Assessing predictive accuracy in panel data models with long-range dependence

Daniel Borup, Bent Jesper Christensen
and Yunus Emre Ergemen

CREATES Research Paper 2019-4

Department of Economics and Business Economics
Aarhus University
Fuglesangs Allé 4
DK-8210 Aarhus V
Denmark

Email: oekonomi@au.dk
Tel: +45 8716 5515

Assessing predictive accuracy in panel data models with long-range dependence*

Daniel Borup† Bent Jesper Christensen‡ Yunus Emre Ergemen§

Abstract

This paper proposes tests of the null hypothesis that model-based forecasts are uninformative in panels, allowing for individual and interactive fixed effects that control for cross-sectional dependence, endogenous predictors, and both short-range and long-range dependence. We consider a Diebold-Mariano style test based on comparison of the model-based forecast and a nested no-predictability benchmark, an encompassing style test of the same null, and a test of pooled uninformativeness in the entire panel. A simulation study shows that the encompassing style test is reasonably sized in finite samples, whereas the Diebold-Mariano style test is oversized. Both tests have non-trivial local power. The methods are applied to the predictive relation between economic policy uncertainty and future stock market volatility in a multi-country analysis.

Keywords: Panel data, predictability, long-range dependence, Diebold-Mariano test, encompassing test

JEL Classification: C12, C23, C33, C52, C53

This version: March 28, 2019

*We are grateful to Carlos Vladimir Rodriguez Caballero, Federico Carlini, Bo Honoré, Morten Ørregaard Nielsen, Carlos Velasco, and participants at the Long Memory Conference 2018 at Aalborg University and the joint Econometrics-Finance seminar at the Center for Research in Econometric Analysis of Time Series (CREATES) at Aarhus University for useful comments, and to CREATES (DNRF78), the Dale T. Mortensen Center, Aarhus University, and the Danish Social Science Research Council for research support.

†CREATES, Department of Economics and Business Economics, Aarhus University. Email: dborup@econ.au.dk.

‡Corresponding author. CREATES and the Dale T. Mortensen Center, Department of Economics and Business Economics, Aarhus University. Email: bjchristensen@econ.au.dk.

§CREATES, Department of Economics and Business Economics, Aarhus University. Email: ergemen@econ.au.dk.

I. Introduction

Many macroeconomic and financial variables are presented in the form of panels, describing dynamic characteristics of the individual units such as countries or assets. Examples include cross-country panels of GDP and inflation, international panels of stock returns or their volatility, and intraday electricity prices, see, e.g., Ergemen et al. (2016). In the interest of forecasting such variables, it should be natural to treat data as a panel rather than separate time series. Relative to a pure time series approach, a panel approach has the potential to yield efficiency gains and improved forecasts by accounting for the interaction between cross-sectional units, see, e.g., Canova and Ciccarelli (2004), Groen (2005), and Baltagi (2013).

Many macroeconomic and financial time series have been shown to exhibit possibly fractional long-range dependence, see, e.g., Gil-Alaña and Robinson (1997) and Andersen et al. (2003). Model-based forecasting accounting for such features has been considered, e.g., by Christensen and Nielsen (2005), Corsi (2009), Busch et al. (2011), and Bollerslev et al. (2013). As argued by, among others, Robinson and Velasco (2015), Ergemen and Velasco (2017), and Ergemen (2019), panel data models should also account for these features, both in order to obtain valid inference, see Kruse et al. (2018), and for possibly more accurate forecasts, see, e.g., Bos et al. (2002), Bhardwaj and Swanson (2006), Deo et al. (2006), and Chiriac and Voev (2011).

In this paper, we study out-of-sample predictive accuracy in a general fractionally integrated panel data model and develop formal tests of (un)informativeness of the model-based forecasts. We consider the data-generating process proposed by Ergemen (2019), which allows for individual and interactive fixed effects, endogenous predictors, and both short-range and long-range dependence. The model nests stationary $I(0)$ and nonstationary $I(1)$ panel data models and features a multifactor structure that accounts for cross-sectional dependence in data. It allows for potentially different integration orders of factor components and the possibility of cointegrating relations, which may improve forecasting via an error-correction mechanism, see, e.g., Engel et al. (2008) for an application to exchange rate modelling in a panel context. Our approach allows for heterogeneity in both slope parameters and persistence characteristics, providing flexibility and wider applicability than a model restricting all panel units to share common dynamics. The model also nests popular forecasting frameworks, such as panel vector-autoregressive systems (Westerlund et al., 2016), predictive regressions (Welch and Goyal, 2008), and autoregressive forecasting (Stock and Watson, 1999), possibly with the addition of exogenous or endogenous predictors (Clark and McCracken, 2006).

The estimation approach is based on proxying the multifactor structure by a cross-sectional averaging procedure, following Ergemen and Velasco (2017) and Ergemen (2019). Our main interest is then in testing the null hypothesis of uninformativeness, at a given forecast horizon, of forecasts obtained from the general model, after removing the multifactor structure. Rejection in this case should imply rejection with the factors included, as factor augmentation generally improves forecasting performance, see, e.g., McCracken and Ng (2016). First, for each unit in the panel, we develop a Diebold and Mariano (1995) style test statistic comparing the loss associated with the model-based forecasts relative to that from a no-predictability benchmark, i.e., the unconditional mean. We estimate the unconditional mean on the evaluation sample, thus also facilitating an evaluation of the informativeness of externally obtained forecasts, such as those from surveys or market-implied values. Next, we develop an encompassing version of our predictability test, which is as easy to implement as the Diebold-Mariano style tests. Finally, besides correcting individual forecasts for panel features such as interaction across units, it may also be of interest to evaluate predictive accuracy for the entire panel. To this end, we propose pooled versions of our test statistics, in a spirit similar to Pesaran (2007).

We contribute to the literature on forecast evaluation in several ways. Primarily, we provide the first tests of predictive accuracy in a panel with long-range dependence. This extends the results in Kruse et al. (2018) by treating a panel, rather than working with individual time series. Moreover, our approach considers the underlying process of the forecasts, rather than being silent about the origin of these. This distinction is important, since it allows us to compare model-based forecasts with the nested unconditional mean. Accordingly, the resulting Diebold-Mariano (DM) style statistic has a non-standard, half-normal distribution under the null hypothesis. As in Breitung and Knüppel (2018), we circumvent this issue by providing a simple modification leading to a chi-squared distributed test statistic. The encompassing test on the other hand possesses a standard normal distribution. Our framework extends Hjalmarson (2010), Westerlund and Narayan (2015a,b), and Westerlund et al. (2016) who present in-sample analyses of predictive regressions in short-range dependent panel settings which may allow for either endogenous predictors or factor structure, but not both simultaneously. These studies focus on stock return predictability, whereas our framework would be relevant for the analysis of general panels involving other variables such as stock return volatility or aggregate macroeconomic variables, because it features out-of-sample analysis, long-range dependence, and the co-existence of endogenous predictors and a factor structure. Moreover, to the best of our knowledge, there are only a few papers examining out-of-sample forecast evaluation within the typical short-range dependent panel literature. Pesaran et al. (2009) and Chudik et al. (2016) propose pooled DM tests, possibly accounting for cross-sectional dependence, whereas Liu et al. (2015) employ a panel-wide Giacomini and White (2006) conditional test of predictive ability.

We explore the finite-sample properties of our testing procedures by means of Monte Carlo experiments. The encompassing test is reasonably sized, whereas the DM style test suffers from oversizing, paralleling the findings in Breitung and Knüppel (2018), and the pooled test is quite conservative. Both tests have non-trivial power against local departures from the null. In an empirical application, we apply our methodology to estimate the relationship between a newspaper-based index of economic policy uncertainty (Baker et al., 2016) and stock market volatility in 14 countries, obtain forecasts, and evaluate the predictive accuracy of the economic policy index for future stock market volatility.

The rest of the paper is laid out as follows. Section II presents the model framework and the conditions imposed to study it based on Ergemen (2019). Section III introduces the forecast setting based on the model framework and discusses the null hypothesis of interest. It also provides the main results for DM and encompassing style tests and related panel-wide generalization. Section IV examines the finite-sample properties of the proposed tests based on Monte Carlo experiments and Section V presents the empirical application. Section VI concludes.

Throughout the paper, "$(n,T)_j$" and "$(n,T,m)_j$" denote the joint asymptotics in which the sample is growing in multiple dimensions, with n the cross-section dimension, T the length of the in-sample and m the out-of-sample window, "$\Rightarrow$" denotes weak convergence, "$\xrightarrow{p}$" convergence in probability, "$\xrightarrow{d}$" convergence in distribution, and

$\|A\| = (\text{trace}(AA'))^{1/2}$ for a matrix A. All proofs are collected in the Appendix.

II. Model framework

We describe the model framework and estimation procedure of Ergemen (2019) as the basis for our forecasting discussions in the next sections. The basis of the approach is a triangular array describing a long-range dependent panel data model of the observed series (y_{it}, x_{it}) given by

$$\begin{aligned} y_{it} &= \alpha_i + \beta_{i0}' x_{it} + \lambda_i' f_t + \Delta_t^{-d_{i0}} \epsilon_{1it}, \\ x_{it} &= \mu_i + \gamma_i' f_t + \Delta_t^{-\vartheta_{i0}} \epsilon_{2it}, \end{aligned} \tag{1}$$

where, for $i = 1, \ldots, n$ and $t = 1, \ldots, T$, the scalar y_{it} and the k-vector of covariates x_{it} are observable, α_i and μ_i are unobserved individual fixed effects, f_t is the q-vector of unobserved common factors whose j-th component is fractionally integrated of order δ_j (so we write $f_{jt} \sim I(\delta_j)$), $j = 1, \ldots, q$, and the q-vector λ_i and the $q \times k$ matrix γ_i contain the corresponding unobserved factor loadings indicating how much each cross-section unit is impacted by f_t. Both k and q are fixed throughout. In (1), with prime denoting transposition, $\epsilon_{it} = (\epsilon_{1it}, \epsilon_{2it}')'$ is a covariance stationary process, allowing for $\text{Cov}[\epsilon_{1it}, \epsilon_{2it}] \neq 0$, with short-range vector-autoregressive (VAR) dynamics described by

$$B(L; \theta_i)\epsilon_{it} \equiv \left(I_{k+1} - \sum_{j=1}^{p} B_j(\theta_i) L^j \right) \epsilon_{it} = v_{it}, \tag{2}$$

where L is the lag operator, θ_i the short-range dependence parameters, I_{k+1} the $(k+1) \times (k+1)$ identity matrix, B_j are $(k+1) \times (k+1)$ upper-triangular matrices, and v_{it} is a $(k+1) \times 1$ sequence that is identically and independently distributed across i and t with zero mean and variance-covariance matrix $\Omega_i > 0$. Throughout the paper, the operator Δ_t^{-d} applied to a vector or scalar ϵ_{it} is defined by

$$\Delta_t^{-d} \epsilon_{it} = \Delta^{-d} \epsilon_{it} \mathbb{1}(t > 0) = \sum_{j=0}^{t-1} \pi_j(-d) \epsilon_{it-j}, \quad \pi_j(-d) = \frac{\Gamma(j+d)}{\Gamma(j+1)\Gamma(d)},$$

where $\mathbb{1}(\cdot)$ is the indicator function and $\Gamma(\cdot)$ the gamma function, such that $\Gamma(d) = \infty$ for $d = 0, -1, -2, \ldots$, and $\Gamma(0)/\Gamma(0) = 1$.

For the analysis of the system in (1), Ergemen (2019) considers $d_{i0} \in \mathscr{D}_i = [\underline{d}_i, \overline{d}_i]$ and $\vartheta_{i0} \in \mathscr{V}_i = [\underline{\vartheta}_i, \overline{\vartheta}_i]^k$ with $\underline{d}_i, \underline{\vartheta}_i > 0$, implying that the observable series are fractionally integrated. In particular, $y_{it} \sim I(\max\{\vartheta_{i0}, d_{i0}, \delta_{max}\})$ and $x_{it} \sim I(\max\{\vartheta_{i0}, \delta_{max}\})$ where $\delta_{max} = \max_j \delta_j$. Further, setting

$$\vartheta_{max} = \max_i \vartheta_{i0} \quad \text{and} \quad d_{max} = \max_i d_{i0}$$

and letting d^* denote a prewhitening parameter chosen by the econometrician, the following conditions are imposed on (1).

