only Hölder continuous for any Hölder coefficient less than 1/2 compared to the Lipschitz property of μ, i.e. a Hölder coefficient of unity).

3.2 Nonparametric adjustment

To this end, consider the exponentially smoothed estimate of α_t given for $t = 2, \ldots, T$ by

$$\bar{y}_t = (1 - \pi)\, y_{t-1} + \pi \bar{y}_{t-1} = \hat{\alpha}_t,$$

with $\bar{y}_1 = y_1$. While one could explore alternative ways of adjusting for a time-varying mean that do not imply over-differencing (e.g. splines regression), exponential smoothing is a well-established method in time series analysis. It is a particualr case of the Kalman filter, and has computational advantages in that it can be represented with the help of autoregressive filters.

The mean-adjusted dependent variable is then given by

$$\tilde{y}_t = y_t - \bar{y}_t = y_t - (1-\pi)\sum_{j=0}^{t-2}\pi^j y_{t-j-1} - \pi^{t-1}y_1.$$

For $\pi = 0$, one recovers differencing, $\tilde{y}_t = y_t - y_{t-1}$, while, for $\pi = 1$, we essentially subtract the initial observation y_1 – which is counterproductive, given that y_t is (locally) stationary. The trick is to use a coefficient π close, but not too close to, unity. Intuitively, this places weight on enough observations to ensure a good estimation of the mean, while still remaining local in nature.

Concretely, we pick

$$\pi = 1 - \frac{p}{T^\gamma}, \qquad \gamma \in (0,1),$$

where we will impose some additional restrictions on γ. This IVX analog reduces persistence of both signal and time-varying mean, but in a different manner – such that one may still distinguish between them reliably.

The modified IVX t-statistic is then

$$\tilde{t}_{vx} = \frac{\sum_{t=2}^T z_{t-1}\tilde{y}_t}{\sqrt{\sum_{t=2}^T z_{t-1}^2\tilde{y}_t^2}};$$

we only need to mean-adjust the dependent variable, since z_t is invariant to μ_x by construction. The statistic is asymptotically standard normal just like the usual IVX statistic when the intercept is constant, as shown in

Proposition 2 *Let Assumptions 1–6 hold. Under the sequence of local alternatives* $\beta = \frac{b}{T^{1/2+\eta/2}}$, *we have as* $T \to \infty$ *that*

$$\tilde{t}_{vx} \stackrel{d}{\to} \mathcal{Z} + b\psi\sqrt{\frac{2}{a}}\frac{1+p}{p}\frac{J_{c,H}^2(1) - \int_0^1 J_{c,H}(s)\,\mathrm{d}J_{c,\mathrm{H}}(s)}{\sqrt{\int_0^1 \sigma_v^2(s)\,\sigma_u^2(s)\,\mathrm{d}s}},$$

if $1/2 + \eta/2 < \gamma < 1 - \eta$, *where* $\mathcal{Z}$ *is a standard normal variate.*

Proof: *See Appendix B.*

Remark 3.1 *If the conditional covariance of* u_t *and* v_t *is only deterministically varying (as is the case with serial independence), it can be shown that* $\mathcal{Z}$ *is independent of* $J_{c,\mathrm{H}}(s)$.

Remark 3.2 *The modified IVX based test has power under the same kind of local alternatives as the original IVX statistic, i.e. in* $T^{-1/2-\eta/2}$ *neighbourhoods of the null. We do note a loss of local power, however: it can be seen that Proposition 2 requires* $\eta < 1/3$ *in any case, whereas only* $\eta < 1$ *is needed for the original IVX. But the original IVX only works for constant intercept, so the loss of some local power buys robustness.*

Remark 3.3 *A natural question is hence whether these restrictions could be relaxed so as to increase η could favor of some power gains. Note that, from the proof of proposition 2 we learn that $\sum_{t=2}^{T} z_{t-1}^2 \tilde{y}_t^2 = O_p(T^{1+\eta})$ which is the denominator of $\tilde{t}_{vx}$, and this is regardless of the restrictions placed on η and γ. Therefore, the numerator of the $\tilde{t}_{vx}$ must necessarily be $O_p\big(T^{1/2+\eta/2}\big)$. This can guaranteed only if $\gamma > 1/2 + \eta/2$ and $\gamma < 1 - \eta$, otherwise $\tilde{t}_{vx}$ either vanishes or diverges as $T \to \infty$.*

Remark 3.4 *The procedure is applied without any essential modification in the multiple regression case. Should there be K near-integrated regresssors of interest, each regressor x_{kt-1} is instrumented using "its own" IVX instrument $z_{kt-1} = (1 - \varrho L)_{+}^{-1} \Delta x_{kt-1}$, $k = 1, \ldots, K$, whereas the dependent variable is adjusted for a time-varying mean in the exact same manner as discusssed above. The analogous result to Proposition 2 follows imemdiately and we do not include the details.*

3.3 A bootstrap implementation

Preliminary simulations (not reported here) show that the statistic is not close enough to standard normality in finite samples; this is not surprising, given that the usual IVX statistic is not behaving too well in finite samples either; see Kostakis et al. (2015), Demetrescu and Hosseinkouchack (2021) and Demetrescu et al. (2022a). Demetrescu et al. (2022a) propose a wild bootstrap implementation of IVX regressions which largely eliminates such problems; adapted to our situation, the algorithm is as follows.

1. Get residuals $\hat{v}_t$ from an autoregression of order k where k is chosen using AIC with a maximum lag length of $4\lfloor (T/100)^{0.25} \rfloor$, i.e. $\hat{v}_t = x_t - \hat{\mu}_x - \sum_{j=1}^{k} \hat{\theta}_j x_{t-j}$ with back-ward demeaning of x_t.

2. Get residuals $\hat{u}_t = \tilde{y}_t$ (i.e. under the null).

3. Get scalar bootstrap multipliers R_t.

4. Generate $u_t^* = R_t \hat{u}_t$ and $v_t^* = R_t \hat{v}_t$.

5. Recolor v_t^* to obtain x_t^* using the estimated coefficients from step 1 above.

6. Add estimated trend $y_t^* = u_t^* + \bar{y}_t$ to obtain bootstrap data satisfying the null hypothesis.

7. Define $\tilde{t}_{vx}^*$ the same way as $\tilde{t}_{vx}$, but on the basis of the adjusted $\tilde{y}_t^*$ and x_t^*.

8. Use quantiles of $\tilde{t}_{vx}^*$ rather than those of the standard normal for inference.

This bootstrap is consistent in the sense that the bootstrapped $\tilde{t}_{vx}^*$ converges weakly in probability to the same limit as $\tilde{t}_{vx}$ so the use of the bootstrap critical values is justified.

Proposition 3 *Under the assumptions of Proposition 2, we have as $T \to \infty$ that*

$$\tilde{t}^*_{vx} \overset{p}{\Rightarrow} \mathcal{Z}.$$

Proof: *See Appendix B.*

This justifies the use of bootstrap critical values for $\tilde{t}_{vx}$. The behavior of this test is studied in the following section.

4 Finite-sample behavior

We generate the putative predictor, x_t, based on equation (2) with $\rho = 1 - c/T$ when $c \in \{0, 1, 5, 10, 30, 50\}$ and $\Psi(L) = (1 - \phi L)^{-1}$ with $\phi = 0$ or 0.5. Furthermore, the shocks are iid following $(v_t, u_t)' \sim \mathcal{N}(0, \Sigma)$ where $\Sigma = (1, -0.95; -0.95, 1)$. We generate the regressand based on equation (1) where

$$\alpha_t \quad = \quad \alpha_\mu \left(\tanh \left(5 \left(\frac{t}{T} - \frac{1}{2} \right) \right) + \frac{1}{2} \right),$$

represents a smoothly varying mean for y_t for which $\alpha_\mu \in \{0, \frac{1}{4}, \frac{1}{2}\}$ controls the strength of the deterministic component of y_t. Alternatively, we have a sudden break,

$$\alpha_t = \tau \mathbb{I}_{\{t > \lfloor \tau T \rfloor\}},$$

for $\tau \in \{0.3, 0.5, 0.7\}$.

The first set of results, given in Table 1, shows the size properties of the bootstrap version of the proposed test statistic, $\tilde{t}^*_{vx}$, compared to the original IVX test statistic, t^W_{vx}, and an IVX test statistic computed using the series y_t adjusted for breaks, t^{WBP}_{vx}. To adjust for breaks, we used Bai and Perron (1998) to fit the break times, which allows for detecting multiple breaks in the model parameters, particularly for the mean of y_t in our specification. Since there is a shift in the mean of y_t, the hope is that applying Bai and Perron (1998) would potentially alleviate the adverse effect of leaving the mean change unattended.

The results are generated using 5,000 replications with $\beta = 0$ and $T = 250$ or $T = 500$. For the smoothing step we consider $\pi = 1 - 1/T^{0.75}$, i.e. $p = 1$ and $\gamma = 0.75$. We generate the IVX instrument for $\tilde{t}_{vx}$ using $\varrho = 1 - 1/T^{0.15}$, i.e. $a = 1$ and $\eta = 0.15$, to stay conformable with the parameter restrictions required by Proposition 2, while to construct the instrument for t^W_{vx} and t^{WBP}_{vx} we set $\varrho = 1 - 1/T^{0.95}$, as suggested by Kostakis et al. (2015).

Table 1: Size properties of different tests under short-run dynamics and strong contemporaneous shock correlation

		$T=250$									$T=500$								
		2-sided			left-sided			right-sided			2-sided			left-sided			right-sided		
α_μ	c	t_{vx}^{W}	t_{vx}^{WBP}	$\tilde{t}_{vx}^{*}$	t_{vx}^{W}	t_{vx}^{WBP}	$\tilde{t}_{vx}^{*}$	t_{vx}^{W}	t_{vx}^{WBP}	$\tilde{t}_{vx}^{*}$	t_{vx}^{W}	t_{vx}^{WBP}	$\tilde{t}_{vx}^{*}$	t_{vx}^{W}	t_{vx}^{WBP}	$\tilde{t}_{vx}^{*}$	t_{vx}^{W}	t_{vx}^{WBP}	$\tilde{t}_{vx}^{*}$
0	0	4.44	4.68	4.98	0.04	0.04	1.04	8.66	9.38	7.82	4.34	4.58	4.18	0.10	0.04	1.74	8.98	8.92	6.24
	1	4.96	5.14	4.40	0.14	0.26	1.70	9.68	9.92	7.30	4.16	4.72	4.58	0.20	0.14	2.58	8.72	9.40	5.82
	5	5.66	5.54	5.58	0.90	1.18	3.98	10.02	9.72	6.70	5.38	4.80	5.40	0.86	0.90	4.16	10.02	9.74	6.66
	10	5.92	6.10	5.56	1.66	2.08	4.44	9.70	10.00	6.04	5.64	5.98	5.82	1.52	1.80	4.78	9.34	9.64	6.14
	30	6.24	6.08	5.80	3.40	3.14	4.92	8.40	8.58	5.68	5.76	6.32	6.64	2.86	2.96	5.60	7.96	8.26	5.90
	50	5.86	5.64	5.78	3.14	3.06	5.16	7.96	7.68	5.68	4.74	5.44	5.82	2.62	2.84	5.50	7.42	7.60	5.32
0.25	0	10.26	10.34	4.76	0.62	0.66	1.24	16.46	17.04	7.38	13.80	15.64	4.20	1.90	2.32	1.90	20.26	21.38	6.24
	1	10.82	10.86	5.38	1.50	1.72	2.22	17.34	17.34	7.80	15.64	16.06	4.90	2.84	3.62	2.26	21.76	22.54	7.24
	5	10.32	10.26	5.76	3.38	4.12	3.96	14.84	14.64	6.64	16.44	16.14	4.46	6.82	6.88	3.92	19.96	18.54	5.56
	10	10.48	10.32	5.64	4.22	4.00	4.38	13.74	13.98	6.12	13.02	12.96	5.56	5.92	5.92	4.68	15.36	16.02	6.06
	30	7.94	7.06	5.64	3.72	4.06	4.90	10.06	9.30	5.72	8.58	8.88	5.36	4.36	5.06	5.00	10.38	10.22	5.22
	50	6.36	7.32	5.60	3.36	4.16	5.08	8.08	8.84	5.50	7.70	8.10	5.48	5.04	4.62	4.98	9.64	9.50	5.68
0.5	0	25.80	27.68	4.98	5.30	6.34	1.30	29.84	31.74	7.34	38.60	40.32	4.80	11.62	12.36	2.00	36.32	37.00	6.94
	1	29.32	30.78	4.82	6.82	8.82	1.92	31.60	32.40	7.42	39.58	42.94	4.38	13.68	14.46	2.40	34.72	37.36	6.20
	5	27.16	28.76	5.46	10.64	11.08	3.56	26.84	27.28	6.90	41.26	42.74	5.02	16.02	16.46	3.74	34.02	35.44	6.12
	10	21.10	22.92	5.58	9.02	10.30	4.50	22.12	22.94	6.18	32.46	33.60	5.06	14.26	14.86	4.34	27.36	28.96	5.94
	30	12.42	13.88	6.00	6.54	7.20	5.32	12.86	13.96	5.88	18.00	18.18	5.20	9.28	9.52	5.12	17.14	16.68	5.38
	50	10.24	11.28	5.90	5.76	6.88	5.32	10.86	11.34	5.90	13.92	13.64	5.56	8.64	8.00	4.76	12.46	12.92	6.18
τ																			
0.3	0	23.34	25.06	6.52	3.16	4.30	2.20	29.88	31.00	9.84	35.82	36.82	6.52	8.98	10.74	3.42	36.36	35.54	8.20
	1	24.04	25.24	7.04	5.44	6.44	3.20	28.18	28.80	8.98	36.98	37.86	5.64	11.00	12.64	3.34	34.84	33.96	7.38
	5	26.42	27.76	6.14	9.68	10.36	4.30	26.84	26.78	6.96	38.24	40.64	6.10	15.10	16.00	4.36	31.52	33.12	6.66
	10	23.18	25.16	5.66	9.84	11.22	4.68	22.66	23.60	6.16	34.36	36.18	5.80	15.82	15.60	4.46	28.24	29.06	6.28
	30	14.28	15.24	5.34	7.76	8.64	4.64	14.48	14.48	5.94	21.68	22.60	5.44	11.84	12.58	5.32	18.56	19.16	5.12
	50	11.84	12.24	6.10	6.82	7.36	5.04	11.54	12.66	5.72	14.98	16.58	4.92	9.52	9.86	4.80	14.06	14.70	4.94
0.5	0	32.20	35.26	6.32	8.34	10.36	2.30	32.86	34.56	8.98	45.26	47.86	6.44	16.02	16.52	2.70	37.56	39.50	8.52
	1	36.04	37.58	7.36	9.94	12.32	3.34	34.34	34.76	9.50	49.06	51.56	6.08	16.54	19.06	3.30	39.72	39.42	8.20
	5	36.18	39.66	6.24	13.98	14.70	4.18	32.06	34.34	7.02	50.36	53.54	5.48	20.20	21.68	4.16	37.40	39.52	6.40
	10	29.26	32.02	5.86	12.44	12.90	4.10	25.96	27.70	6.44	43.56	43.02	5.52	18.58	16.54	4.58	33.48	32.42	6.24
	30	16.76	17.22	5.44	8.48	9.96	4.48	16.28	15.88	5.56	23.84	25.88	5.42	12.20	12.44	4.74	20.94	22.26	5.42
	50	12.12	14.06	5.46	6.72	8.06	5.20	12.52	13.42	5.32	17.16	19.18	5.36	9.86	10.40	5.24	15.36	17.46	5.32
0.7	0	28.96	31.62	6.94	6.66	7.72	2.40	31.16	33.38	9.60	40.58	43.00	6.44	12.70	13.66	2.46	35.90	37.00	8.78
	1	31.82	33.44	7.06	9.40	10.32	3.28	32.42	33.32	9.28	44.38	47.00	5.96	14.44	17.02	3.34	37.70	37.78	7.38
	5	31.60	33.34	6.70	11.34	12.80	4.46	30.52	30.98	7.38	44.34	47.66	6.54	18.60	19.44	4.68	34.16	36.28	7.16
	10	25.74	27.42	6.02	10.92	12.26	4.80	24.08	24.98	6.62	38.38	39.60	5.66	16.28	17.46	4.54	31.26	30.84	6.30
	30	15.38	16.54	5.42	8.48	8.26	4.94	14.92	15.76	4.94	21.90	22.90	5.72	11.30	12.42	5.34	18.96	19.62	5.32
	50	11.42	12.72	5.48	6.70	7.70	4.66	11.40	12.48	5.64	15.46	17.00	5.50	8.82	10.92	5.50	14.62	14.66	5.32

Note: Data generated with (1) and (2) with $v_t = \phi v_t + \nu_t$ for $\phi = 0.5$, where $(u_t, \nu_t) \sim iiN(0, \mathbf{\Sigma})$ and $\mathbf{\Sigma}$ exhibits constant correlation $\delta = -0.95$. We set $\rho = 1 - {}^{c}/_{T}$ for various c. t_{vx}^{W} is defined as in equation (3) for which we use the small sample corrections proposed by Kostakis et al. (2015) and $\tilde{t}_{vx}^{*}$ is the bootstrap implementation of the test with mean-adjusted y_t. For t_{vx}^{W} and t_{vx}^{WBP} we set $\varrho = 1 - {}^{1}/_{T^{0.95}}$. For $\tilde{t}_{vx}^{*}$ we set $\varrho = 1 - {}^{1}/_{T^{0.15}}$ and $\pi = 1 - {}^{1}/_{T^{0.75}}$.

From Table 1 one may note that, as predicted by Proposition 1, t_{vx}^{W} diverges with an increase in T when $\alpha_u \neq 0$, as does t_{vx}^{WBP}. The table also shows that $\tilde{t}_{vx}^{*}$ exhibits a good size control for almost all the parameter constellations considered here. This implies that the test statistic proposed in this paper is capable of controlling the adverse effect of the time varying component pestering the standard IVX procedure. As mentioned earlier, the smoothing step comes with a loss in the power compared to the usual IVX statistic.

To compare the power properties of $\tilde{t}_{vx}^{*}$ and t_{vx}^{W}, we consider a local alternative of the form $\beta = \frac{b}{T^{1/2+\eta/2}}$ where $b \in \{-2, -1.9, -1.8, ..., 2\}$ where, as mentioned before, for $\tilde{t}_{vx}^{*}$ we set $\eta = 0.15$ and keep $\eta = 0.95$ for t_{vx}^{W}. To save on space, we focus on $c \in \{0, 10, 30\}$ for power illustrations. Figures (1)–(3) show, respectively, the left, right, and the 2-sided power curves of $\tilde{t}_{vx}^{*}$ and t_{vx}^{W} where $\phi = 0.5$ and $T = 250$. As these figures show, $\tilde{t}_{vx}^{*}$ is outperformed in terms of power by t_{vx}^{W} when $\alpha_\mu = 0$. This reflects the power loss that one has to accept when one needs some insurance against a potentially present time varying mean component. At the same time we observe overrejections by t_{vx}^{W} when $\alpha_\mu \neq 0$.