Assumption A (Long-range dependence and common-factor structure). *Persistence and cross-section dependence are introduced according to the following:*

1. *The fractional integration parameters, with true values $\vartheta_{i0} \neq d_{i0}$, satisfy $d_{i0} \in \mathscr{D}_i = [\underline{d}_i, \overline{d}_i] \subset (0, 3/2)$, $\vartheta_{i0} \in \mathscr{V}_i = [\underline{\vartheta}_i, \overline{\vartheta}_i]^k \subset (0, 3/2)^k$, $\vartheta_{max} - \underline{\vartheta}_i < 1/2$, $\vartheta_{max} - \underline{d}_i < 1/2$, $\delta_{max} - \underline{\vartheta}_i < 1/2$, $\delta_{max} - \underline{d}_i < 1/2$, $d_{max} - \underline{d}_i < 1/2$, and $d^* > \max\{\vartheta_{max}, d_{max}, \delta_{max}\} - 1/4$.*

2. *For $j = 1, \ldots, q$, the j-th component of the common factor vector satisfies $f_{jt} = \alpha_j^f + \Delta_t^{-\delta_j} z_{jt}^f$, $\delta_j \geq 0$, for $\delta_{max} < 3/2$, where the vector z_t^f containing the $I(0)$ series z_{jt}^f satisfies $z_t^f = \Psi^f(L)\varepsilon_t^f$, with $\Psi^f(s) = \sum_{l=0}^{\infty} \Psi_l^f s^l$, $\sum_{l=0}^{\infty} l \|\Psi_l^f\| < \infty$, $\det\left(\Psi^f(s)\right) \neq 0$ for $|s| \leq 1$, and $\varepsilon_t^f \sim iid(0, \Sigma_f)$, $\Sigma_f > 0$, $\mathbb{E}\|\varepsilon_t^f\|^4 < \infty$.*

3. *f_t and ϵ_{it} are independent, and independent of the factor loadings λ_i and γ_i, for all i and t.*

4. *The factor loadings λ_i and γ_i are independent across i, and $rank(\overline{\boldsymbol{C}}_n) = q \leq k+1$ for all n, where the $(k+1) \times q$ matrix $\overline{\boldsymbol{C}}_n$ containing cross-sectionally averaged factor loadings is defined as*

$$\overline{\boldsymbol{C}}_n = \begin{pmatrix} \overline{\beta_0' \gamma'}_n + \overline{\lambda'}_n \\ \overline{\gamma'}_n \end{pmatrix},$$

with $\overline{\gamma}_n = n^{-1} \sum_{i=1}^n \gamma_i$, $\overline{\lambda}_n = n^{-1} \sum_{i=1}^n \lambda_i$, and $\overline{\beta_0' \gamma'}_n = n^{-1} \sum_{i=1}^n \beta_{i0}' \gamma_i'$.

Assumption A.1 imposes restrictions on the range of memory orders allowed, motivated by the use of first differences (to remove fixed effects) in the methodology. The requirement on the lower bounds of the sets $\mathscr{D}_i$ and $\mathscr{V}_i$ is necessary to ensure that the initial-condition terms, arising due to the use of truncated filters and uniformly of size $O_p(T^{-\underline{d}_i})$ and $O_p(T^{-\underline{\vartheta}_i})$, vanish asymptotically. The conditions that restrict the distance between the parameter values allowed and the lower bounds of the sets are necessary to control for the unobserved individual fixed effects, see Robinson and Velasco (2015), and cross-section dependence, see Ergemen and Velasco (2017), as well as to ensure that the projection approximations adopted below work well with the original integration orders of the series, see Ergemen (2019) for rigorous details. The projection method based on the cross-section averages of the d^*-differenced observables is guaranteed to work under Assumption A.1 since the projection errors vanish asymptotically with the prescribed choice of d^*. For most applications, first differences, $d^* = 1$, would suffice, anticipating $\vartheta_{i0}, \delta_{max}, d_{i0} < 5/4$.

Assumption A.2 allows for a fractionally integrated common factor vector that may also exhibit short-memory dynamics, where the $I(0)$ innovations of f_t are not collinear and each common factor can have different memory, unlike the homogeneity restriction imposed by Ergemen and Velasco (2017). The upper bound condition on the maximal factor memory is not restrictive and is motivated by working with $d^* \geq 1$. The non-zero mean possibility in common factors, i.e., when $\alpha_j^f \neq 0$, allows for a drift.

Assumption A.3 is standard in the factor model literature and has been used, e.g., by Pesaran (2006) and Bai (2009). When $\lambda_i \neq 0$ and $\gamma_i \neq 0$, further endogeneity is induced by the common factors, in addition to that stemming from $\mathrm{Cov}[\epsilon_{1it}, \epsilon_{2it}] \neq 0$ in (1).

Assumption A.4 states that sufficiently many covariates whose sample averages can span the factor space are required. When the system in (1) is written for $z_{it} = (y_{it}, x_{it}')'$, the matrix $\overline{\mathbf{C}}_n$ basically contains the cross-sectionally averaged factor loadings. The full rank condition on $\overline{\mathbf{C}}_n$ simplifies the identification of q factors with $k+1$ cross-section averages of observables. This condition is also imposed by Pesaran (2006) in establishing the asymptotics of heterogeneous slope parameters.

Assumption B (System errors). *In the representation*

$$B(L;\theta_i)\epsilon_{it} \equiv \left(I_{k+1} - \sum_{j=1}^{p} B_j(\theta_i) L^j \right) \epsilon_{it} = v_{it},$$

1. $B_j(\cdot)$ *are upper-triangular matrices satisfying* $\sum_{j=1}^{\infty} j \|B_j\| < \infty$, $\det(B(s;\theta_i)) \neq 0$ *for*

$|s| = 1$, $\theta_i \in \Theta_i$.

2. *the v_{it} are identically and independently distributed vectors across i and t with zero mean and positive-definite covariance matrix Ω_i, and have bounded fourth-order moments.*

Assumption B.1 rules out possible collinearity in the innovations by imposing a standard summability requirement and ensures well-defined functional behaviour at zero frequency. This ensures invertibility and thereby allows for a VARMA representation, see Robinson and Hualde (2003). Finally, Assumption B.2 imposes a standard moment requirement.

For the estimation of both linear (slope) and memory parameters in (1), Ergemen (2019) works with the projected series that proxy the common-factor structure up to an asymptotically vanishing projection error. First-differencing (1) to remove the fixed effects,

$$\begin{aligned} \Delta y_{it} &= \beta_{i0}' \Delta x_{it} + \lambda_i' \Delta f_t + \Delta_t^{1-d_{i0}} \epsilon_{1it}, \\ \Delta x_{it} &= \gamma_i' \Delta f_t + \Delta_t^{1-\vartheta_{i0}} \epsilon_{2it}, \end{aligned} \tag{3}$$

for $i = 1, \dots, n$ and $t = 2, \dots, T$, (3) can be prewhitened from idiosyncratic long-range dependence for some fixed exogenous differencing choice, d^*, as prescribed in Assumption A.1, by which all variables become asymptotically stationary with their sample means converging to population limits.

Using the notation $a_{it}(\tau) = \Delta_{t-1}^{\tau-1} \Delta a_{it}$ for any τ, the prewhitened model is given by

$$\begin{aligned} y_{it}(d^*) &= \beta_{i0}' x_{it}(d^*) + \lambda_i' f_t(d^*) + \epsilon_{1it}(d^* - d_{i0}), \\ x_{it}(d^*) &= \gamma_i' f_t(d^*) + \epsilon_{2it}(d^* - \vartheta_{i0}). \end{aligned} \tag{4}$$

Thus, using the notation $z_{it}(\tau_1, \tau_2) = \left(y_{it}(\tau_1), x_{it}'(\tau_2)\right)'$, (4) can be written in the vectorized form as

$$z_{it}(d^*, d^*) = \zeta \beta_{i0}' x_{it}(d^*) + \Lambda_i' f_t(d^*) + \epsilon_{it}\left(d^* - d_{i0}, d^* - \vartheta_{i0}\right), \tag{5}$$

where $\zeta = (1, 0, \dots, 0)'$, and $\Lambda_i = (\lambda_i \ \gamma_i)$. The structure $\Lambda_i' f_t(d^*)$ in (5) induces cross-sectional correlation across individual units i through the common factor structure $f_t(d^*)$. A standard method for dealing with this unobserved structure is projection based on proxies obtained by sample cross-section averages of (differenced) data, see

Pesaran (2006). Write $\bar{z}_t(d^*,d^*) = n^{-1}\sum_{i=1}^{n} z_{it}(d^*,d^*)$ for the cross-section average of (5), so that

$$\bar{z}_t(d^*,d^*) = \overline{\zeta\beta_0' x_t(d^*)} + \bar{\Lambda}' f_t(d^*) + \bar{\epsilon}_t\left(d^* - d_0, d^* - \vartheta_0\right), \tag{6}$$

where $\bar{\epsilon}_t(d^* - d_0, d^* - \vartheta_0)$ is $O_p(n^{-1/2})$ for sufficiently large d^*. Thus, $\bar{z}_t(d^*,d^*)$ and $\overline{\zeta\beta_0' x_t(d^*)}$ asymptotically capture all the information provided by the common factors, provided that $\bar{\Lambda}$ is of full rank. Note that $\bar{x}_t(d^*)$ is readily contained in $\bar{z}_t(d^*,d^*) = (\bar{y}_t(d^*), \bar{x}_t(d^*)')'$ and that the β_{i0} do not introduce any dynamics in $\overline{\zeta\beta_0' x_t(d^*)}$ since they are fixed for each i. Therefore, $\bar{z}_t(d^*,d^*)$ alone can span the factor space.

Write the prewhitened time-stacked observed series as $\mathbf{x}_i(d^*) = (x_{i2}(d^*),\ldots,x_{iT}(d^*))'$ and $\mathbf{z}_i(d^*,d^*) = (z_{i2}(d^*,d^*),\ldots,z_{iT}(d^*,d^*))'$ for $i = 1,\ldots,n$. Then, for each $i = 1,\ldots,n$,

$$\mathbf{z}_i(d^*,d^*) = \mathbf{x}_i(d^*)\beta_{i0}\zeta' + \mathbf{F}(d^*)\Lambda_i + \mathbf{E}_i\left(d^* - d_{i0}, d^* - \vartheta_{i0}\right), \tag{7}$$

where $\mathbf{E}_i(d^* - d_{i0}, d^* - \vartheta_{i0}) = (\epsilon_{i2}(d^* - d_{i0}, d^* - \vartheta_{i0},),\ldots,\epsilon_{iT}(d^* - d_{i0}, d^* - \vartheta_{i0}))'$ and $\mathbf{F}(d^*) = (f_2(d^*),\ldots,f_T(d^*))'$. Writing $T_1 = T - 1$, the common factor structure can asymptotically be removed by the $T_1 \times T_1$ feasible projection matrix

$$\overline{\mathbf{M}}_{T_1}(d^*) = \mathbf{I}_{T_1} - \bar{\mathbf{z}}(d^*,d^*)(\bar{\mathbf{z}}'(d^*,d^*)\bar{\mathbf{z}}(d^*,d^*))^{-}\bar{\mathbf{z}}'(d^*,d^*), \tag{8}$$

where $\bar{\mathbf{z}}(d^*,d^*) = n^{-1}\sum_{i=1}^{n}\mathbf{z}_i(d^*,d^*)$, and P^- denotes the generalized inverse of a matrix P. When the projection matrix is built with the original (possibly nonstationary) series, it is impossible to ensure the asymptotic replacement of the latent factor structure by cross-section averages because the noise in (5) may be too persistent when $d^* = 0$. On the other hand, using some $d^* > \max\{\vartheta_{max}, d_{max}, \delta_{max}\} - 1/4$ for prewhitening guarantees that the projection errors vanish asymptotically.

Introducing the infeasible projection matrix based on unobserved factors,

$$\mathbf{M}_F(d^*) = \mathbf{I}_{T_1} - \mathbf{F}(d^*)(\mathbf{F}(d^*)'\mathbf{F}(d^*))^{-}\mathbf{F}(d^*)',$$

and adopting Pesaran (2006)'s argument under Assumption A.2 and A.4, we have the approximation as $(n,T)_j \to \infty$ that

$$\overline{\mathbf{M}}_{T_1}(d^*)\mathbf{F}(d^*) \approx \mathbf{M}_F(d^*)\mathbf{F}(d^*) = 0. \tag{9}$$

This implies that the feasible and infeasible projection matrices can be employed interchangeably for factor removal in the asymptotics, provided that the rank condition in Assumption A.4 holds. Based on (7) and using the approximation in (9), the defactored observed series for each $i = 1, \ldots, n$ is

$$\tilde{\mathbf{z}}_i(d^*, d^*) \approx \tilde{\mathbf{x}}_i(d^*)\beta_{i0}\zeta' + \tilde{\mathbf{E}}_i\left(d^* - d_{i0}, d^* - \vartheta_{i0}\right), \tag{10}$$

where $\tilde{\mathbf{z}}_i(d^*, d^*) = \overline{\mathbf{M}}_{T_1}(d^*)\mathbf{z}_i(d^*, d^*)$, $\tilde{\mathbf{x}}_i(d^*) = \overline{\mathbf{M}}_{T_1}(d^*)\mathbf{x}_i(d^*)$ and $\tilde{\mathbf{E}}_i(d^* - d_{i0}, d^* - \vartheta_{i0}) = \overline{\mathbf{M}}_{T_1}(d^*)\mathbf{E}_i(d^* - d_{i0}, d^* - \vartheta_{i0})$.

Integrating the defactored series back by d^* to their original integration orders, disregarding the projection errors that are negligible as $n \to \infty$ under Assumption A.1 as shown by Ergemen and Velasco (2017), and defining the generic parameter vector $\tau_i = (d_i, \vartheta_i')'$,

$$\tilde{z}^*_{it}(\tau_i) = \zeta\beta_{i0}'\tilde{x}^*_{it}(d_i) + \tilde{\varepsilon}^*_{it}(d_i - d_{i0}, \vartheta_i - \vartheta_{i0}), \tag{11}$$

where the first and second equations of (11) are for the transformed series

$$\tilde{y}^*_{it}(d_i) = \Delta_{t-1}^{d_i - d^*}\tilde{y}_{it}(d^*) \quad \text{and} \quad \tilde{x}^*_{it}(\vartheta_i) = \Delta_{t-1}^{\vartheta_i - d^*}\tilde{x}_{it}(d^*),$$

respectively, omitting the dependence on d^* and the (asymptotically negligible) initial conditions in the notation. Both the linear and long-range dependence parameters are identified, and may be estimated with the standard parametric convergence rates with centered asymptotic normal distributions. Ergemen and Velasco (2017) consider estimation under the assumptions of common memory orders δ_j of factors and exogeneity of the defactored series $\tilde{x}^*_{it}(\vartheta_i)$ for $\tilde{y}^*_{it}(d_i)$, and Ergemen (2019) relaxes these assumptions.

To see the identification, apply the operator $B(L;\theta_i)$ giving the short-range dynamics in (2) to each side of (11) and invoke Assumption B to get

$$\tilde{z}^*_{it}(\tau_i) - \sum_{j=1}^{p} B_j(\theta_i)\tilde{z}^*_{it-j}(\tau_i)$$
$$= \zeta\beta'_{i0}\tilde{x}^*_{it}(d_i) - \sum_{j=1}^{p} B_j(\theta_i)\zeta\beta'_{i0}\tilde{x}^*_{it-j}(d_i) + \tilde{v}^*_{it}(d_i - d_{i0}, \vartheta_i - \vartheta_{i0}). \quad (12)$$

Here, the error terms are serially uncorrelated, given that v_{it} are identically and independently distributed, and clearly $B_j(\theta_i)$ is identified. Writing the system out in more detail, noting that $\tilde{z}^*_{it}(d_i, \vartheta_i) = (\tilde{y}^*_{it}(d_i), \tilde{x}^*_{it}(\vartheta_i)')'$, the first equation in (12) is

$$\tilde{y}^*_{it}(d_i) = \beta'_{i0}\tilde{x}^*_{it}(d_i) + \sum_{j=1}^{p} B_{1j}(\theta_i)\tilde{z}^*_{it-j}(\tau_i)$$
$$- \sum_{j=1}^{p} B_{1j}(\theta_i)\zeta\beta'_{i0}\tilde{x}^*_{it-j}(d_i) + \tilde{v}^*_{1it}(d_i - d_{i0}), \quad (13)$$

and the second equation is

$$\tilde{x}^*_{it}(\vartheta_i) - \sum_{j=1}^{p} B_{2j}(\theta_i)\tilde{z}^*_{it-j}(d_i, \vartheta_i) = -\sum_{j=1}^{p} B_{2j}(\theta_i)\zeta\beta'_{i0}\tilde{x}^*_{it-j}(d_i) + \tilde{v}^*_{2it}(\vartheta_i - \vartheta_{i0}), \quad (14)$$

with $B_j = (B'_{1j}, B'_{2j})'$. Combining (13) and (14), we obtain

$$\tilde{y}^*_{it}(d_i) = \beta'_{i0}\tilde{x}^*_{it}(d_i) + \rho'_i\tilde{x}^*_{it}(\vartheta_i) + \sum_{j=1}^{p} \left(B_{1j}(\theta_i) - \rho'_i B_{2j}(\theta_i)\right)\left(\tilde{z}^*_{it-j}(\tau_i) - \zeta\beta'_{i0}\tilde{x}^*_{it-j}(d_i)\right)$$
$$+ \tilde{v}^*_{1it}(d_i - d_{i0}) - \rho'_i\tilde{v}^*_{2it}(\vartheta_i - \vartheta_{i0}), \quad (15)$$

where $\rho_i = \mathbb{E}[\tilde{v}^*_{2it}\tilde{v}^{*\prime}_{2it}]^{-1}\mathbb{E}[\tilde{v}^*_{2it}\tilde{v}^*_{1it}]$, suppressing the dependence on fractional orders. The error term in (15) is orthogonal to $\tilde{v}^*_{2it}$, given that v_{it} are identically and independently distributed, so that $\tilde{v}^*_{1it} - \rho'_i\tilde{v}^*_{2it}$ is uncorrelated with $\tilde{v}^*_{2it}$ as a result of the triangular array structure of the system. Ergemen (2019) imposes $\vartheta_{i0} \neq d_{i0}$ to avoid collinearity when $\vartheta_{i0} = d_{i0}$. On this basis, both β_{i0} and ρ_i are identified.

Following Ergemen (2019), we can now consistently estimate the linear parameters, β_{i0}, ρ_i, and $B_j(\theta_i)$, by least squares for each i. The feasible estimate requires estimation of the long-range dependence parameters, which feature non-linearly in the system given by (14) and (15). They can be estimated equation-by-equation by conditional sum of squares (CSS), where d_{i0} is estimated from (15), and ϑ_{i0} can be estimated from any of the k equations for x_{it} in (14), or by an objective function constructed over all k equations. Both linear and non-linear parameters enjoy standard $\sqrt{T}$ convergence rates and centered normal asymptotic distributions.

III. Forecasting

In this section, we study predictive accuracy of forecasts using the projected series. We develop Diebold-Mariano and encompassing style test statistics and determine their asymptotic distribution. Forecasting the original (non-defactored) series instead would require a different methodology, since the fractionally integrated unobserved factors would need to be estimated. Asymptotic results for such factor estimates are yet to be provided in the literature. Instead, we focus on testing uninformativeness in the defactored series. Rejection in this case should imply rejection in the original series.

Employing (15), the h-step-ahead direct forecast, $h = 0, 1, \ldots, H$ with H fixed, can be written as

$$\tilde{y}^*_{it+h}(\tau_i) = \beta'_{i0}\tilde{x}^*_{it}(d_i) + \rho'_i\tilde{x}^*_{it}(\vartheta_i) + \sum_{j=1}^{p}\left(B_{1j}(\theta_i) - \rho'_i B_{2j}(\theta_i)\right)\left(\tilde{z}^*_{it-j}(\tau_i) - \zeta\beta'_{i0}\tilde{x}^*_{it-j}(d_i)\right)$$
$$+ \tilde{v}^*_{1it+h}(d_i - d_{i0}) - \rho'_i\tilde{v}^*_{2it+h}(\vartheta_i - \vartheta_{i0}), \qquad (16)$$

making the dependence on d_i and ϑ_i explicit, so that conditional on information through time t, we can write down the forecasting equation as

$$\tilde{y}^*_{it+h|t}(\tau_i) = \beta'_{i0}\tilde{x}^*_{it}(d_i) + \rho'_i\tilde{x}^*_{it}(\vartheta_i) + \sum_{j=1}^{p}\left(B_{1j}(\theta_i) - \rho'_i B_{2j}(\theta_i)\right)\left(\tilde{z}^*_{it-j}(\tau_i) - \zeta\beta'_{i0}\tilde{x}^*_{it-j}(d_i)\right). \qquad (17)$$

Thus, forecasts are derived from a heterogeneous predictive model, facilitating unit-specific inferences on predictive accuracy, while still treating each unit as part of a panel.

It is important to understand the generality of this model, which nests several models employed in the literature, while still accounting for long-range and cross-sectional dependence. If $B_j(\theta_i) = 0$ and $\rho_i = 0$, for all i, we recover the well-known predictive regression for each i, with coefficients β_{i0}. This is especially popular in the financial literature on stock return and volatility prediction, typically relating aggregate financial ratios or macroeconomic conditions to future stock returns and their volatility. In the pure time series setting, Welch and Goyal (2008) is a classical example of an examination of the in- and out-of-sample predictive performance of stock characteristics, interest-rate related and macroeconomic indicators for future stock market returns, whereas Hjalmarson (2010) and Westerlund and Narayan (2015a,b) employ panel predictive regressions. If the zero condition on ρ_i is relaxed, the framework allows for endogenous predictors, even after accounting for cross-sectional dependence, which is a common issue in finance (Stambaugh, 1999). If, instead, $\beta_{i0} = \rho_i = B_{2j}(\theta_i) = 0$ for all i, we obtain a conventional short-range dependent autoregressive model in y_{it+h} with coefficients $B_{1j}(\theta_i)$, as often used in macroeconomic settings, for instance for inflation forecasting (Stock and Watson, 1999). Again, once the zero conditions on β_{i0} or ρ_i are relaxed, other exogenous or endogenous predictors are accommodated, as, e.g., in Phillips curve forecasting models, with x_{it} a measure of economic activity, such as the output gap or unemployment rate. Clark and McCracken (2006) explore each of these in a univariate, short-range dependent setting. The predictor may enter with additional lags in (17) by allowing non-zero $B_{2j}(\theta_i)$. Once $k \geq 1$, we have an upper-triangular panel VAR system, which is typical in forecasting settings where y_{it} can depend on lagged values of itself and the predictors, whereas each component in x_{it} depends only its own past values, see, e.g., Westerlund et al. (2016).

For an evaluation sample (out-of-sample) that includes m observations, indexed by $t = 1, \ldots, m$, we may distinguish between three different estimation schemes for the forecasting model, noting that its parameters can be estimated consistently, cf. Section II. First, the recursive scheme fixes the starting point of the estimation sample (in-sample) window at $t = -T+1$ and increases its endpoint recursively with t. Second, the rolling scheme fixes the length of the estimation sample to T observations and, thus, increases both the starting and end point with t. Third, the fixed window scheme estimates parameters only once on the (initial) estimation sample of size T. Given the parameter estimates at time t, we then form the h-step forecasts from the forecasting equation in (17).

The theoretical forecast error is

$$
\begin{aligned}
e_{it+h|t}(\tau_i) &:= \tilde{y}^*_{it+h}(\tau_i) - \tilde{y}^*_{it+h|t}(\tau_i) \\
&= \tilde{v}^*_{1it+h}(d_i - d_{i0}) - \rho_i' \tilde{v}^*_{2it+h}(\vartheta_i - \vartheta_{i0}).
\end{aligned} \tag{18}
$$

Since under Assumption A.1 the difference between the maximum integration orders and the lower bounds of allowed set values is less than 1/2, the forecast error $e_{it+h|t}(\tau_i)$ is stationary and exhibits long memory whenever $d_i \neq d_{i0}$ or $\vartheta_i \neq \vartheta_{i0}$, which must be accounted for when conducting inference on predictive accuracy.

We are interested in testing the null hypothesis that the forecast function $\tilde{y}^*_{it+h|t}(\tau_i)$ is uninformative for $\tilde{y}^*_{it+h}(\tau_i)$ for fixed i,

$$
H_0 : \mathbb{E}\left[e^2_{it+h|t}(\tau_i)\right] \geq \mathbb{E}\left[\left(\tilde{y}^*_{it+h}(\tau_i) - \overline{\tilde{y}^*_{it}(\tau_i)}\right)^2\right], \tag{19}
$$

where $\overline{\tilde{y}^*_{it}(\tau_i)}$ is the expectation of $\tilde{y}^*_{it}(\tau_i)$ for fixed i. In order to gain more insight on the forecast error, we can decompose the mean squared error (MSE) as

$$
\begin{aligned}
\mathbb{E}\left[e^2_{it+h|t}(\tau_i)\right] = \mathbb{E}&\left[\left(\tilde{y}^*_{it+h}(\tau_i) - \overline{\tilde{y}^*_{it}(\tau_i)}\right)^2\right] \\
&- 2\mathbb{E}\left[\left(\tilde{y}^*_{it+h}(\tau_i) - \overline{\tilde{y}^*_{it}(\tau_i)}\right)\left(\tilde{y}^*_{it+h|t}(\tau_i) - \overline{\tilde{y}^*_{it}(\tau_i)}\right)\right] \\
&+ \mathbb{E}\left[\left(\tilde{y}^*_{it+h|t}(\tau_i) - \overline{\tilde{y}^*_{it}(\tau_i)}\right)^2\right].
\end{aligned} \tag{20}
$$

Clearly, $\mathbb{E}\left[\left(\tilde{y}^*_{it+h}(\tau_i) - \overline{\tilde{y}^*_{it}(\tau_i)}\right)\left(\tilde{y}^*_{it+h|t}(\tau_i) - \overline{\tilde{y}^*_{it}(\tau_i)}\right)\right] = 0$ is a sufficient condition for the forecast to be uninformative, as $\mathbb{E}\left[\left(\tilde{y}^*_{it+h|t}(\tau_i) - \overline{\tilde{y}^*_{it}(\tau_i)}\right)^2\right] \geq 0$. Furthermore, for a rational forecast satisfying $\mathbb{E}\left[e_{it+h|t}(\tau_i)\tilde{y}^*_{it+h|t}(\tau_i)\right] = 0$ and since under Assumption B.2 $\mathbb{E}[e_{i,t+h|t}(\tau_i)] = 0$, we have that

$$
\begin{aligned}
&\mathbb{E}\left[\left(\tilde{y}^*_{it+h}(\tau_i) - \overline{\tilde{y}^*_{it}(\tau_i)}\right)\left(\tilde{y}^*_{it+h|t}(\tau_i) - \overline{\tilde{y}^*_{it}(\tau_i)}\right)\right] \\
&= \mathbb{E}\left[\left(e_{it+h|t}(\tau_i) + \tilde{y}^*_{it+h|t}(\tau_i) - \overline{\tilde{y}^*_{it}(\tau_i)}\right)\left(\tilde{y}^*_{it+h|t}(\tau_i) - \overline{\tilde{y}^*_{it}(\tau_i)}\right)\right] \\
&= \mathbb{E}\left[\left(\tilde{y}^*_{it+h|t}(\tau_i) - \overline{\tilde{y}^*_{it}(\tau_i)}\right)^2\right].
\end{aligned} \tag{21}
$$

Therefore, by combining (20) and (21), any rational forecast with positive variance is informative and the null hypothesis (19) is then equivalent to

$$\text{Cov}\left[\tilde{y}^*_{it+h}(\tau_i), \tilde{y}^*_{it+h|t}(\tau_i)\right] = 0,$$

which will be utilized further below in constructing an alternative to a Diebold-Mariano style test statistic.