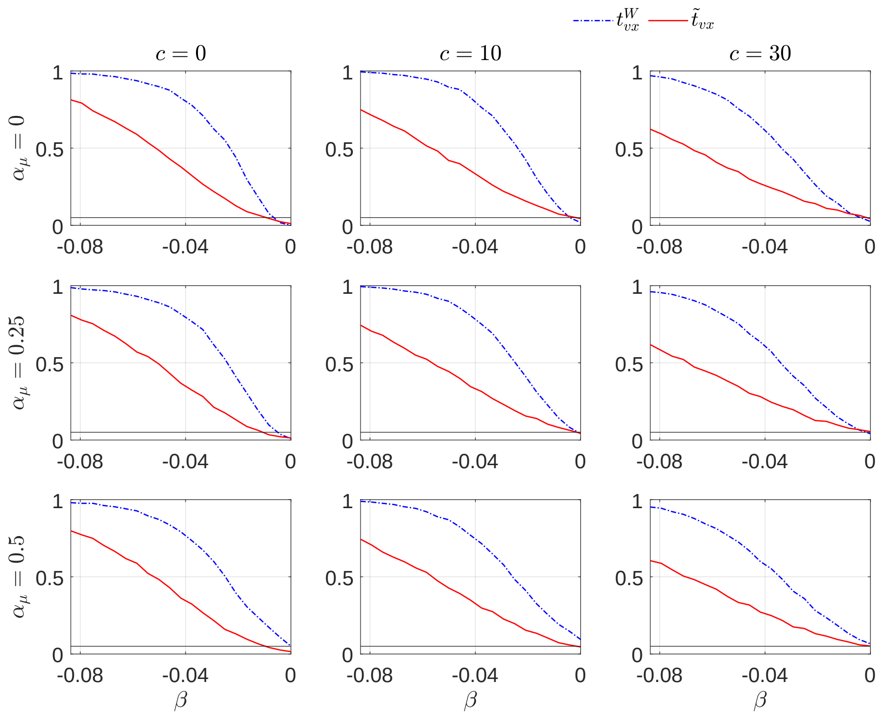

Figure 1: Left sided power curves for $\tilde{t}_{vx}$ (when $\gamma = 0.75$ and $\eta = 0.15$) and t_{vx}^{W} (when $\eta = 0.95$) under a local alternative. See the text for more details.

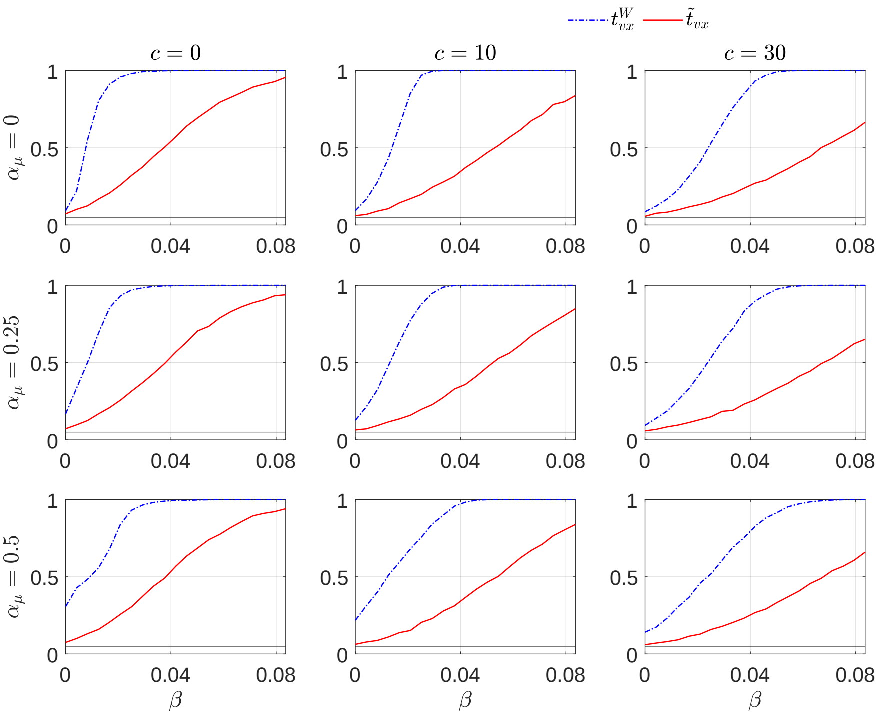

Figure 2: Right-sided power curves for $\tilde{t}_{vx}$ (when $\gamma = 0.75$ and $\eta = 0.15$) and t_{vx}^W (when $\eta = 0.95$) under a local alternative. See the text for more details.

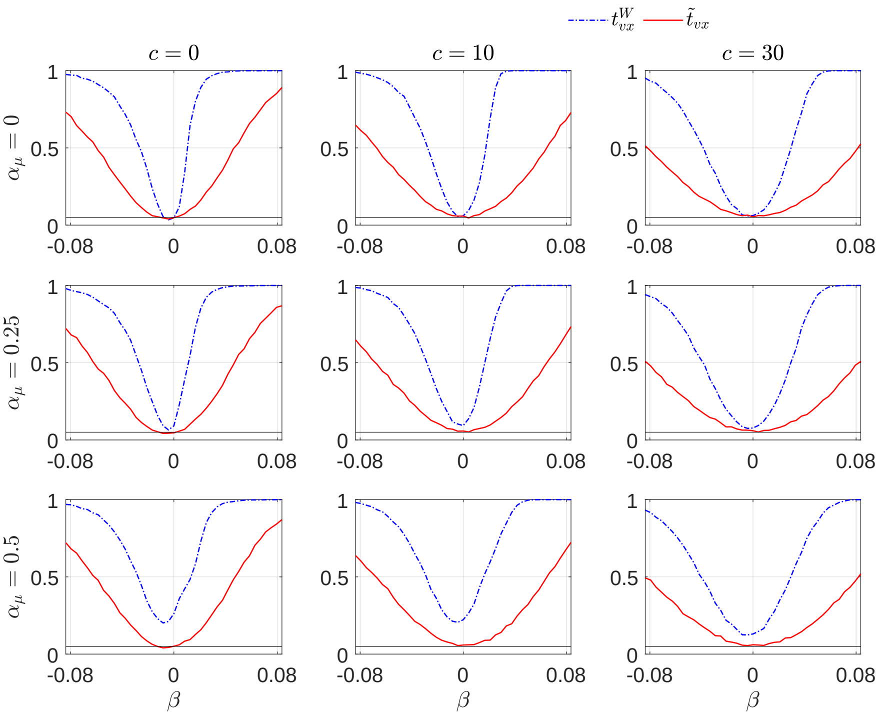

Figure 3: Two sided power curves for $\tilde{t}_{vx}$ (when $\gamma = 0.75$ and $\eta = 0.15$) and t_{vx}^W (when $\eta = 0.95$) under a local alternative. See the text for more details.

Figure 4 addresses the size-power trade-off. We notice that the largest effect on power is coming from the variation in η, which characterizes the instrument persistence. Varying γ does not have a visible effect on the power curves. Selecting other combinations did have a sizeable effect on size control, and we do not report them. The Appendix contains additional simulation results.

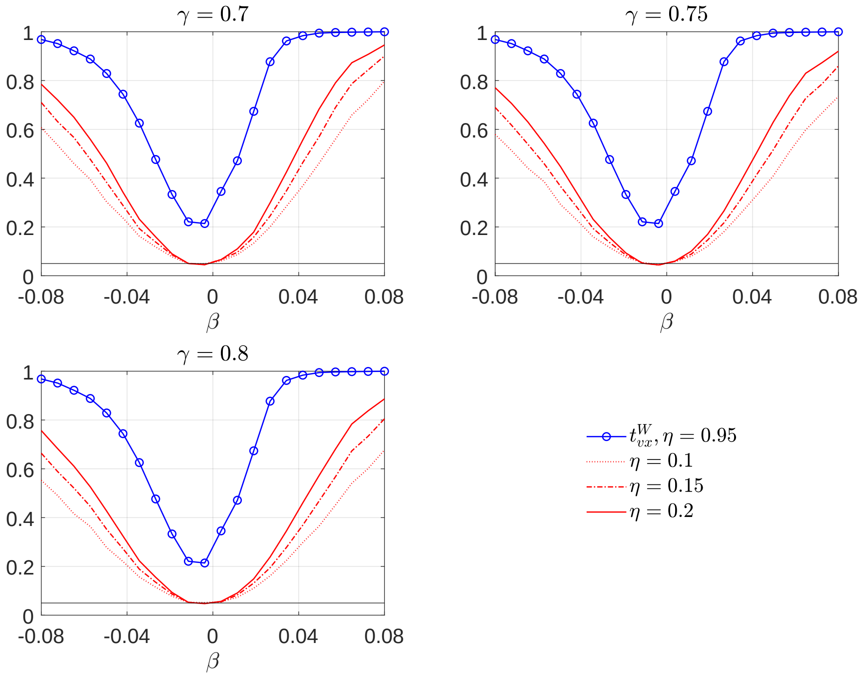

Figure 4: Two sided power curves for $\tilde{t}_{vx}$ and t^W_{vx} under a local alternative when $c = 0$ and $\alpha_\mu = 0.5$. The power curves for t^W_{vx} are the same for all the three sub figures, while power curves for $\tilde{t}_{vx}$ are calculated with different values for η and γ. See the text for more details.

Summing up, the proposed filtering procedure leads to robust inference, but does come at a price in terms of power loss. The power losses are minor compared to the gains in size control preventing spurious findings of predictability.

5 Concluding remarks

The thorny issue of inference on stock return predictability was addressed, when regressors are imperfect. To allow for inference without resorting to additional, often not available, information, we assume here that the problematic components of the conditional mean of the stock returns and of the predictors are smoothly varying in time. For such situations, we propose a filtering scheme that has the property of not affecting the limiting behavior of IVX test statistics.