Furthermore, an iterative scheme can be used if, instead of (16), we write

$$\begin{aligned}\tilde{y}^*_{it+h}(\tau_i) = \beta'_{i0}\tilde{x}^*_{it+h}(d_i) + \rho'_i\tilde{x}^*_{it+h}(\vartheta_i) \\ + \sum_{j=1}^{p}\left(B_{1j}(\theta_i) - \rho'_i B_{2j}(\theta_i)\right)\left(\tilde{z}^*_{it+h-j}(\tau_i) - \zeta\beta'_{i0}\tilde{x}^*_{it+h-j}(d_i)\right) \\ + \tilde{v}^*_{1it+h}(d_i - d_{i0}) - \rho'_i\tilde{v}^*_{2it+h}(\vartheta_i - \vartheta_{i0}),\end{aligned} \tag{22}$$

in which both the dependent variable and regressors need to be forecast. This error-correction representation of the system may form the basis of iterative forecasts, including in the case of possible fractional cointegration among the original series y_{it} and x_{it}, in which case the cointegration can potentially be used to improve forecasting performance. In general, the h-step-ahead iterative forecast error exhibits an $MA(h-1)$ structure in the usual way, in addition to stationary long-range dependence as discussed for (18). For the iterative forecasting scheme, a vector autoregressive moving average (VARMA) specification or seemingly unrelated regression estimation (SURE) can be used, as in Pesaran et al. (2011). However, VARMA models are not commonly used in practice and can have stability and convergence problems in the face of large-dimensional data.

A. Panel Diebold-Mariano test

In order to construct the test statistic, we concentrate on the direct forecasting scheme and give the details accordingly. This is due to the fact that for iterative forecasting, the same asymptotic arguments follow under Assumption B by taking further into account the resulting MA forecast errors, and the main ideas are better motivated under the direct scheme whose treatment avoids further notational complexity.

We use $\overline{\tilde{y}^*_{ih}(\tau_i)} = m^{-1}\sum_{t=h+1}^{m+h}\tilde{y}^*_{it}(\tau_i)$ as a consistent estimator for $\mathbb{E}\left[\tilde{y}^*_{it}(\tau_i)\right]$, focusing on information in the evaluation sample. Denote

$$u_{it+h}(\tau_i) = \tilde{y}^*_{it+h}(\tau_i) - \overline{\tilde{y}^*_{ih}(\tau_i)}. \tag{23}$$

Then, to be able to work with the forecasting functions $\tilde{y}^*_{it+h|t}(\tau_i)$ and $\hat{\tilde{y}}^*_{it+h|t}(\hat{\tau}_i)$, we adapt conditions used by Breitung and Knüppel (2018) to our panel setting in the following.

Assumption C (Forecasting functions).

1. *Under the null hypothesis in (19), $u_{it+h}(\tau_i)$ is independent of the estimation error; $\mathbb{E}[u_{it+h}|\hat{\tau}_{is} - \tau_i] = 0$ for all i and $s = t, t-1, \ldots$.*
2. *The parameter vector τ_i is consistently estimated and the convergence rates satisfy $\hat{\tau}_{i0} - \tau_i = O_p(T^{-1/2})$ and $\hat{\tau}_{it} - \hat{\tau}_{i0} = O_p(\sqrt{t}/T)$, $t = 1, \ldots, m$.*

Assumption C.1 states that the time series is not predictable given the information set at time t which includes the estimation error $\hat{\tau}_{it} - \tau_i$ and is implied by (19). The conditions in Assumption C.2 set the usual convergence rate in the in-sample period and limit the variation in the recursive estimation, respectively. The first condition in Assumption C.2 is shown to be satisfied using conditional-sum-of-squares estimation for memory parameters under Assumptions A and B by Ergemen (2019).

Note also that we can only observe the actual forecast error,

$$\begin{aligned} \hat{e}_{it+h|t}(\hat{\tau}_i) &:= \hat{\tilde{y}}^*_{it+h}(\hat{\tau}_i) - \hat{\tilde{y}}^*_{it+h|t}(\hat{\tau}_i) \\ &= \hat{\tilde{v}}^*_{1it+h}\left(\hat{d}_i - d_{i0}\right) - \hat{\rho}_i' \hat{\tilde{v}}^*_{2it+h}\left(\hat{\vartheta}_i - \vartheta_{i0}\right). \end{aligned} \tag{24}$$

Define the loss differential for fixed i as

$$\hat{\varphi}^h_{it} = \hat{e}^2_{it+h|t}(\hat{\tau}_i) - \hat{u}^2_{it+h}(\hat{\tau}_i)$$

and let $\hat{\omega}^2_\varphi$ denote the consistent long-run variance estimator with long-memory/anti-persistence correction (MAC) due to Robinson (2005) applied to $\hat{\varphi}^h_{it}$. Then we can construct a Diebold-Mariano type test as

$$DM_{ih} = m^{1/2-\hat{\kappa}_i} \frac{1}{\hat{\omega}_\varphi m} \sum_{t=1}^{m} \hat{\varphi}^h_{it} \tag{25}$$

for $i = 1,\ldots,n$, where $\hat{\kappa}_i$ is a consistent estimator of the integration order of the squared loss differential $\hat{\varphi}_{it}^h$, which may be determined cf. Propositions 2-4 of Kruse et al. (2018). Here, it is important to note that under Assumptions A and B, Ergemen (2019) establishes the $\sqrt{T}$-consistency of both individual memory estimates, $\hat{d}_i$ and $\hat{\vartheta}_i$, so if we consider, e.g., the full-sample estimation, $\hat{\kappa}_i$ also enjoys the $\sqrt{T}$-consistency under any scenario prescribed in Propositions 2-4 of Kruse et al. (2018). This means that for all i, the $(\log T)(\hat{\kappa}_i - \kappa_i) = o_p(1)$ condition as imposed by Robinson (2005) and Abadir, Distaso, and Giraitis (2009) for consistent estimation of ω_φ^2 is naturally satisfied. In practice, it can be much simpler to estimate κ_i by resorting to semi-parametric, e.g., local Whittle, methods.

It should be noted that it is also possible to consider a modified version of (19) for comparing the model forecast to an external forecast if we write

$$H_0^\dagger : \mathbb{E}\left[e_{it+h|t}^2(\tau_i)\right] \geq \mathbb{E}\left[\xi_{it+h|t}^2\right], \tag{26}$$

where $\xi_{it+h|t}^2$ corresponds to the squared forecast error resulting from an external forecast, e.g., a survey, for cross-section unit i. The methodologies described here can be adapted to test (26), noting that the memory of the loss differential, κ_i, needs to be estimated in this case. It cannot be deduced based on the model memory parameter estimates, since the memory transmission rules listed in Kruse et al. (2018) are no longer guaranteed to apply.

To study the asymptotic behavior of the test statistic, we further impose the following rate conditions.

Assumption D (Rate conditions for DM type test). *As* $(n,T,m)_j \to \infty$,

$$m^{1/2-\kappa_i} n^{-1} + m^{1/2-\kappa_i} n^{-1/2} T^{-1/2} + m^{3/2-3\kappa_i} T^{-1/2} \to 0$$

and $\kappa_i \in (-1/2, 1/2)$ *for all* i.

The requirement on the relative asymptotic size of the number of forecasts and the number of cross-section units, $m^{1/2-\kappa_i} n^{-1} + m^{1/2-\kappa_i} n^{-1/2} T^{-1/2} \to 0$, is to control for the projection error, see Appendix C. The condition $m^{3/2-3\kappa_i} T^{-1/2} \to 0$ ensures that the first-order approximations of conditional forecasts depending on estimated memory parameters around the true memory parameters work well, imposing that there is enough information gathered in the estimation period so that forecasts can be evaluated in the following m periods. The condition on κ_i is imposed in order to include the case where the predictive accuracy comparison is made to an external forecast, i.e., when testing (26) (it is automatically satisfied when testing the null in (19)). This condition also enables the estimation of the long-run variance of the loss differential using readily available techniques such as local Whittle methods, see Robinson (2005) and Abadir et al. (2009).

The following result establishes the asymptotic distribution of the test statistic in (25).

Theorem 1. *Under Assumptions A-D and H_0 in (19), as $(n,T,m)_j \to \infty$,*

$$DM_{ih} \Rightarrow \frac{|w_i|}{2},$$

for fixed i, where w_i is a standard normally distributed random variable.

This result states that for each cross-section unit i, the Diebold-Mariano type (henceforth DM) test statistic converges weakly to a random variable that has an asymptotic half-normal distribution with unit variance under the null hypothesis in (19). Breitung and Knüppel (2018) obtain the same limiting result, but under a setup in which the series and thus the forecasts are $I(0)$ time series, imposing only that $m/T \to 0$ as both m and T diverge, in contrast to our Assumption D. Therefore, showing the result in Theorem 1 is quite different under our setup, particularly because of the estimation of the long-run variance, due to the allowance for long memory as well as the proxying for the common-factor structure in the panel.

Corollary 1. *Under the conditions of Theorem 1,*

$$2DM_{ih} \xrightarrow{d} |N(0,1)|,$$
$$\widetilde{DM}_{ih} = m^{1-2\hat{\kappa}_i} \frac{1}{\hat{\omega}_{iu}^2 m} \sum_{t=1}^{m} \hat{\varphi}_{it}^h \xrightarrow{d} \chi_1^2,$$

for fixed i, where $\hat{\omega}_{iu}^2$ is a consistent long-run variance estimator of u_{it+h} as defined in (23).

These two test statistics are in spirit of the adjusted DM test statistic in (25) and are direct consequences of the result in Theorem 1. It is important to note here that the null hypothesis is rejected for smaller values of the test statistic $\widetilde{DM}_{ih}$, which contrasts with most chi-squared distributed test statistics, see also Breitung and Knüppel (2018) for a discussion.

The DM type test statistic in (25) can be adopted for use under (26), yielding the same asymptotic properties, if we replace $\hat{\varphi}_{it}^{h}$ by $\hat{\varphi}_{it}^{h\dagger}$ defined as

$$\hat{\varphi}_{it}^{h\dagger} = \hat{e}^2_{it+h|t}(\hat{\tau}_i) - \hat{\xi}^2_{it+h|t},$$

where $\hat{\xi}^2_{it+h|t}$ is the (actual) observed squared forecast error from the external forecast, and $\hat{\omega}^2_{\varphi}$ by $\hat{\omega}^2_{\varphi\dagger}$, obtained by applying the MAC estimator to $\hat{\varphi}_{it}^{h\dagger}$.

B. Panel encompassing test

The DM type and related test statistics considered so far encounter size problems in small samples due to the null in (19) being rejected for small values of the test statistic, see also Breitung and Knüppel (2018) and our simulation results in Section IV. Furthermore, the rate condition $m^{3/2-3\kappa_i}T^{-1/2} \to 0$ in Assumption D can be too stringent, particularly when $\kappa_i \in (-1/2, 0)$. To offer a remedy, we reformulate the null hypothesis as

$$H_0' : \mathbb{E}\left[\left(\tilde{y}^*_{it+h}(\tau_i) - \tilde{y}^*_{it+h|t}(\tau_i)\right)\left(\tilde{y}^*_{it+h|t}(\tau_i) - \overline{\tilde{y}^*_{ih}(\tau_i)}\right)\right] = 0, \tag{27}$$

which, for a rational forecast satisfying $\mathbb{E}\left[\tilde{y}^*_{it+h}(\tau_i) - \tilde{y}^*_{it+h|t}(\tau_i)\middle|\tilde{y}^*_{it+h|t}(\tau_i)\right] = 0$, is equivalent to (19). The null hypothesis is rejected when there is positive correlation between $\tilde{y}^*_{it+h}(\tau_i)$ and $\tilde{y}^*_{it+h|t}(\tau_i)$. To motivate this further, writing

$$\begin{aligned}\sum_{t=1}^{m} \varphi_{it}^{h} &= \sum_{t=1}^{m}\left[\tilde{y}^*_{it+h}(\tau_i) - \overline{\tilde{y}^*_{ih}(\tau_i)} - (\tilde{y}^*_{it+h|t}(\tau_i) - \overline{\tilde{y}^*_{ih}(\tau_i)})\right]^2 - \left(\tilde{y}^*_{it+h}(\tau_i) - \overline{\tilde{y}^*_{ih}(\tau_i)}\right)^2 \\ &= \sum_{t=1}^{m}\left(\tilde{y}^*_{it+h|t}(\tau_i) - \overline{\tilde{y}^*_{ih}(\tau_i)}\right)^2 - 2\sum_{t=1}^{m}\left[(\tilde{y}^*_{it+h}(\tau_i) - \overline{\tilde{y}^*_{ih}(\tau_i)})(\tilde{y}^*_{it+h|t}(\tau_i) - \overline{\tilde{y}^*_{ih}(\tau_i)}\right]\end{aligned}$$

shows the link to the DM type test statistic. The covariance between $\tilde{y}^*_{it+h}(\tau_i)$ and $\tilde{y}^*_{it+h|t}(\tau_i)$ plays the key role in terms of deciding the power of the test, since the first term is non-negative.

Our approach may be considered as a one-sided Mincer-Zarnowitz regression,

$$\tilde{y}^*_{it+h}(\tau_i) = \phi_{i0,h} + \phi_{i1,h}\tilde{y}^*_{it+h|t}(\tau_i) + e_{it+h},$$

in which we test $\phi_{i1,h} = 0$ versus $\phi_{i1,h} > 0$ for fixed i and unrestricted $\phi_{i0,h}$. In the literature, the errors from similar regressions are typically mean zero $I(0)$ processes, and this is asymptotically so for e_{it+h} in our setup, given the consistency of $\hat{\tau}_i$. We consider the null as one of uninformativeness of the forecast, whereas Mincer and Zarnowitz (1969) focused on the joint null of informativeness and unbiasedness, $\phi_{i1,h} = 1$ and $\phi_{i0,h} = 0$. Our test can also be seen as a forecast encompassing test by writing

$$\begin{aligned}\tilde{y}^*_{it+h}(\tau_i) &= \psi_{ih}\tilde{y}^*_{it+h|t}(\tau_i) + (1-\psi_{ih})\overline{\tilde{y}^*_{ih}(\tau_i)} + e_{it+h}\\ \tilde{y}^*_{it+h}(\tau_i) - \overline{\tilde{y}^*_{ih}(\tau_i)} &= \psi_{ih}(\tilde{y}^*_{it+h|t}(\tau_i) - \overline{\tilde{y}^*_{ih}(\tau_i)}) + e_{it+h},\end{aligned}$$

since testing for $\phi_{i1,h} = 0$ is the same as testing for $\psi_{ih} = 0$. Further, in analogy with Breitung and Knüppel (2018), our DM type statistics can be interpreted as likelihood ratio tests of the uninformativeness null in the Mincer-Zarnowitz or encompassing regressions, against the joint informativeness and unbiasedness alternative.

We focus here on the encompassing type test statistic given by

$$\varrho_{ih} = m^{1/2-\hat{\upsilon}_i}\frac{1}{\hat{\omega}_\Xi m}\sum_{t=1}^{m}\hat{\Xi}^h_{it}, \tag{28}$$

essentially an LM type statistic, where

$$\hat{\Xi}^h_{it} = (\hat{\tilde{y}}^*_{it+h}(\hat{\tau}_i) - \overline{\hat{\tilde{y}}^*_{ih}(\hat{\tau}_i)})(\hat{\tilde{y}}^*_{it+h|t}(\hat{\tau}_i) - \overline{\hat{\tilde{y}}^*_{ih}(\hat{\tau}_i)}),$$

$\hat{\upsilon}_i$ is a consistent memory estimate of $\hat{\Xi}^h_{it}$ satisfying $(\log T)(\hat{\upsilon}_i - \upsilon_i) = o_p(1)$, and $\hat{\omega}^2_\Xi$ is the MAC-robust long-run variance estimator of Robinson (2005) applied to $\hat{\Xi}^h_{it}$. We impose the following condition to study the asymptotic behavior of the test statistic in (28).

Assumption E (Rate conditions for encompassing type test). *As* $(n,T,m)_j \to$

∞,

$$m^{1/2-v_i}n^{-1} + m^{1/2-v_i}n^{-1/2}T^{-1/2} + mT^{-1} \to 0,$$

and $v_i \in (-1/2, 1/2)$ *for all* i.

The first two terms ensure that the projection errors in the panel setting do not have any asymptotic contribution, see Appendix C. The third condition, $m/T \to 0$, is standard and is also imposed by Breitung and Knüppel (2018). It simply states that the out-of-sample length must be smaller than the in-sample length so that there is enough information at hand for prediction.

The next result establishes the asymptotic behavior of the encompassing type test.

Theorem 2. *Under Assumptions A-C and E,* H_0' *in (27), and a recursive estimation scheme, as* $(n, T, m)_j \to \infty$,

$$\varrho_{ih} \xrightarrow{d} N(0,1),$$

for fixed i.

The null in (27) is rejected when ϱ_{ih} is large compared to the critical value from the standard normal distribution.

Given our panel setup, it is interesting in addition to analyze the cross-sectional average of the test statistic in (28). Noting that, as $n \to \infty$, the projection errors, of size $O_p(n^{-1} + (nT)^{-1/2})$, become $o_p(1)$, making the cross-section units asymptotically independent of each other under Assumption B.2. Thus, the individual ϱ_{ih} test statistics are asymptotically approximately independent for large n. Thus, we can simply consider the test statistic

$$\overline{\varrho}_h := n^{-1/2} \sum_{i=1}^{n} \varrho_{ih}, \tag{29}$$

based on the first result in Theorem 2, in a similar spirit to the CIPS test statistic proposed by Pesaran (2007). We present the asymptotic behavior of $\overline{\varrho}_h$ in the next result.

Theorem 3. *Under the conditions of Theorem 2,*

$$\bar{\varrho}_h \xrightarrow{d} N(0,1).$$

Theorem 3 shows that the cross-sectionally averaged test statistic is asymptotically distributed as standard normal, under the conditions of Theorem 2. Note that although $\bar{\varrho}_h$ uses equal weighting, it is also possible to allow for different weights for cross-section units, as in Chudik et al. (2016), and the asymptotic normality result in Theorem 3 still holds under suitable conditions imposed on the weights, see, e.g., Pesaran (2006), although in this case the asymptotic mean and variance are characterized based on those weights. It would be also possible to consider the combination of p-values of the individual encompassing test statistics. For example, the inverse chi-squared test statistics defined by

$$P(n,T) = -2\sum_{i=1}^{n} ln(p_{iT}),$$

where p_{iT}, the p-value corresponding to cross-section unit i, see also Pesaran (2007), can be used when n is large.

C. Local power analysis

In order to study the local power properties of DM and encompassing type tests, we work with a restricted version of (16) in which $\rho_i = 0$ and $B_j(\theta_i) = 0$, for all i, so that we end up with a predictive regression setup taking $h = 1$,

$$\tilde{y}^*_{it+1}(\tau_i) = \beta'_{i0}\tilde{x}^*_{it}(d_i) + \tilde{v}^*_{1it+1}(d_i - d_{i0}). \tag{30}$$

Our choice in (30) is motivated by a desire to contrast our setup to the popular predictive regression setup in the literature, and we present Monte Carlo results and an empirical application to this case in the following sections.

Under Assumptions A.1 and B.2, using the asymptotic independence of $\tilde{v}^*_{1it+1}(d_i - d_{i0})$ and $\{\tilde{x}^*_{it}(d_i), \tilde{x}^*_{it-1}(d_i), \ldots\}$, as $T \to \infty$, we have that

$$\mathbb{E}\left[\left(\tilde{y}^*_{it+1}(\tau_i) - \overline{\tilde{y}^*_{i1}(\tau_i)}\right)^2\right] = \sigma^2_{v_1} + \beta^2_{i0}\sigma^2_x.$$

So, if $\beta_{i0} \neq 0$, the forecast is informative, and $\widetilde{DM}_{i1}$ and ϱ_{i1} are $O_p(m^{-1/2})$. Accordingly, both DM and encompassing type tests are consistent against fixed alternatives $\beta_{i0} \neq 0$. We consider local alternatives of the form $\beta_{i0} = c_i/\sqrt{m}$, for all i, extending the case in Breitung and Knüppel (2018) to the panel setup with factor projection and long-range dependence. It is also possible to consider deviations in short/long-range dependence, as well as contemporaneous correlation parameters, but we focus on the simplest case to show our tests have nontrivial local power.

In relation to the aggregate test statistic in (29), we note that it is possible to allow for $c_i = 0$ for some non-negligible, but non-dominating, fraction of the cross-section units, as in Su and Chen (2013), under further regularity conditions, but this is beyond the scope of the current paper.

The next result establishes the local asymptotic behavior of both DM and encompassing type test statistics.

Theorem 4. *Under the conditions of Theorem 2 for $\widetilde{DM}_{i1}$ and additionally imposing $m/\sqrt{T} \to 0$ for ϱ_{i1}, and $\beta_{i0} = c_i/\sqrt{m}$ in (30), as $(n,T,m)_j \to \infty$,*

$$\widetilde{DM}_{i1} \xrightarrow{d} w_{1i}^2 - 2\eta_i w_{2i} - \eta_i^2 \tag{31}$$

$$\varrho_{i1} \xrightarrow{d} sign(c_i) w_{2i} + \eta_i \tag{32}$$

for fixed i where $\eta_i^2 = c_i^2\sigma_x^2/\sigma_{v_1}^2$ is the signal-to-noise ratio and w_{1i} and w_{2i} are two independent standard normally distributed random variables.

This result states that both DM and encompassing type test statistics have non-trivial power against local alternatives. The rate condition is the same as the one imposed in Theorem 4 of Breitung and Knüppel (2018) under the alternative model considered in (30) since it carries along the error term $\tilde{v}_{1it+1}^*(d_i - d_{i0})$, which is stationary under Assumption A.1. Figure 1 depicts the resulting power curves for $c_i \geq 0$ for three conventional significance levels. In each case, the DM type test is more powerful in the vicinity of the null than the encompassing style test, whereas the encompassing type test is most powerful when the signal-to-noise ratio grows large. Note also that the power curves are symmetric in c since the asymptotic distribution in (31) and (32) is unchanged by replacing w_2 with $-w_2$.

IV. Monte Carlo study

In this section, we carry out a Monte Carlo experiment to study the finite-sample properties of our proposed tests. The simulation experiment builds upon the setup considered by Ergemen (2019). We simulate scalar y_{it} and x_{it} and draw the idiosyncratic vector $(\varepsilon_{1it}, \varepsilon_{2it})'$ as standard normal with covariance matrix

$$\Omega = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix},$$

with signal-to-noise ratio $s = a_{22}/a_{11}$ and correlation $\rho = a_{12}/(a_{11}a_{22})^{1/2}$. Without loss of generality, we set $a_{11} = 1$ and introduce short-memory dynamics via $B_1(\theta_i) = \text{diag}\{\theta_{1i}, \theta_{2i}\}$. We generate a serially correlated common factor via $f_t = 1/2 f_{t-1} + \Delta_t^{-\delta} z_t^f$ based on iid innovations z_t^f drawn as standard normal and then fractionally integrated to the order δ. The individual fixed effects are left unspecified since they are removed via first differencing and projections are based on the first-differenced data. We then obtain y_{t+1} from the general DGP

$$\begin{aligned} y_{it+1} &= \alpha_i + \beta_{i0} x_{it} + \lambda_i f_t + \Delta_{t+1}^{-d_{i0}} \varepsilon_{1it+1}, \\ x_{it} &= \mu_i + \gamma_i f_t + \Delta_t^{-\vartheta_{i0}} \varepsilon_{2it}. \end{aligned} \tag{33}$$

Under the null, y_{it+1} is obtained by setting $\beta_{i0} = 0$, $\theta_1 = \theta_2 = 0$, and $\rho = 0$ for all i, whereas forecasts are based on the predictive model including x_{it} as a predictor. This model is estimated recursively, using information known through t, and forecasts are made in a direct manner according to (17). We employ an alternative that fixes $\theta_1 = \theta_2 = 0$ and $\rho = 0$. The memory orders κ_i and υ_i of the loss series $\hat{\varphi}_{it}^h$ and $\hat{\Xi}_{it}^h$ are estimated on the evaluation sample with the CSS parametric estimator from Ergemen and Velasco (2017). We consider two choices of bandwidth length, $\lfloor m^q \rfloor$, in the MAC estimator, setting $q = 3/5$ and $q = 4/5$, where the latter corresponds to the MSE optimal boundary (Abadir et al., 2009). We focus on different cross-section, in-sample and out-of-sample sizes, n, T, and m, and consider $d_{i0} = 0.4, 1$ and $\vartheta_{i0} = 0.4, 0.7, 1$. For this study, we fix $s = 1$ and $d^* = 1$. Simulations are carried out via 20,000 replications.