The proposed adjustment scheme affects the finite-sample behavior of the IVX based test. Therefore, we employ a wild bootstrap implementation of the test, which performs well in simulations.

The adjustment leads to some loss of power under the alternative, which is however more than offset by the gain in size control under the null hypothesis. This prevents spurious predictive regression findings.

Appendix

A Auxiliary results

Let C denote a generic constant and note that for $t/T \to s > 0$, $\varrho^{kt} = \left(1 - \frac{a}{T^\eta}\right)^{kt} = \left(\left(1 - \frac{a}{T^\eta}\right)^{-\frac{T^\eta}{a}}\right)^{-ka\frac{t}{T}T^{1-\eta}} \to 0$ so $\sum_{j=0}^{t-1} \varrho^{kj} \sim \frac{1}{ka}T^\eta$ as $T \to \infty$. Clearly, this holds analogously for $\pi = 1 - \frac{p}{T^\gamma}$. Before moving on to the proofs, we state some preliminary results.

Lemma A.1 *Under the assumptions of Proposition 2, we have as $T \to \infty$ that*

1. $T^{-\gamma/2} \sum_{j=0}^{t-3} \pi^j u_{t-1-j}$ *is uniformly L_4-bounded;*
2. $\tilde{y}_t = \tilde{u}_t + \beta \tilde{x}_{t-1} + \tilde{\alpha}_t$, *where* $\tilde{u}_t = u_t - \frac{1}{T^\gamma} \sum_{j=0}^{t-2} \pi^j u_{t-1-j} - \pi^{t-1} u_1$, $\tilde{x}_{t-1} = x_{t-1} - \frac{1}{T^\gamma} \sum_{j=0}^{t-2} \pi^j x_{t-2-j} - \pi^{t-1} x_0$ *and* $\tilde{\alpha}_t = O\left(T^{\gamma-1}\right)$ *uniformly in* $t = 2, \ldots, T$.

Proof: *See Appendix B.*

Lemma A.2 *Under Assumptions 1–6, we have for $\eta, \gamma \in (0, 1)$ as $T \to \infty$ that*

1. $\sum_{t=2}^{T} \tilde{\alpha}_t z_{t-1} = O_p\left(T^{-1/2+\gamma+\eta}\right)$ *and* $\sum_{t=2}^{T} \pi^{t-1} z_{t-1} = O_p\left(T^{\gamma/2+\min\{\eta,\gamma\}}\right)$.
2. *Let* $r_{t-1} = \sum_{j=0}^{t-2} \pi^j u_{t-1-j}$, *then we have* $\sum_{t=2}^{T} r_{t-1} z_{t-1} = O_p\left(T^{1+\eta}\right)$ *whenever* $\eta < \gamma$.
3. $T^{-1-\eta} \sum z_{t-1} \tilde{x}_{t-1} \Rightarrow \psi^2 \frac{1+p}{ap}\left(J_{c,\mathrm{H}}^2(1) - \int J_{c,\mathrm{H}}(s)\, \mathrm{d}J_{c,\mathrm{H}}(s)\right)$.

Proof: *See Appendix B.*

B Proofs

Proof of Lemma A.1

1. We have that

$$\mathrm{E}\left(\left(T^{-\gamma/2} \sum_{j=0}^{t-3} \pi^j u_{t-1-j}\right)^4\right) = \frac{1}{T^{2\gamma}} \sum_{j_1=0}^{t-3} \sum_{j_2=0}^{t-3} \sum_{j_3=0}^{t-3} \sum_{j_4=0}^{t-3} \pi^{j_1}\pi^{j_2}\pi^{j_3}\pi^{j_4}\, \mathrm{E}\left(u_{t-1-j_1} u_{t-1-j_2} u_{t-1-j_3} u_{t-1-j_4}\right)$$

$$= O\left(\frac{1}{T^{2\gamma}}\sum_{j_1=0}^{t-3}\sum_{j_2=0}^{t-3}\pi^{2j_1}\pi^{2j_2} + \frac{1}{T^{2\gamma}}\sum_{j_1=0}^{t-3}\pi^{4j_1}\right)$$

since u_t is uniformly L_4-bounded and the expectation is nonzero only for equal or pairwise equal indices. Therefore,

$$\mathrm{E}\left(\left(T^{-\gamma/2}\sum_{j=0}^{t-3}\pi^j u_{t-1-j}\right)^4\right) = O\left(\frac{1}{T^{2\gamma}}\sum_{j_1=0}^{t-3}\sum_{j_2=0}^{t-3}\pi^{2j_1}\pi^{2j_2}\right) = O\left(\frac{1}{T^{2\gamma}}\left(\sum_{j_1=0}^{t-3}\pi^{2j_1}\right)^2\right) = O(1).$$

2. Notice that

$$\tilde{y}_t = \tilde{u}_t + \beta\tilde{x}_{t-1} + \tilde{\alpha}_t$$

with $1-\pi = T^{-\gamma}$, where, after re-arranging sum terms,

$$\tilde{\alpha}_t = \sum_{j=0}^{t-2}\pi^j\Delta\alpha_{t-j}.$$

Note now that $|\mu(s_2)-\mu(s_1)| = \left|\int_0^{s_2}\mu'(x)\,\mathrm{d}x - \int_0^{s_1}\mu'(x)\,\mathrm{d}x\right| \leqslant \int_{s_1}^{s_2}|\mu'(x)|\,\mathrm{d}x \leqslant C|s_2-s_1|$ hence $|\Delta\alpha_t| = |\mu(t/T)-\mu((t-1)/T)| \leqslant \frac{C}{T}$. Therefore,

$$\left|\sum_{j=0}^{t-2}\pi^j\Delta\alpha_{t-j}\right| = O\left(T^{\gamma-1}\right),$$

as required.

Proof of Lemma A.2

1. We need to just bound the variance for each sum.

   (a) For the first part note that $z_{t-1} = \psi\check{z}_{t-1} + \delta_t$ where δ_t is uniformly L_4-bounded (see Demetrescu and Hosseinkouchack, 2021, proof of Lemma 2) and $\check{z}_{t-1} = \sum_{j=1}^{t-1}c_{j,t-1}v_j$. Therefore we have

$$\sum_{t=2}^{T}\tilde{\alpha}_t z_{t-1} = \psi\sum_{t=2}^{T}\tilde{\alpha}_t\check{z}_{t-1} + \sum_{t=2}^{T}\tilde{\alpha}_t\delta_t.$$

   For the first term we have

$$\mathrm{Var}\left(\sum_{t=2}^{T}\tilde{\alpha}_t\check{z}_{t-1}\right) \leqslant C\sum_{j=1}^{T}\left(\sum_{t=j+1}^{T}\tilde{\alpha}_t c_{j,t-1}\right)^2$$

$$
\begin{aligned}
&= C\sum_{j=1}^{T}\sum_{t=j+1}^{T}\sum_{s=j+1}^{T} c_{j,t-1}c_{j,s-1}\tilde{\alpha}_t\tilde{\alpha}_s \\
&\leqslant C\sum_{j=1}^{T}\sum_{t=j+1}^{T}\sum_{s=j+1}^{T} \left|c_{j,t-1}c_{j,s-1}\right|\left|\tilde{\alpha}_t\tilde{\alpha}_s\right| \\
&\leqslant CT^{-2}\sum_{j=1}^{T}\sum_{t=j+1}^{T}\sum_{s=j+1}^{T} \left|c_{j,t-1}c_{j,s-1}\right|\left|\frac{\pi^2-\pi^{t-2}}{\pi-\pi^2}\frac{\pi^2-\pi^{s-2}}{\pi-\pi^2}\right| \\
&\leqslant CT^{-2+2\gamma}\sum_{j=1}^{T}\sum_{t=j+1}^{T}\sum_{s=j+1}^{T} \left|c_{j,t-1}c_{j,s-1}\right| \\
&= O\left(T^{-2+2\gamma}T^{1+2\eta}\right).
\end{aligned}
$$

For the second term, using the L_4-boundedness of δ_t we obtain $T^{-1/2-\eta/2}\sum_{t=2}^{T}\tilde{\alpha}_t\delta_t = O_p\left(T^{\gamma\text{-}1/2\text{-}\eta/2}\right)$. Therefore $T^{-1/2-\eta/2}\sum_{t=2}^{T}\tilde{\alpha}_t z_{t-1} = O_p\left(T^{-1+\gamma+\eta/2}\right)$.

(b) For the second part again we may write $\sum_{t=2}^{T}\pi^{t-1}z_{t-1} = \psi\sum_{t=2}^{T}\pi^{t-1}\check{z}_{t-1} + \sum_{t=2}^{T}\pi^{t-1}\delta_t$. We have with $c_{j,t-1} = \frac{\varrho^{t-1-j}(1-\varrho)-\rho^{t-1-j}(1-\rho)}{\rho-\varrho}$

$$
\sum_{t=2}^{T}\pi^{t-1}\check{z}_{t-1} = \sum_{t=2}^{T}\sum_{j=1}^{t-1}\pi^{t-1}c_{j,t-1}v_j = \sum_{j=1}^{T-1}\sum_{t=j+1}^{T}\pi^{t-1}c_{j,t-1}v_j,
$$

hence

$$
\begin{aligned}
\mathrm{Var}\left(\sum_{t=2}^{T}\pi^{t-1}\check{z}_{t-1}\right) &= C\sum_{j=1}^{T-1}\left(\sum_{t=j+1}^{T}\pi^{t-1}c_{j,t-1}\right)^2 \\
&= C\left(\frac{1-\pi}{(1-\pi\varrho)(1-\pi\rho)}\right)^2\frac{\pi^2-\pi^{2T}}{1-\pi^2} \\
&+ C\left(\frac{(1-\varrho)\pi^T}{(1-\pi\varrho)(\varrho-\rho)}\right)^2\frac{\varrho^2-\varrho^{2T}}{1-\varrho^2} \\
&+ C\left(\frac{(1-\rho)\pi^T}{(1-\pi\rho)(\varrho-\rho)}\right)^2\frac{\rho^2-\rho^{2T}}{1-\rho^2} \\
&+ C\frac{2(1-\pi)(1-\varrho)\pi^T}{(1-\pi\varrho)^2(1-\pi\rho)(\varrho-\rho)}\frac{\pi^T\varrho-\pi\varrho^T}{\pi-\varrho} \\
&- C\frac{2(1-\pi)(1-\rho)\pi^T}{(1-\pi\varrho)(1-\pi\rho)^2(\varrho-\rho)}\frac{\pi^T\rho-\pi\rho^T}{\pi-\rho} \\
&- C\frac{2(1-\varrho)(1-\rho)\pi^{2T}}{(1-\pi\varrho)(1-\pi\rho)(\varrho-\rho)^2}\frac{\varrho\rho-\varrho^T\rho^T}{1-\varrho\rho}, \\
&= O\left(T^{\gamma+2\min\{\eta,\gamma\}}\right).
\end{aligned}
$$

Further, using the uniformly L_4-boundedness of δ_t we have $\sum_{t=2}^{T}\pi^{t-1}\delta_t = O_p\left(T^{\gamma/2}\right)$. Therefore $\sum_{t=2}^{T}\pi^{t-1}z_{t-1} = O_p\left(T^{\gamma/2+\min\{\eta,\gamma\}}\right)$.