In terms of size, Tables 1 and 2 report results for the DM style tests from Corollary 1, $2DM_{ih}$ and $\widetilde{DM}_{ih}$, respectively, with $n = 20$ and various combinations of the length of the in-sample and out-of-sample windows. The DM style tests suffer from considerable oversizing. This is not entirely surprising, since the critical values of the half-normal and the χ_1^2 distribution are very close to zero, leaving a considerable amount of probability mass in the vicinity of the critical values. That is, small asymptotically negligible terms may have a large impact on the size of the test. These results are also consistent with the findings in Breitung and Knüppel (2018). On the contrary, from Table 3, the size properties of the encompassing style test from Theorem 2, which rejects for large, positive values in the standard normal distribution, are much more reasonable, though somewhat conservative. The choice of bandwidth in the MAC estimator has little impact on the size properties of the Diebold-Mariano style tests, whereas the encompassing style test provides the best results when $q = 3/5$.

To examine the power properties of the tests, we consider a local departure from the null via $\beta_{i0} = c_i/\sqrt{m}$, motivated by the analysis of the local alternative in Theorem 4. For ease of exposition, we analyse the case in which the tests depart from the null for all i with $c_i = 2$ and $c_i = 5$, corresponding to $\beta_{i0} = 0.2$ and $\beta_{i0} = 0.5$, respectively, for $m = 100$. The second of the two values of β_{i0} is similar to that in Breitung and Knüppel (2018). Tables 4 and 5 show that when the departure from the null is proportional to $c_i = 2$, the power of the Diebold-Mariano style test is reasonably high, while the power of the encompassing style test (cf. Table 6) is moderate at the 5% and 10% significance levels and quite low otherwise. Increasing the length of the in-sample and out-of-sample window generally improves power. Also, $q = 3/5$ provides the best results. When $c_i = 5$ (Tables 7-9), the power of both Diebold-Mariano and encompassing style tests is generally high. According to the local power analysis in Theorem 4, the Diebold-Mariano and encompassing style tests have local asymptotic power equal to 76.2 and 63.0, respectively, for $c_i = 2$ and 99.0 and 99.9, respectively, for $c_i = 5$ and the given parameter values. These theoretical predictions correspond well to the experimental findings reported in Tables 4-9. Finally, Table 10 reports size and power properties for the pooled test statistic in (29). Its size properties are quite conservative, but the local power is high for both choices of c_i.

V. Predictive relation between stock market volatility and economic policy uncertainty

The relation between stock market return volatility and political uncertainty has gained considerable interest in the literature. Recent structural models of Pástor and

Veronesi (2012, 2013) (PV) show that increasing uncertainty about economic policy makes stocks more volatile. The main reason is that as uncertainty about government policy actions increases, stock markets become unsettled in view of the uncertain prospects of the economy. Such political uncertainty may be triggered by, among other things, elections (see, e.g., Bialkowski et al. (2008)) or economic crises (see, e.g., Schwert (1989) and Bittlingmayer (1998)). It may arguably be prominent not only in the USA, but also in Europe and the developing countries. In this empirical application we apply our methodology to the predictive relation between stock market volatility and economic policy uncertainty (EPU) in a multi-country panel analysis, treating the former as the forecast objective and the latter as the predictor. We use the Baker et al. (2016) EPU indices for 14 countries, quantifying newspaper coverage of policy-related economic uncertainty in a given month. We compute the realized volatility (RV) from daily returns on each stock market index within the month, obtained from Global Financial Data and Yahoo! Finance (see, e.g., Chernov (2007), Rapach et al. (2013), and Luo and Qi (2017) for use of similar data sources). The data set spans the time period 2001-2017, totalling 204 time series observations. We find that both RV and EPU exhibit cross-sectional variation in volatility characteristics, which we take into account by standardizing within country so that valid comparisons can be made on the estimates. Figure 2 depicts RV and EPU. Both series show spikes in accordance with the global financial crisis. The EPU of the UK rose sharply in 2016 at the time of the Brexit referendum and has remained elevated since then. Table 11 reports the full-sample integration order of RV and EPU for each country. Both RV and EPU are found to exhibit long-range dependence with integration orders in the neighborhood of 0.5. The estimates are strongly significant and exhibit cross-sectional variation, thus calling for the methods outlined in this paper.

Following PV, we specify a linear relation between RV and EPU, including a first-order short-range dependent component of RV,

$$\begin{aligned} RV_{it} &= \alpha_i + \beta_i EPU_{it} + \theta_i RV_{it-1} + \lambda_i' f_t + \Delta_t^{-d_i} \epsilon_{1it}, \\ EPU_{it} &= \mu_i + \gamma_i' f_t + \Delta_t^{-\vartheta_i} \epsilon_{2it}. \end{aligned} \tag{34}$$

In the analysis, we allow for an unobserved common factor structure that on average captures other relevant indicators for the study. This allowance can be considered a more flexible way of modelling the relation between RV and EPU, appropriately controlling for relevant factors, as opposed to adding separate observable series. We also allow the innovations of RV and EPU to carry correlation after accounting for this common factor structure. For example, Bittlingmayer (1998) notes that the highly volatile stock markets of the 1930s in the USA may very well have reflected a non-negligible probability that the USA would "go socialist", whereas a critic would argue that the highly volatile markets were driven mainly by the business cycle, eventually causing uncertainty about government (in)actions as response (Schwert, 1989).

With these specifications, we estimate (34) to get the results collected in Table 12 for a contemporaneous analysis. For comparison with PV, we also report estimates of the coefficient on EPU from a short-range regression of RV on EPU and its own lag. The structural model of PV suggests that the contemporaneous relation should be positive. The results indicate a positive relation for most countries, which is significant at conventional levels for Australia, Canada, Greece, Ireland, Japan, and Mexico. Across the entire panel, the average (mean-group) estimate is 0.025 with a (one-sided) p-value of 0.086. Interestingly, the relation is insignificant for the USA, which contrasts with the findings in PV of a strongly significant relation (also confirmed in our data set when applying their procedure, cf. Table 12). Our results differ in this regard because we account for the evident long-range dependent features in the data to obtain consistent slope estimates, and control for possible endogeneity of EPU with respect to RV, as well as for other relevant (global) factors.

Given the presence of this contemporaneous relation, which is the only one studied in PV, it is natural to ask whether it also translates into a predictive relation with relevance for, e.g., risk management, asset allocation, and policy-making. For instance, market unsettlement driven by current political uncertainty may likely persist for several periods. Specifically, we ask whether EPU is informative about future RV. To examine this, we construct direct forecasts with an expanding estimation scheme based on (17) and the same specification as above. We set the initial in-sample window to the 2001-2007 pre-crisis period, and for $h = 1,2,3,6,9,12$ construct the first forecast for month h of 2008. This splits data roughly equally, matching the parameter choices in the simulation study in Section IV. Based on the findings in the simulation study, we use the encompassing type ϱ_{ih} statistic from (28) with MAC bandwidth $\lfloor m^{3/5} \rfloor$ for inference on uninformativeness of EPU for future RV. Table 13 reports the results. We find that the contemporaneous relation documented above generally does not translate into an out-of-sample predictive relation. On the basis of our proposed inferential framework, uninformativeness of EPU for future RV is only rejected at conventional levels for the UK ($h = 1$), Greece ($h = 3$), Spain ($h = 6$), Mexico ($h = 6$), and Australia ($h = 12$). Moreover, the pooled statistic $\overline{\varrho}_h$ does not reject uninformativeness across the entire panel for any of the forecast horizons.

Assuming that no factor structure is present, Panel B of the table, there is some evidence of informativeness of EPU for future RV in the case of Canada, Spain, Greece, Japan, and Mexico at various forecast horizons. Indeed, the pooled statistic rejects uninformativeness at the 10% level for one-quarter and one-year ahead forecast horizons. This suggests that some (weak) informativeness of EPU for future RV is contained in the common (international) component, leaving, however, the domestic predictive relation mostly non-existent when accounted for (Panel A). We also ran the procedure including the global crisis period in the initial in-sample window such that the first forecast is constructed for month h of 2010. Conclusions were qualitatively unaltered.

In summary, our panel-wide treatment of the relation between stock market volatility and economic policy uncertainty generally supports the logic from PV of a contemporaneous relation between economic political uncertainty and stock market volatility. However, after a proper account of common factors (obtained from the panel structure of the data) as well as the presence of long-range dependence, the relation in the USA is insignificant, in contrast to the finding obtained using the short-range dependence, univariate procedure in PV. Moreover, using the encompassing style test of the present paper, we find only weak evidence of informativeness in economic policy uncertainty for future stock market volatility, and it is mainly contained in the common factor structure. These results demonstrate the importance of long-range dependence treatments in practice and underpin the value of panel data modelling.

VI. Concluding remarks

This paper develops a forecast evaluation framework for testing the null hypothesis that model-based forecasts at a given horizon are uninformative in panels under potential long-range dependence. We consider the fractionally integrated panel data system of Ergemen (2019) with individual stochastic components, cross-section dependence, and endogenous predictors. In this setup, a Diebold-Mariano style test statistic has a half-normal asymptotic distribution, but suffers from oversizing in finite samples. For an equivalent null hypothesis, an alternative test derived from the encompassing principle is reasonably sized. Both tests have non-trivial power for local departures from the null. We also provide natural generalizations for evaluating pooled uninformativeness of the model-based forecasts in the entire panel.

An interesting direction for future research is the estimation of the fractionally integrated latent factor structure. Estimates of the factors may be used in a plug-in scheme to exploit the potential predictive content in the factors themselves, see, e.g., Gonçalves et al. (2017). Asymptotic results for such factor estimates are, however, yet to be developed. Moreover, although our model framework allows for heterogeneity in parameters, it is not necessarily given that this results in superior forecasting performance relative to a homogeneous specification, see, e.g., Baltagi (2013). A thorough analysis of the relative merits of the two specifications is left for future research. Finally, it is also common in macroeconomics to employ a univariate autoregressive model as benchmark in forecasting exercises, see, e.g., Stock and Watson (2003), and our framework can be extended to this case, too.

References

ABADIR, K. M., W. DISTASO, AND L. GIRAITIS (2009): "Two estimators of the long-run variance: beyond short memory," *Journal of Econometrics*, 150(1), 56–70.

ANDERSEN, T. G., T. BOLLERSLEV, F. X. DIEBOLD, AND P. LABYS (2003): "Modeling and forecasting realized volatility," *Econometrica*, 71, 579–625.

BAI, J. (2009): "Panel data models with interactive fixed effects," *Econometrica*, 77(4), 1229–1279.

BAKER, S. R., N. BLOOM, AND S. J. DAVIS (2016): "Measuring economic policy uncertainty," *The Quarterly Journal of Economics*, 131, 1593–1636.

BALTAGI, B. H. (2013): "Panel data forecasting," in *Handbook of economic forecasting*, ed. by G. Elliott and A. Timmermann, North Holland, vol. 2B, 995–1024, 1st ed.

BHARDWAJ, G. AND N. R. SWANSON (2006): "An empirical investigation of the usefulness of ARFIMA models for predicting macroeconomic and financial time series," *Journal of Econometrics*, 131, 539–578.

BIALKOWSKI, J., K. GOTTSCHALK, AND T. P. WISNIEWSKI (2008): "Stock market volatility around national elections," *Journal of Banking and Finance*, 32, 1941–1953.

BITTLINGMAYER, G. (1998): "Output, stock volatility, and political uncertainty in a natural experiment: Germany 1880-1940," *The Journal of Finance*, 53, 2243–2257.

BOLLERSLEV, T., D. OSTERRIEDER, N. SIZOVA, AND G. TAUCHEN (2013): "Risk and return: long-run relationships, fractional cointegration, and return predictability," *Journal of Financial Economics*, 108(2), 409–424.

BOS, C. S., P. H. FRANSES, AND M. OOMS (2002): "Inflation, forecast intervals and long memory regression models," *International Journal of Forecasting*, 18, 243–264.

BREITUNG, J. AND M. KNÜPPEL (2018): "How far can we forecast? Statistical tests of the predictive content," *Deutsche Bundesbank working paper*.

BUSCH, T., B. J. CHRISTENSEN, AND M. Ø. NIELSEN (2011): "The role of implied volatility in forecasting future realized volatility and jumps in foreign exchange, stock, and bond markets," *Journal of Econometrics*, 160, 48–57.

CANOVA, F. AND M. CICCARELLI (2004): "Forecasting and turning point predictions in a Bayesian panel VAR model," *Journal of Econometrics*, 120, 327–359.