2. Again note that $\sum_{t=2}^{T} r_{t-1}z_{t-1} = \psi \sum_{t=2}^{T} r_{t-1}\check{z}_{t-1} + \sum_{t=2}^{T} r_{t-1}\delta_t$. It is then straightforward to show $\mathrm{E}\left(\left|\sum_{t=2}^{T} r_{t-1}\check{z}_{t-1}\right|\right) = O\left(T^{1+\min\{\gamma,\eta\}}\right)$ as follows. Write

$$
\begin{aligned}
\mathrm{E}\left(\left(\sum_{t=2}^{T} r_{t-1}\check{z}_{t-1}\right)^2\right) &= \mathrm{E}\left(\sum_{t=2}^{T}\sum_{s=2}^{T} r_{t-1}\check{z}_{t-1}r_{s-1}\check{z}_{s-1}\right) \\
&= \mathrm{E}\left(\sum_{t,s=2}^{T}\sum_{k,j=1}^{t-1}\sum_{l,m=1}^{s-1} \pi^{t-k-1}\pi^{s-l-1}c_{j,t-1}c_{m,s-1}v_j u_k v_m u_l\right).
\end{aligned}
$$

For the latter, letting S to denote the summations over $s \leq t$, we have

$$
\begin{aligned}
S &= \sum_{t=2}^{T}\sum_{s=2}^{t}\sum_{k=1}^{s-1}\sum_{j=1,j\neq k}^{s-1} \pi^{t-k-1}\pi^{s-k-1}c_{j,t-1}c_{j,s-1}\,\mathrm{E}\left(v_j^2 u_k^2\right) \\
&+ \sum_{t=2}^{T}\sum_{s=2}^{t}\sum_{k=1}^{s-1}\sum_{j=1,j\neq k}^{s-1} \pi^{t-k-1}\pi^{s-k-1}c_{j,t-1}c_{j,s-1}\,\mathrm{E}\left(v_j^2 u_k^2\right) \\
&+ \sum_{t=2}^{T}\sum_{s=2}^{t}\sum_{j=1}^{t-1}\sum_{m=1}^{s-1} \pi^{t-j-1}\pi^{s-m-1}c_{j,t-1}c_{m,s-1}\,\mathrm{E}\left(v_j u_j v_m u_m\right) \\
&+ \sum_{t=2}^{T}\sum_{s=2}^{t}\sum_{k=1}^{s-1}\sum_{j=1}^{s-1} \pi^{t-k-1}\pi^{s-j-1}c_{j,t-1}c_{k,s-1}\,\mathrm{E}\left(v_j u_k v_k u_j\right), \\
&= S_1 + S_2 + S_3 + S_4.
\end{aligned}
$$

For S_1 we have

$$
\begin{aligned}
S_1 &= C\sum_{t=2}^{T}\sum_{s=2}^{t}\sum_{k=1}^{s-1} \pi^{t-k-1}\pi^{s-k-1}c_{k,t-1}c_{k,s-1}, \\
&= \sum_{t=2}^{T}\sum_{s=2}^{t}\sum_{k=1}^{s-1} \pi^{s+t-2k-2}\left(\frac{(1-\varrho)^2}{(\rho-\varrho)^3}\varrho^{s+t-2-2k} - \frac{(1-\varrho)(1-\rho)}{(\rho-\varrho)^2}\varrho^{t-1-k}\rho^{s-1-k}\right. \\
&\left. - \frac{(1-\varrho)(1-\rho)}{(\rho-\varrho)^2}\rho^{t-1-k}\varrho^{s-1-k} + \frac{(1-\rho)^2}{(\rho-\varrho)^2}\rho^{s+t-2-2k}\right), \\
&= S_{1,1} + S_{1,2} + S_{1,3} + S_{1,4}.
\end{aligned}
$$

Using geometric summations we obtain $S_{1,1} = O\left(T^{1+2\min\{\eta,\gamma\}}\right)$ which dominates $S_{1,2}$, $S_{1,3}$ and $S_{1,4}$, therefore $S_1 = O\left(T^{1+2\min\{\eta,\gamma\}}\right)$. We now turn our attention to S_2 which is a bit more involved. We have

$$
\begin{aligned}
S_2 &= C\sum_{t=2}^{T}\sum_{s=2}^{t}\sum_{k=1}^{s-1} \pi^{s+t-2k-2}\sum_{j=1}^{s-1} c_{j,t-1}c_{j,s-1}, \\
&= \frac{C}{1-\pi^2}\sum_{t=2}^{T}\sum_{s=2}^{t}(\pi^{t-s} - \pi^{t+s-2})\sum_{j=1}^{s-1} c_{j,t-1}c_{j,s-1}.
\end{aligned}
$$

For $c_{j,t-1}c_{j,s-1}$ we have

$$
c_{j,t-1}c_{j,s-1} = \frac{(1-\varrho)^2}{(\rho-\varrho)^2}\varrho^{s+t-2-2j} - \frac{(1-\varrho)(1-\rho)}{(\rho-\varrho)^2}\rho^{s-1-j}\varrho^{t-1-j}
$$

$$-\frac{(1-\varrho)(1-\rho)}{(\rho-\varrho)^2}\rho^{t-1-j}\varrho^{s-1-j}+\frac{(1-\rho)^2}{(\rho-\varrho)^2}\rho^{s+t-2-2j},$$

hence

$$\begin{aligned}\sum\nolimits_{j=1}^{s-1}c_{j,t-1}c_{j,s-1} &= \frac{(1-\varrho)}{(\rho-\varrho)^2(1+\varrho)}\left(\varrho^{t-s}-\varrho^{s+t-1}\right)-\frac{(1-\varrho)(1-\rho)}{(\rho-\varrho)^2\rho(1-\varrho\rho)}\left(\rho\varrho^{t-s}-\rho^s\varrho^{t-1}\right)\\ &\quad -\frac{(1-\varrho)(1-\rho)}{(\rho-\varrho)^2\varrho(1-\rho\varrho)}\left(\rho\varrho^{t-s}-\rho^{t-1}\varrho^{s}\right)+\frac{(1-\rho)}{(\rho-\varrho)^2(1+\rho)}\left(\rho^{t-s}-\rho^{t+s-2}\right).\end{aligned}$$

Plugging the latter in the expression for S_2 we obtain

$$\begin{aligned}S_2 &= \frac{C}{1-\pi^2}\frac{(1-\varrho)}{(\rho-\varrho)^2(1+\varrho)}\sum_{t=2}^{T}\sum\nolimits_{s=2}^{t}\left(\pi^{t-s}-\pi^{t+s-2}\right)\left(\varrho^{t-s}-\varrho^{s+t-1}\right)\\ &+ \frac{C}{1-\pi^2}\frac{(1-\varrho)(1-\rho)}{(\rho-\varrho)^2\rho(1-\varrho\rho)}\sum\nolimits_{t=2}^{T}\sum\nolimits_{s=2}^{t}\left(\pi^{t-s}-\pi^{t+s-2}\right)\left(\rho\varrho^{t-s-1}-\rho^s\varrho^{t-1}\right)\\ &+ \frac{C}{1-\pi^2}\frac{(1-\varrho)(1-\rho)}{(\rho-\varrho)^2\varrho(1-\rho\varrho)}\sum\nolimits_{t=2}^{T}\sum\nolimits_{s=2}^{t}\left(\pi^{t-s}-\pi^{t+s-2}\right)\left(\rho\varrho^{t-s-1}-\rho^{t-1}\varrho^{s}\right)\\ &+ \frac{C}{1-\pi^2}\frac{(1-\rho)}{(\rho-\varrho)^2(1+\rho)}\sum\nolimits_{t=2}^{T}\sum\nolimits_{s=2}^{t}\left(\pi^{t-s}-\pi^{t+s-2}\right)\left(\rho^{t-s}-\rho^{t+s-2}\right),\\ &= S_{2,1}+S_{2,2}+S_{2,3}+S_{2,4}.\end{aligned}$$

$S_{2,1}$, $S_{2,2}$, $S_{2,3}$ and $S_{2,4}$ can be bounded using elementary arguments to obtain $S_2 = O\left(T^{1+\gamma+\eta+\min\{\gamma,\eta\}}\right)$. Further we obtain for S_3 and S_4

$$\begin{aligned}S_3 &= O\left(T^{2+2\min\{\gamma,\eta\}}\right),\\ S_4 &= O\left(T^{1+3\min(\gamma,\eta)}\right)+O\left(T^{\eta+2\gamma+\min(\gamma,\eta)}\right)+O\left(T^{-1+2\eta+3\gamma}\right).\end{aligned}$$

Therefore $S = O\left(T^{2+2\min\{\gamma,\eta\}}\right)$, which in turn, with $\eta<\gamma$, implies that $\sum_{t=2}^{T} r_{t-1}\check{z}_{t-1} = O_p\left(T^{1+\eta}\right)$. Now, using Lemma A.1 we have that $T^{-\gamma/2}r_{t-1}$ is uniformly L_2-bounded, hence $\sum_{t=2}^{T} r_{t-1}\delta_t$ is dominated by $\sum_{t=2}^{T} r_{t-1}\check{z}_{t-1}$ and hence the result follows.

3. We have $\tilde{x}_{t-1} = x_{t-1}-\frac{1}{T^\gamma}\sum_{j=0}^{t-2}\pi^j x_{t-2-j}-\pi^{t-1}x_0$ where for $\sum_{j=0}^{t-2}\pi^j x_{t-2-j}$ we may write

$$\begin{aligned}T^{-1/2-\gamma}\sum_{j=0}^{t-2}\pi^j x_{t-2-j} &= T^{-1/2-\gamma}\sum_{j=0}^{t-2}\left(\sum_{k=0}^{j}\rho^{j-k}\pi^k\right)v_{t-2-j}\\ &= T^{-1/2-\gamma}\sum_{j=0}^{t-2}\frac{\pi^{j+1}-\rho^{j+1}}{\pi-\rho}v_{t-2-j}\\ &= -p^{-1}\left(1+o_p(1)\right)T^{-1/2}\sum_{j=0}^{t-2}\rho^{j+1}v_{t-2-j}+O\left(T^{-1/2+\gamma/2}\right)\end{aligned}$$

where $T^{-1/2}\sum_{j=0}^{[sT]}\rho^{j+1}v_{t-2-j}\Rightarrow J_{c,\mathrm{H}}(s)$, which in turn implies $T^{-1/2}\tilde{x}_{[sT]}\Rightarrow\psi\left(1+\frac{1}{p}\right)J_{c,\mathrm{H}}(s)$. Therefore using the arguments in the proof of Lemma A.4 in Demetrescu and Rodrigues (2022) we obtain

$$\frac{1}{T^{1+\eta}}\sum_{t=2}^{T}z_{t-1}\tilde{x}_{t-1}\quad\Rightarrow\quad\frac{1+p}{ap}\psi^{2}\left(J_{c,\mathrm{H}}^{2}(1)-\int J_{c,\mathrm{H}}(s)\,\mathrm{d}J_{c,\mathrm{H}}(s)\right).$$

Proof of Proposition 1

Under the null $\beta=0$, we have for $\frac{1}{T^{1+\eta}}\sum_{t=2}^{T}(z_{t-1}-\bar{z})^{2}(y_{t}-\bar{y})^{2}$ that

$$\begin{aligned}\frac{1}{T^{1+\eta}}\sum_{t=2}^{T}(z_{t-1}-\bar{z})^{2}(y_{t}-\bar{y})^{2} &= \frac{1}{T^{1+\eta}}\sum_{t=2}^{T}(z_{t-1}-\bar{z})^{2}(u_{t}-\bar{u}+\alpha_{t}-\bar{\alpha})^{2}\\ &= \frac{1}{T^{1+\eta}}\sum_{t=2}^{T}(z_{t-1}-\bar{z})^{2}(u_{t}-\bar{u})^{2}\\ &+ \frac{1}{T^{1+\eta}}\sum_{t=2}^{T}(z_{t-1}-\bar{z})^{2}(\alpha_{t}-\bar{\alpha})^{2}\\ &+ \frac{2}{T^{1+\eta}}\sum_{t=2}^{T}(z_{t-1}-\bar{z})^{2}(u_{t}-\bar{u})(\alpha_{t}-\bar{\alpha})\\ &\overset{d}{\to} \frac{\psi^{2}}{2a}\int\sigma_{v}^{2}(s)\,\sigma_{u}^{2}(s)\,\mathrm{d}s+\frac{\psi^{2}}{2a}\int\sigma_{v}^{2}(s)\,(\mu(s)-\bar{\mu})^{2}\,\mathrm{d}s\end{aligned}$$

where the first term comes from Demetrescu and Rodrigues (2022, Lemma A.4), the 2nd follows analogously to the first, and the third is easily seen to have vanishing variance since it can be reduced to a sum of martingale differences.

Proof of Proposition 2

Examining the numerator of the t statistic, we have

$$\frac{1}{T^{1/2+\eta/2}}\sum_{t=2}^{T}z_{t-1}\tilde{y}_{t}=\frac{1}{T^{1/2+\eta/2}}\sum_{t=2}^{T}z_{t-1}\tilde{u}_{t}+\frac{\beta}{T^{1/2+\eta/2}}\sum_{t=2}^{T}z_{t-1}\tilde{x}_{t-1}+\frac{1}{T^{1/2+\eta/2}}\sum_{t=2}^{T}z_{t-1}\tilde{\alpha}_{t},$$

where the third term on the r.h.s. is negligible (Lemma A.2). Then, with $r_{t-1}=\sum_{j=0}^{t-2}\pi^{j}u_{t-1-j}$,

$$\frac{1}{T^{1/2+\eta/2}}\sum_{t=2}^{T}z_{t-1}\tilde{u}_{t}=\frac{1}{T^{1/2+\eta/2}}\sum_{t=2}^{T}z_{t-1}u_{t}-\frac{1}{T^{1/2+\gamma+\eta/2}}\sum_{t=2}^{T}r_{t-1}z_{t-1}-\frac{u_{1}}{T^{1/2+\eta/2}}\sum_{t=2}^{T}\pi^{t-1}z_{t-1}$$

where Lemma A.2 indicates that the second summand is negligible since $\gamma>1/2+\eta/2$ implies $\gamma>\eta$; also, by the same Lemma, the third summand vanishes too since $\frac{1}{2}+\frac{\eta}{2}>$

$\frac{\gamma}{2} + \eta$ by assumption.

Then, we learn from Lemma A.2 that, under the considered local alternative,

$$\frac{\beta}{T^{1/2+\eta/2}} \sum z_{t-1}\tilde{x}_{t-1} \Rightarrow b \cdot \frac{1+p}{ap}\psi^2 \left(J_{c,\mathrm{H}}^2 (1) - \int J_{c,\mathrm{H}} (s) \, \mathrm{d}J_{c,\mathrm{H}} (s) \right).$$

Finally, it is easily shown that

$$\begin{aligned} \frac{1}{T^{1+\eta}} \sum_{t=2}^{T} z_{t-1}^2 \tilde{y}_t^2 &= \frac{1}{T^{1+\eta}} \sum_{t=2}^{T} z_{t-1}^2 u_t^2 + o_p(1) \\ &\xrightarrow{p} \frac{\psi^2}{2a} \int \sigma_v^2 (s) \, \sigma_u^2 (s) \, \mathrm{d}s \end{aligned}$$

(where the probability limit follows from Lemma A.4 in Demetrescu and Rodrigues, 2022). Since $\frac{\sum_{t=2}^{T} z_{t-1} u_t}{\sqrt{\sum_{t=2}^{T} z_{t-1}^2 u_t^2}} \to \mathcal{Z}$, see Demetrescu and Rodrigues (2022, Lemma A.3), the desired result follows.

Proof of Proposition 3

It is not difficult to show that, conditional on the data,

$$\tilde{t}_{vx}^* = \frac{\sum_{t=2}^{T} z_{t-1}^* u_t R_t}{\sqrt{\sum_{t=2}^{T} z_{t-1}^{*2} u_t^2 R_t^2}} + o_p^* (1) \, .$$

The desired result then follows with the results provided by Demetrescu et al. (2022a).

References

Amihud, Y. and C. M. Hurvich (2004). Predictive regressions: A reduced-bias estimation method. *Journal of Financial and Quantitative Analysis 39*(4), 813–841.

Amihud, Y., C. M. Hurvich, and Y. Wang (2009). Multiple-predictor regressions: Hypothesis testing. *Review of Financial Studies 22*(1), 413–434.

Andersen, T. G. and R. T. Varneskov (2021). Consistent local spectrum (lcm) inference for predictive return regressions. Technical Report w28569, National Bureau of Economic Research.

Bai, J. and P. Perron (1998). Estimating and testing linear models with multiple structural changes. *Econometrica 66*(1), 47–78.

Bauer, D. and A. Maynard (2012). Persistence-robust surplus-lag granger causality testing. *Journal of Econometrics 169*(2), 293–300.

Breitung, J. and M. Demetrescu (2015). Instrumental variable and variable addition based inference in predictive regressions. *Journal of Econometrics 187*(1), 358–375.

Campbell, J. Y. (2008). Viewpoint: Estimating the equity premium. *Canadian Journal of Economics/Revue canadienne d'économique 41*(1), 1–21.

Campbell, J. Y. and M. Yogo (2006). Efficient tests of stock return predictability. *Journal of Financial Economics 81*(1), 27–60.

Cavaliere, G., A. Rahbek, and A. M. R. Taylor (2010). Testing for Co-Integration in Vector Autoregressions with Non-Stationary Volatility. *Journal of Econometrics 158*(1), 7–24.

Demetrescu, M., I. Georgiev, P. Rodrigues, and A. M. R. Taylor (2022a). Extensions to IVX methods of inference for return predictability. *Journal of Econometrics forthcoming.*

Demetrescu, M., I. Georgiev, P. Rodrigues, and A. M. R. Taylor (2022b). Testing for episodic predictability in stock returns. *Journal of Econometrics 227*(1), 85–113.