CHERNOV, M. (2007): "On the role of risk premia in volatility forecasting," *Journal of Business and Economic Statistics*, 25, 411–426.

CHIRIAC, R. AND V. VOEV (2011): "Modelling and forecasting multivariate realized volatility," *Journal of Applied Econometrics*, 26, 922–947.

CHRISTENSEN, B. J. AND M. Ø. NIELSEN (2005): "Asymptotic normality of narrow-band least squares in the stationary fractional cointegration model and volatility forecasting," *Journal of Econometrics*, 133, 343–371.

CHUDIK, A., V. GROSSMAN, AND M. H. PESARAN (2016): "A multi-country approach to forecasting output growth using PMIs," *Journal of Econometrics*, 192, 349–365.

CLARK, T. E. AND M. W. MCCRACKEN (2006): "The predictive content of the output gap for inflation: resolving in-sample and out-of-sample evidence," *Journal of Money, Credit and Banking*, 1127–1148.

CORSI, F. (2009): "A simple approximate long-memory model of realized volatility," *Journal of Financial Econometrics*, 7, 174–196.

DEO, R., C. HURVICH, AND Y. LU (2006): "Forecasting realized volatility using a long-memory stochastic volatility model: estimation, prediction and seasonal adjustment," *Journal of Econometrics*, 131, 29–58.

DIEBOLD, F. X. AND R. S. MARIANO (1995): "Comparing predictive accuracy," *Journal of Business Economics and Statistics*, 13, 253–263.

ENGEL, C., N. C. MARK, AND K. D. WEST (2008): "Exchange rate models are not as bad as you think," *NBER Macroeconomics Annual 2007*, 22, 381–441, National Bureau of Economic Research.

ERGEMEN, Y., N. HALDRUP, AND C. RODRÍGUEZ-CABALLERO (2016): "Common long-range dependence in a panel of hourly Nord Pool electricity prices and loads," *Energy Economics*, 60, 79–96.

ERGEMEN, Y. E. (2019): "System estimation of panel data models under long-range dependence," *Journal of Business and Economics Statistics*, 37, 13–26.

ERGEMEN, Y. E. AND C. VELASCO (2017): "Estimation of fractionally integrated panels with fixed effects and cross-section dependence," *Journal of Econometrics*, 248–258.

GIACOMINI, R. AND H. WHITE (2006): "Tests of conditional predictive ability," *Econometrica*, 74, 1545–1578.

GIL-ALAÑA, L. AND P. ROBINSON (1997): "Testing of unit root and other nonstationary hypotheses in macroeconomic time series," *Journal of Econometrics*, 80(2), 241–268.

GONÇALVES, S., M. W. MCCRACKEN, AND B. PERRON (2017): "Tests of equal accuracy for nested models with estimated factors," *Journal of Econometrics*, 198, 231–252.

GROEN, J. J. J. (2005): "Exchange rate predictability and monetary fundamentals in a small multi-country panel," *Journal of Money, Credit and Banking*, 37, 495–516.

HJALMARSON, E. (2010): "Predicting global stock returns," *Journal of Financial and Quantitative Analysis*, 45, 49–80.

KRUSE, R., C. LESCHINSKI, AND M. WILL (2018): "Comparing predictive accuracy under long memory, with an application to volatility forecasting," *Journal of Financial Econometrics*, 1–49.

LIU, L. Y., A. J. PATTON, AND K. SHEPPARD (2015): "Does anything beat 5-minute RV? A comparison of realized measures across multiple asset classes," *Journal of Econometrics*, 187, 293–311.

LUO, X. AND X. QI (2017): "The dynamic correlations between the G7 economies and China: Evidence from both realized and implied volatilities," *Journal of Futures Markets*, 37, 989–1002.

MCCRACKEN, M. AND S. NG (2016): "FRED-MD: A monthly database for macroeconomic research," *Journal of Business and Economic Statistics*, 34, 574–589.

MINCER, J. A. AND V. ZARNOWITZ (1969): *The evalution of economic forecasts*, NBER, chap. 1, 1–46.

PÁSTOR, L. AND P. VERONESI (2012): "Uncertainty about government policy and stock prices," *The Journal of Finance*, 67, 1219–1264.

——— (2013): "Political uncertainty and risk premia," *Journal of Financial Economics*, 110, 520–545.

PESARAN, H. (2006): "Estimation and inference in large heterogeneous panels with a multifactor error structure," *Econometrica*, 74(4), 967–1012.

——— (2007): "A simple panel unit root test in the presence of cross-section dependence," *Journal of Applied Econometrics*, 22(2), 265–312.

PESARAN, M. H., A. PICK, AND A. TIMMERMANN (2011): "Variable selection, estimation and inference for multi-period forecasting problems," *Journal of Econometrics*, 164(1), 173–187.

PESARAN, M. H., T. SCHUERMANN, AND L. V. SMITH (2009): "Forecasting economic and financial variables with global VARs," *International Journal of Forecasting*, 25, 642–675.

RAPACH, D. E., J. K. STRAUSS, AND G. ZHOU (2013): "International stock return predictability: What is the role of the United States?" *The Journal of Finance*, 68, 1633–1662.

ROBINSON, P. M. (2005): "Robust covariance matrix estimation : 'HAC' estimates with long memory/antipersistence correction," *Econometric Theory*, 21(1), 171–180.

ROBINSON, P. M. AND J. HUALDE (2003): "Cointegration in fractional systems with unknown integration orders," *Econometrica*, 71(6), 1727–1766.

Robinson, P. M. and C. Velasco (2015): "Efficient estimation of fractionally integrated panel data models with fixed effects," *Journal of Econometrics*, 185, 435–452.

Schwert, G. W. (1989): "Business cycles, financial crisis, and stock volatility," *Carnegie-Rochester Conference Series on Public Policy*, 31, 83–125, National Bureau of Economic Research.

Stambaugh, R. F. (1999): "Predictive regressions," *Journal of Financial Economics*, 54, 375–421.

Stock, J. H. and M. W. Watson (1999): "Forecasting inflation," *Journal of Monetary Economics*, 44, 293–335.

——— (2003): "Forecasting output and inflation: the role of asset prices," *Journal of Economic Literature*, 41, 788–829.

Su, L. and Q. Chen (2013): "Testing homogeneity in panel data models with interactive fixed effects," *Econometric Theory*, 29, 1079–1135.

Welch, I. and A. Goyal (2008): "A comprehensive look at the empirical performance of equity premium prediction," *The Review of Financial Studies*, 21, 1455–1508.

Westerlund, J., H. Karabiyik, and P. Narayan (2016): "Testing for predictability in panels with general predictors," *Journal of Applied Econometrics*, 32, 554–574.

Westerlund, J. and P. Narayan (2015a): "A random coefficient approach to the predictability of stock returns in panels," *Journal of Financial Econometrics*, 13, 605–664.

——— (2015b): "Testing for predictability in conditionally heteroskedastic stock returns," *Journal of Financial Econometrics*, 13, 342–375.

A. Tables

A. Monte Carlo results

Table 1: Size of Diebold-Mariano style test, $2DM_{ih}$

This table reports size for test at level 10%, 5%, and 1% of the $2DM_{ih}$ statistic with $n = 20$ ($\delta = 0.4, \theta_1 = \theta_2 = 0, \rho = 0$), various integration orders d_i and ϑ_i and length of in-sample (T in the column) and out-of-sample (m in the row) windows. Panel A (B) contains results for MAC estimator bandwidth $\lfloor m^{3/5} \rfloor$ ($\lfloor m^{4/5} \rfloor$). For $T = \infty$ the test statistics are computed using true parameter values in the in-sample window. Results are based on 20,000 simulations.

	Significance level = 10%						*Significance level = 5%*						*Significance level = 1%*					
	$d_i = 0.4$			$d_i = 1$			$d_i = 0.4$			$d_i = 1$			$d_i = 0.4$			$d_i = 1$		
T\m	50	100	200	50	100	200	50	100	200	50	100	200	50	100	200	50	100	200
							Panel A: MAC bandwidth = $\lfloor m^{3/5} \rfloor$											
$\vartheta_i = 0.4$:																		
50	0.159	0.136	0.129	0.147	0.120	0.106	0.150	0.130	0.121	0.139	0.113	0.100	0.144	0.125	0.116	0.134	0.108	0.096
100	0.179	0.166	0.155	0.167	0.156	0.138	0.170	0.157	0.148	0.159	0.149	0.132	0.163	0.150	0.142	0.153	0.143	0.126
200	0.188	0.195	0.191	0.178	0.173	0.172	0.178	0.186	0.182	0.169	0.165	0.164	0.171	0.180	0.175	0.163	0.158	0.156
400	0.197	0.209	0.222	0.191	0.194	0.194	0.188	0.199	0.212	0.181	0.185	0.185	0.180	0.192	0.205	0.174	0.178	0.178
∞	0.093	0.096	0.100	0.088	0.091	0.092	0.045	0.049	0.049	0.043	0.047	0.045	0.008	0.011	0.010	0.008	0.010	0.010
$\vartheta_i = 0.7$:																		
50	0.150	0.136	0.124	0.150	0.126	0.107	0.142	0.127	0.115	0.143	0.119	0.099	0.136	0.120	0.108	0.137	0.115	0.094
100	0.166	0.159	0.143	0.166	0.157	0.140	0.157	0.150	0.135	0.159	0.150	0.132	0.150	0.143	0.129	0.153	0.144	0.127
200	0.177	0.180	0.176	0.182	0.182	0.177	0.168	0.171	0.167	0.173	0.174	0.169	0.161	0.165	0.161	0.166	0.168	0.162
400	0.191	0.198	0.209	0.190	0.202	0.200	0.180	0.188	0.200	0.180	0.192	0.189	0.172	0.182	0.193	0.173	0.185	0.182
∞	0.090	0.097	0.096	0.088	0.094	0.094	0.044	0.048	0.049	0.044	0.047	0.047	0.008	0.009	0.009	0.007	0.010	0.010
$\vartheta_i = 1$:																		
50	0.107	0.097	0.097	0.150	0.132	0.113	0.100	0.089	0.087	0.142	0.125	0.106	0.096	0.083	0.081	0.137	0.119	0.101
100	0.116	0.104	0.093	0.172	0.161	0.145	0.109	0.098	0.087	0.164	0.153	0.138	0.104	0.093	0.082	0.157	0.147	0.132
200	0.125	0.119	0.107	0.183	0.183	0.177	0.118	0.113	0.100	0.174	0.174	0.169	0.113	0.108	0.096	0.168	0.168	0.163
400	0.134	0.124	0.121	0.192	0.200	0.202	0.126	0.118	0.115	0.181	0.190	0.192	0.121	0.113	0.112	0.174	0.183	0.185
∞	0.069	0.069	0.066	0.087	0.090	0.095	0.035	0.034	0.034	0.044	0.045	0.048	0.007	0.006	0.007	0.009	0.009	0.009
							Panel B: MAC bandwidth = $\lfloor m^{4/5} \rfloor$											
$\vartheta_i = 0.4$:																		
50	0.158	0.136	0.128	0.146	0.120	0.106	0.150	0.130	0.120	0.139	0.112	0.100	0.144	0.125	0.116	0.134	0.108	0.096
100	0.178	0.165	0.155	0.166	0.155	0.138	0.170	0.156	0.148	0.159	0.148	0.132	0.163	0.150	0.142	0.153	0.143	0.126
200	0.187	0.195	0.191	0.177	0.172	0.172	0.178	0.185	0.182	0.169	0.165	0.163	0.171	0.179	0.175	0.163	0.158	0.156
400	0.196	0.208	0.222	0.190	0.194	0.194	0.187	0.199	0.212	0.181	0.184	0.185	0.180	0.192	0.205	0.174	0.177	0.178
∞	0.088	0.094	0.098	0.084	0.089	0.092	0.044	0.047	0.048	0.041	0.045	0.044	0.008	0.011	0.009	0.008	0.009	0.010
$\vartheta_i = 0.7$:																		
50	0.149	0.135	0.123	0.150	0.125	0.107	0.142	0.126	0.114	0.143	0.119	0.099	0.136	0.120	0.108	0.137	0.114	0.094
100	0.165	0.159	0.142	0.166	0.157	0.139	0.156	0.149	0.135	0.158	0.150	0.132	0.150	0.143	0.129	0.153	0.144	0.127
200	0.176	0.179	0.176	0.181	0.181	0.177	0.167	0.171	0.167	0.173	0.174	0.169	0.161	0.165	0.161	0.166	0.168	0.162
400	0.190	0.198	0.209	0.189	0.202	0.199	0.179	0.188	0.200	0.179	0.192	0.189	0.172	0.182	0.193	0.173	0.185	0.182
∞	0.086	0.095	0.094	0.084	0.092	0.093	0.043	0.047	0.048	0.042	0.045	0.046	0.008	0.009	0.009	0.007	0.009	0.010
$\vartheta_i = 1$:																		
50	0.106	0.097	0.096	0.149	0.132	0.113	0.100	0.088	0.087	0.142	0.124	0.106	0.096	0.083	0.081	0.137	0.119	0.101
100	0.115	0.104	0.093	0.172	0.161	0.145	0.109	0.097	0.087	0.163	0.153	0.138	0.104	0.094	0.082	0.157	0.147	0.132
200	0.125	0.118	0.106	0.183	0.183	0.177	0.118	0.113	0.100	0.174	0.174	0.169	0.113	0.108	0.096	0.167	0.168	0.163
400	0.134	0.124	0.120	0.191	0.199	0.202	0.126	0.117	0.115	0.181	0.190	0.192	0.120	0.113	0.112	0.174	0.183	0.185
∞	0.068	0.067	0.066	0.084	0.088	0.094	0.034	0.033	0.034	0.041	0.043	0.047	0.007	0.006	0.007	0.008	0.008	0.009

Table 2: Size of Diebold-Mariano style test, $\widetilde{DM}_{ih}$

This table reports size for test at level 10%, 5%, and 1% of the $\widetilde{DM}_{ih}$ statistic with $n = 20$ ($\delta = 0.4$, $\theta_1 = \theta_2 = 0, \rho = 0$), various integration orders d_i and ϑ_i and length of in-sample (T in the column) and out-of-sample (m in the row) windows. Panel A (B) contains results for MAC estimator bandwidth $\lfloor m^{3/5} \rfloor$ ($\lfloor m^{4/5} \rfloor$). For $T = \infty$ the test statistics are computed using true parameter values in the in-sample window. Results are based on 20,000 simulations.