Demetrescu, M. and B. Hillmann (2022). Nonlinear predictability of stock returns? Parametric vs. nonparametric inference in predictive regressions. *Journal of Business & Economic Statistics 40*(1), 382–397.

Demetrescu, M. and M. Hosseinkouchack (2021). Finite-sample size control of IVX-based tests in predictive regressions. *Econometric Theory 37*(4), 769–793.

Demetrescu, M. and P. M. M. Rodrigues (2022). Residual-augmented IVX predictive regression. *Journal of Econometrics 227*(2), 429–460.

Elliott, G., U. K. Müller, and M. W. Watson (2015). Nearly optimal tests when a nuisance parameter is present under the null hypothesis. *Econometrica 83*(2), 771–811.

Ferson, W. E., S. Sarkissian, and T. T. Simin (2003). Spurious regressions in financial economics? *The Journal of Finance 58*(4), 1393–1413.

Georgiev, I., D. I. Harvey, S. J. Leybourne, and A. R. Taylor (2019). A bootstrap stationarity test for predictive regression invalidity. *Journal of Business & Economic Statistics 37*(3), 528–541.

Gonzalo, J. and J.-Y. Pitarakis (2017). Inferring the predictability induced by a persistent regressor in a predictive threshold model. *Journal of Business & Economic Statistics 35*(2), 202–217.

Jansson, M. and M. J. Moreira (2006). Optimal inference in regression models with nearly integrated regressors. *Econometrica 74*(3), 681–714.

Kostakis, A., T. Magdalinos, and M. P. Stamatogiannis (2015). Robust econometric inference for stock return predictability. *Review of Financial Studies 28*(5), 1506–1553.

Lee, J. H. (2016). Predictive quantile regression with persistent covariates: IVX-QR approach. *Journal of Econometrics 192*(1), 105–118.

Lettau, M. and S. V. Nieuwerburgh (2008). Reconciling the return predictability evidence. *The Review of Financial Studies 21*(4), 1607–1652.

Pástor, L. and R. F. Stambaugh (2009). Predictive systems: Living with imperfect predictors. *The Journal of Finance 64*(4), 1583–1628.

Pavlidis, E. G., I. Paya, and D. A. Peel (2017). Testing for speculative bubbles using spot and forward prices. *International Economic Review 58*(4), 1191–1226.

Phillips, P. C. B. (2015). Pitfalls and possibilities in predictive regression. *Journal of Financial Econometrics 13*(3), 521–555.

Phillips, P. C. B. and J. H. Lee (2013). Predictive regression under various degrees of persistence and robust long-horizon regression. *Journal of Econometrics 177*(2), 250–264.

Yang, B., W. Long, L. Peng, and Z. Cai (2020). Testing the predictability of us housing price index returns based on an IVX-AR model. *Journal of the American Statistical Association 115*(532), 1598–1619.

Table 2: Size properties of different tests under short-run dynamics and strong contemporaneous shock correlation

		$T=250$												$T=500$											
		2-sided				left-sided				right-sided				2-sided				left-sided				right-sided			
α_μ	c	t_{vx}^{W}	t_{vx}^{WBP}	$\tilde{t}_{vx}^{*}$	$\tilde{t}_{vx}^{**}$	t_{vx}^{W}	t_{vx}^{WBP}	$\tilde{t}_{vx}^{*}$	$\tilde{t}_{vx}^{**}$	t_{vx}^{W}	t_{vx}^{WBP}	$\tilde{t}_{vx}^{*}$	$\tilde{t}_{vx}^{**}$	t_{vx}^{W}	t_{vx}^{WBP}	$\tilde{t}_{vx}^{*}$	$\tilde{t}_{vx}^{**}$	t_{vx}^{W}	t_{vx}^{WBP}	$\tilde{t}_{vx}^{*}$	$\tilde{t}_{vx}^{**}$	t_{vx}^{W}	t_{vx}^{WBP}	$\tilde{t}_{vx}^{*}$	$\tilde{t}_{vx}^{**}$
0	0	4.66	4.88	6.12	6.28	0.10	0.00	3.00	3.24	9.56	9.94	8.36	7.96	4.08	4.28	6.28	5.50	0.06	0.02	3.02	3.04	8.52	9.02	7.90	7.56
	1	4.46	4.58	6.76	6.56	0.12	0.00	3.38	3.90	9.18	9.90	8.30	7.54	4.46	4.66	6.66	6.22	0.16	0.00	3.76	3.86	9.18	9.66	8.26	7.74
	5	5.70	5.56	5.58	5.06	1.04	0.06	4.38	4.60	9.64	10.38	6.18	5.60	5.38	5.20	5.64	4.94	0.94	0.12	4.36	4.30	9.94	10.40	6.62	6.10
	10	6.24	5.96	5.30	4.88	1.94	0.48	4.30	4.44	9.52	10.18	6.20	5.46	5.42	4.92	5.44	4.78	1.74	0.52	4.38	4.34	9.34	9.96	6.02	5.70
	30	5.96	5.82	6.06	5.18	2.88	1.72	5.12	5.10	8.52	9.46	5.70	5.16	5.20	5.14	5.76	5.20	2.06	1.30	5.18	5.00	7.46	8.14	5.94	5.56
	50	5.92	5.84	5.66	5.20	3.38	2.44	5.18	4.96	7.82	8.62	5.84	5.48	5.32	5.24	5.70	5.12	2.96	2.04	5.24	5.10	7.50	8.26	5.70	5.12
0.25	0	9.90	10.36	6.34	6.42	0.78	0.00	2.82	3.40	16.14	18.24	8.54	7.98	14.00	12.46	5.70	5.62	1.94	0.00	3.12	3.42	20.58	21.10	7.64	7.20
	1	10.40	11.14	7.04	6.88	1.54	0.00	3.52	3.72	16.54	19.52	9.06	8.32	15.16	15.34	6.08	5.66	2.72	0.00	3.66	3.44	21.52	26.54	7.84	7.12
	5	11.24	12.02	6.54	6.02	3.36	0.10	3.88	4.16	15.44	21.26	7.22	6.74	15.28	16.02	5.80	5.34	6.06	0.00	4.30	4.32	18.42	28.32	6.30	6.12
	10	10.06	10.76	5.80	5.54	4.04	0.20	4.94	5.06	13.34	19.02	5.98	5.66	13.90	15.44	5.88	5.30	6.40	0.16	4.88	4.68	15.74	26.56	5.74	5.14
	30	7.52	9.22	5.64	5.18	3.40	0.58	5.08	5.20	9.88	15.72	5.62	5.24	8.60	11.14	5.26	4.92	4.90	0.52	5.34	5.12	11.04	19.70	4.92	4.50
	50	6.48	8.02	6.08	5.84	3.46	1.08	5.20	5.20	8.38	12.86	5.58	5.30	7.72	9.64	6.40	5.62	4.54	0.70	5.36	5.16	8.82	16.02	6.16	6.02
0.5	0	26.12	13.44	6.82	6.58	5.10	0.00	2.68	3.36	30.88	21.56	9.24	8.38	39.20	10.18	5.92	5.92	11.52	0.00	3.02	3.30	36.62	17.78	7.46	7.20
	1	27.88	17.24	6.96	6.70	7.40	0.00	3.24	3.60	30.40	27.88	8.76	7.74	40.68	12.92	6.14	5.72	13.78	0.00	3.94	3.88	35.62	21.82	7.60	7.26
	5	27.54	21.92	5.14	4.26	10.22	0.06	3.40	3.66	27.42	35.08	6.26	5.78	41.54	19.52	5.66	5.10	15.70	0.00	4.40	4.34	34.62	31.96	6.28	5.80
	10	20.40	21.90	6.40	5.78	8.80	0.02	4.66	4.94	20.54	33.72	6.32	5.82	31.34	20.26	5.48	4.98	13.96	0.00	4.52	4.32	26.82	33.70	5.90	5.44
	30	12.62	15.84	5.58	4.88	6.48	0.14	4.64	4.66	13.66	25.44	5.72	5.62	18.12	16.76	5.18	4.70	9.80	0.06	4.62	4.38	16.76	26.34	5.38	5.06
	50	9.70	13.14	5.40	4.58	6.02	0.44	5.08	4.82	10.64	21.22	5.96	5.30	12.98	13.68	5.78	5.30	7.78	0.30	5.04	4.86	12.28	22.36	5.62	5.22
τ																									
0.25	0	13.16	12.50	6.80	6.60	1.28	0.00	2.70	3.28	19.22	21.30	9.08	8.26	19.66	13.64	5.76	5.44	3.10	0.00	3.00	3.08	25.24	22.28	7.70	7.18
	1	12.70	12.98	6.94	6.62	1.98	0.02	3.64	3.78	20.04	23.76	8.00	7.94	21.14	16.76	6.58	5.84	4.42	0.00	3.50	3.44	25.68	26.50	7.88	7.38
	5	14.44	14.60	5.80	5.80	5.14	0.06	4.00	4.36	18.18	24.74	6.80	6.26	21.30	18.70	5.30	5.24	8.98	0.04	4.50	4.52	22.86	30.56	5.64	5.28
	10	12.42	13.46	5.46	5.10	5.24	0.28	4.80	4.80	15.06	23.10	6.14	5.74	16.78	17.98	5.92	5.40	7.48	0.06	4.88	5.10	18.50	29.20	6.00	5.52
	30	8.22	9.90	6.42	5.84	4.56	0.60	5.52	5.66	10.18	16.42	5.56	5.10	10.10	13.16	5.92	5.48	5.82	0.28	5.46	5.32	11.98	21.74	5.40	5.06
	50	7.28	9.48	5.76	4.84	3.98	0.92	4.84	4.56	8.88	15.68	5.56	4.88	8.74	11.18	5.54	5.20	4.92	0.68	5.54	5.38	9.78	18.20	5.38	5.00
0.5	0	33.70	10.58	7.28	7.22	8.48	0.00	3.00	3.70	34.70	18.38	9.68	8.82	45.66	6.70	6.72	6.36	14.58	0.00	3.74	3.88	38.84	13.58	8.54	7.74
	1	35.82	11.60	7.34	6.78	9.80	0.00	3.72	3.90	34.84	20.04	9.44	8.56	49.98	7.70	7.72	7.28	18.00	0.00	4.14	4.12	39.44	15.42	8.16	7.88
	5	36.94	16.48	6.98	6.52	13.56	0.04	4.94	5.20	32.90	27.30	7.74	7.10	50.32	9.52	6.20	5.50	20.36	0.00	4.14	4.04	37.52	18.18	6.88	6.70
	10	28.44	15.68	6.02	5.72	12.66	0.00	4.98	4.98	25.02	24.86	6.28	5.66	42.54	10.72	5.92	5.38	18.78	0.02	4.28	4.20	32.02	19.18	6.10	5.64
	30	16.82	13.24	5.88	5.20	8.70	0.34	5.06	4.98	16.52	21.78	5.64	5.24	25.92	10.26	6.00	5.28	13.58	0.44	5.18	4.84	21.44	17.94	5.80	5.32
	50	13.40	11.98	6.18	5.74	7.60	0.72	5.32	5.34	12.44	19.28	5.96	5.46	18.26	9.34	5.70	5.00	9.86	0.54	4.82	4.64	15.96	15.52	5.64	5.14
1	0	57.72	6.96	9.12	8.46	22.40	0.00	4.14	4.66	41.72	13.84	11.38	10.46	68.76	5.80	7.98	7.76	31.18	0.00	5.00	5.02	42.62	11.82	9.08	8.72
	1	60.92	8.06	9.14	8.64	24.38	0.00	4.60	5.00	42.12	15.58	10.96	10.24	70.56	6.06	7.98	7.16	30.06	0.00	4.56	4.74	44.84	13.18	8.94	8.38
	5	61.94	9.20	7.36	6.66	26.74	0.04	5.42	5.48	40.60	17.26	7.96	7.40	71.70	7.94	7.64	6.94	31.68	0.02	5.66	5.38	44.30	15.34	7.96	7.74
	10	53.68	8.78	7.08	6.82	24.22	0.12	5.68	5.98	36.00	16.14	6.84	6.44	64.34	8.36	7.12	6.56	29.92	0.12	5.44	5.12	40.08	15.94	6.92	6.32
	30	34.92	8.68	5.78	5.34	16.58	0.66	4.82	4.84	26.12	14.86	5.76	5.34	46.22	7.66	5.80	5.12	23.24	0.58	5.72	5.50	30.10	12.62	5.60	4.90
	50	26.48	8.08	5.40	4.70	13.92	0.98	5.02	4.82	21.14	13.76	5.60	5.14	38.22	6.92	5.60	5.30	21.02	1.10	5.18	5.02	25.68	11.80	5.16	4.90

Note: Data generated with (1) and (2) with $v_t = \phi v_t + \nu_t$ for $\phi = 0.5$, where $(u_t, \nu_t) \sim iiN(0, \mathbf{\Sigma})$ and $\mathbf{\Sigma}$ exhibits constant correlation $\delta = -0.95$. We set $\rho = 1 - {}^{c}/_{T}$ for various c. t_{vx}^{W} is defined as in equation (3) for which we use the small sample corrections proposed by Kostakis et al. (2015) and $\tilde{t}_{vx}^{*}$ is the bootstrap implementation of the test with mean-adjusted y_t. For t_{vx}^{W} and t_{vx}^{WBP} we set $\varrho = 1 - {}^{1}/_{T^{0.95}}$. For $\tilde{t}_{vx}^{*}$ we set $\varrho = 1 - {}^{1}/_{T^{0.20}}$ and $\pi = 1 - {}^{1}/_{T^{0.75}}$.

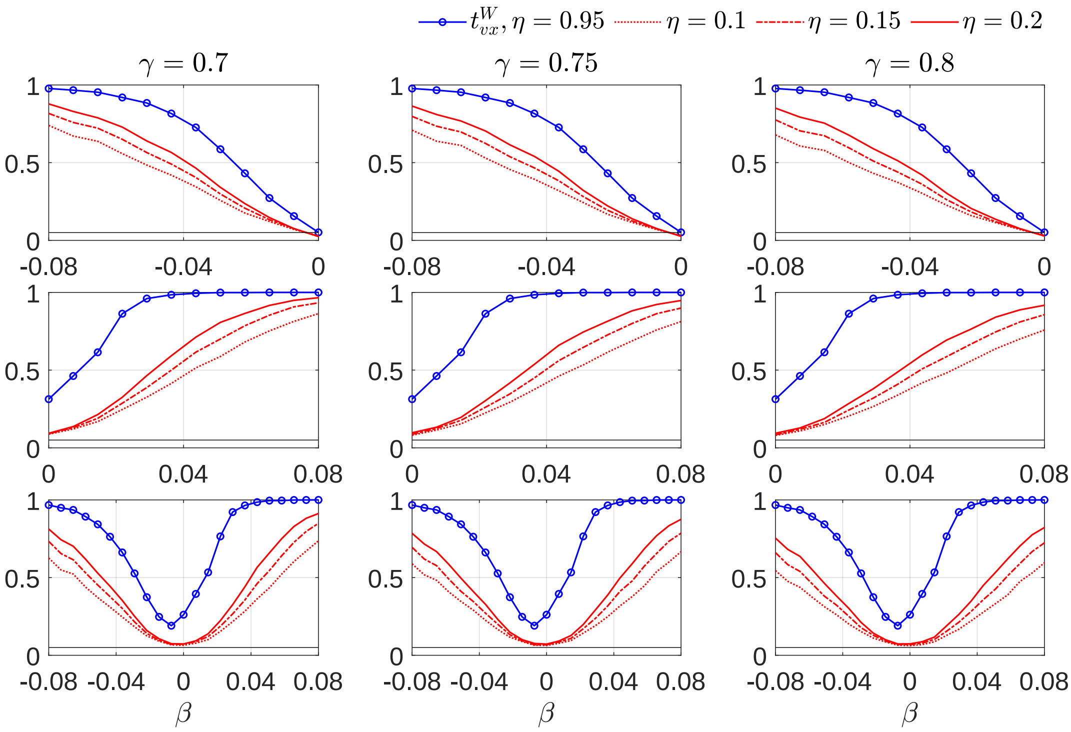

Figure 5: Power curves for $\tilde{t}_{vx}^*$ and t_{vx}^W under a local alternative when $c = 0$ and $\alpha_\mu = 0.5$. The power curves for t_{vx}^W are the same for all the three sub figures, while power curves for $\tilde{t}_{vx}$ are calculated with different values for η and γ. See the text for more details.

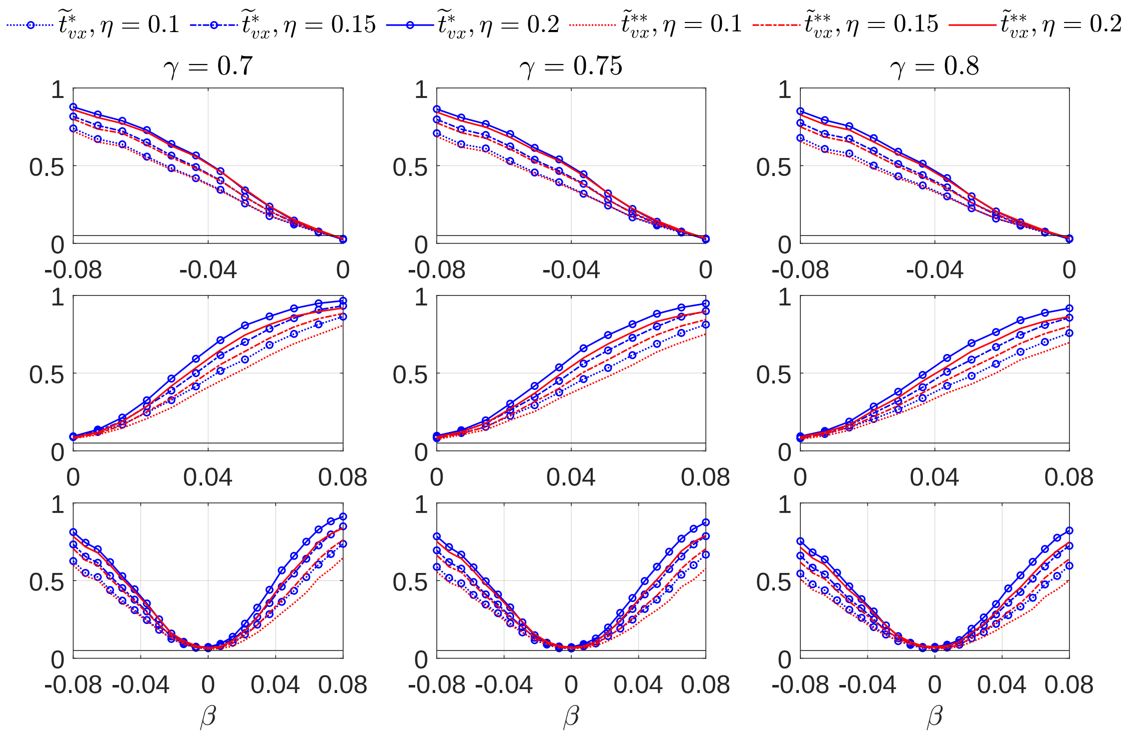

Figure 6: Power curves for $\tilde{t}^{*}_{vx}$ and $\tilde{t}^{**}_{vx}$ under a local alternative when $c = 0$ and $\alpha_{\mu} = 0.5$. See the text for more details.