	Significance level = 10%						*Significance level = 5%*						*Significance level = 1%*					
	$d_i = 0.4$			$d_i = 1$			$d_i = 0.4$			$d_i = 1$			$d_i = 0.4$			$d_i = 1$		
T\m	50	100	200	50	100	200	50	100	200	50	100	200	50	100	200	50	100	200
							Panel A: MAC bandwidth = $\lfloor m^{3/5} \rfloor$											
$\vartheta_i = 0.4$:																		
50	0.146	0.126	0.116	0.136	0.108	0.096	0.144	0.124	0.115	0.133	0.107	0.095	0.143	0.124	0.114	0.133	0.106	0.095
100	0.166	0.151	0.143	0.155	0.144	0.127	0.163	0.149	0.141	0.152	0.142	0.125	0.161	0.148	0.141	0.151	0.141	0.125
200	0.175	0.181	0.176	0.165	0.160	0.157	0.171	0.178	0.174	0.162	0.158	0.155	0.170	0.177	0.173	0.161	0.157	0.154
400	0.185	0.194	0.206	0.179	0.180	0.179	0.180	0.191	0.204	0.173	0.177	0.177	0.179	0.190	0.203	0.172	0.175	0.176
∞	0.093	0.096	0.100	0.088	0.091	0.093	0.045	0.049	0.049	0.043	0.047	0.045	0.008	0.011	0.010	0.008	0.010	0.010
$\vartheta_i = 0.7$:																		
50	0.139	0.121	0.109	0.140	0.115	0.095	0.136	0.119	0.107	0.137	0.114	0.094	0.135	0.119	0.107	0.136	0.113	0.093
100	0.154	0.145	0.130	0.155	0.145	0.128	0.150	0.142	0.129	0.153	0.143	0.126	0.149	0.142	0.128	0.151	0.143	0.126
200	0.164	0.166	0.162	0.170	0.169	0.163	0.160	0.164	0.160	0.166	0.168	0.161	0.159	0.163	0.160	0.165	0.167	0.161
400	0.178	0.185	0.194	0.177	0.188	0.184	0.173	0.181	0.192	0.172	0.184	0.181	0.171	0.180	0.191	0.171	0.183	0.180
∞	0.090	0.097	0.096	0.088	0.094	0.094	0.044	0.048	0.049	0.044	0.047	0.047	0.008	0.009	0.009	0.007	0.010	0.010
$\vartheta_i = 1$:																		
50	0.099	0.085	0.082	0.139	0.120	0.101	0.096	0.083	0.080	0.136	0.119	0.100	0.095	0.082	0.080	0.136	0.118	0.100
100	0.107	0.094	0.082	0.160	0.148	0.133	0.104	0.093	0.081	0.156	0.146	0.131	0.103	0.093	0.081	0.155	0.146	0.131
200	0.116	0.109	0.096	0.172	0.169	0.164	0.113	0.107	0.095	0.167	0.167	0.162	0.112	0.107	0.095	0.166	0.166	0.162
400	0.126	0.115	0.113	0.180	0.185	0.187	0.121	0.112	0.111	0.174	0.182	0.184	0.119	0.111	0.110	0.173	0.181	0.184
∞	0.070	0.069	0.066	0.087	0.090	0.095	0.035	0.034	0.034	0.044	0.045	0.048	0.007	0.006	0.007	0.009	0.009	0.009
							Panel B: MAC bandwidth = $\lfloor m^{4/5} \rfloor$											
$\vartheta_i = 0.4$:																		
50	0.146	0.126	0.116	0.136	0.108	0.096	0.143	0.124	0.115	0.133	0.107	0.095	0.143	0.124	0.114	0.133	0.106	0.095
100	0.166	0.151	0.143	0.155	0.143	0.126	0.163	0.149	0.141	0.152	0.142	0.125	0.161	0.148	0.141	0.151	0.141	0.125
200	0.175	0.181	0.176	0.165	0.160	0.157	0.171	0.178	0.174	0.162	0.158	0.155	0.170	0.177	0.173	0.161	0.157	0.154
400	0.184	0.194	0.206	0.178	0.180	0.179	0.180	0.191	0.204	0.173	0.177	0.177	0.179	0.190	0.203	0.172	0.175	0.176
∞	0.088	0.094	0.098	0.084	0.089	0.092	0.044	0.047	0.048	0.041	0.045	0.044	0.008	0.011	0.009	0.008	0.009	0.010
$\vartheta_i = 0.7$:																		
50	0.139	0.121	0.109	0.139	0.115	0.095	0.135	0.119	0.107	0.137	0.113	0.093	0.135	0.119	0.107	0.136	0.113	0.093
100	0.153	0.145	0.130	0.155	0.145	0.128	0.150	0.143	0.129	0.152	0.143	0.126	0.149	0.142	0.128	0.151	0.143	0.126
200	0.164	0.166	0.162	0.169	0.169	0.162	0.160	0.164	0.160	0.166	0.167	0.161	0.159	0.163	0.160	0.165	0.167	0.161
400	0.177	0.184	0.194	0.176	0.187	0.184	0.172	0.181	0.192	0.172	0.184	0.181	0.171	0.180	0.191	0.171	0.183	0.180
∞	0.086	0.095	0.094	0.084	0.092	0.093	0.043	0.047	0.048	0.042	0.045	0.046	0.008	0.009	0.009	0.007	0.009	0.010
$\vartheta_i = 1$:																		
50	0.098	0.085	0.082	0.139	0.120	0.101	0.096	0.083	0.080	0.136	0.119	0.100	0.095	0.082	0.080	0.136	0.118	0.100
100	0.106	0.094	0.082	0.160	0.148	0.133	0.104	0.093	0.081	0.156	0.146	0.131	0.103	0.093	0.081	0.155	0.146	0.131
200	0.116	0.109	0.096	0.171	0.169	0.163	0.113	0.107	0.095	0.167	0.167	0.162	0.112	0.107	0.095	0.166	0.166	0.162
400	0.125	0.115	0.112	0.179	0.185	0.187	0.121	0.112	0.111	0.174	0.182	0.184	0.119	0.111	0.110	0.173	0.181	0.184
∞	0.068	0.067	0.066	0.084	0.088	0.094	0.034	0.033	0.034	0.041	0.043	0.047	0.007	0.006	0.007	0.008	0.008	0.009

Table 3: Size of encompassing style test, ϱ_{ih}

This table reports size for test at level 10%, 5%, and 1% of the ϱ_{ih} statistic with $n = 20$ ($\delta = 0.4, \theta_1 = \theta_2 = 0, \rho = 0$), various integration orders d_i and ϑ_i and length of in-sample (T in the column) and out-of-sample (m in the row) windows. Panel A (B) contains results for MAC estimator bandwidth $\lfloor m^{3/5} \rfloor$ ($\lfloor m^{4/5} \rfloor$). Results are based on 20,000 simulations.

	Significance level = 10%						*Significance level = 5%*						*Significance level = 1%*					
	$d_i = 0.4$			$d_i = 1$			$d_i = 0.4$			$d_i = 1$			$d_i = 0.4$			$d_i = 1$		
T\m	50	100	200	50	100	200	50	100	200	50	100	200	50	100	200	50	100	200
							Panel A: MAC bandwidth = $\lfloor m^{3/5} \rfloor$											
$\vartheta_i = 0.4$:																		
50	0.076	0.074	0.081	0.070	0.067	0.074	0.032	0.034	0.039	0.032	0.032	0.038	0.005	0.006	0.008	0.005	0.006	0.007
100	0.083	0.083	0.090	0.079	0.076	0.078	0.037	0.040	0.048	0.037	0.037	0.038	0.006	0.007	0.009	0.006	0.006	0.007
200	0.089	0.091	0.099	0.081	0.083	0.090	0.042	0.044	0.053	0.037	0.039	0.046	0.006	0.007	0.011	0.005	0.007	0.010
400	0.099	0.101	0.114	0.089	0.091	0.099	0.046	0.049	0.059	0.039	0.042	0.050	0.006	0.009	0.011	0.007	0.007	0.009
$\vartheta_i = 0.7$:																		
50	0.072	0.067	0.073	0.069	0.072	0.076	0.030	0.030	0.037	0.031	0.033	0.038	0.005	0.005	0.007	0.005	0.005	0.006
100	0.079	0.081	0.079	0.080	0.084	0.080	0.036	0.040	0.040	0.035	0.039	0.038	0.005	0.007	0.007	0.005	0.006	0.008
200	0.086	0.088	0.094	0.086	0.086	0.092	0.038	0.043	0.048	0.037	0.041	0.046	0.006	0.008	0.009	0.006	0.006	0.009
400	0.094	0.102	0.106	0.094	0.096	0.099	0.045	0.050	0.054	0.042	0.046	0.052	0.007	0.009	0.011	0.006	0.008	0.011
$\vartheta_i = 1$:																		
50	0.060	0.062	0.068	0.072	0.074	0.076	0.026	0.029	0.032	0.031	0.035	0.036	0.004	0.005	0.006	0.004	0.005	0.006
100	0.066	0.066	0.078	0.082	0.080	0.084	0.031	0.033	0.039	0.037	0.038	0.042	0.005	0.006	0.008	0.006	0.006	0.008
200	0.076	0.079	0.082	0.083	0.083	0.093	0.035	0.039	0.043	0.037	0.039	0.047	0.006	0.008	0.010	0.006	0.006	0.009
400	0.086	0.085	0.094	0.093	0.095	0.101	0.039	0.044	0.051	0.044	0.046	0.052	0.006	0.008	0.012	0.007	0.007	0.011
							Panel B: MAC bandwidth = $\lfloor m^{4/5} \rfloor$											
$\vartheta_i = 0.4$:																		
50	0.068	0.071	0.078	0.064	0.063	0.070	0.022	0.027	0.035	0.021	0.026	0.033	0.000	0.001	0.004	0.000	0.002	0.004
100	0.075	0.077	0.087	0.072	0.071	0.076	0.024	0.031	0.042	0.025	0.029	0.033	0.000	0.002	0.005	0.001	0.002	0.004
200	0.083	0.085	0.097	0.074	0.078	0.087	0.027	0.037	0.047	0.025	0.033	0.042	0.000	0.002	0.005	0.001	0.002	0.005
400	0.092	0.095	0.108	0.082	0.086	0.094	0.031	0.039	0.052	0.026	0.033	0.044	0.000	0.002	0.006	0.001	0.002	0.005
$\vartheta_i = 0.7$:																		
50	0.067	0.062	0.070	0.063	0.067	0.073	0.019	0.024	0.033	0.022	0.027	0.033	0.000	0.002	0.003	0.001	0.001	0.003
100	0.072	0.077	0.077	0.073	0.078	0.077	0.024	0.032	0.035	0.022	0.031	0.035	0.001	0.002	0.004	0.000	0.002	0.004
200	0.079	0.084	0.091	0.079	0.082	0.087	0.027	0.035	0.042	0.026	0.032	0.040	0.001	0.002	0.005	0.000	0.002	0.005
400	0.088	0.099	0.103	0.086	0.090	0.096	0.031	0.041	0.048	0.028	0.035	0.045	0.000	0.003	0.006	0.001	0.002	0.006
$\vartheta_i = 1$:																		
50	0.055	0.058	0.064	0.064	0.070	0.072	0.017	0.024	0.028	0.022	0.026	0.032	0.000	0.002	0.003	0.000	0.001	0.003
100	0.062	0.064	0.075	0.075	0.074	0.080	0.022	0.027	0.034	0.023	0.031	0.037	0.000	0.002	0.004	0.000	0.002	0.004
200	0.071	0.075	0.080	0.077	0.079	0.090	0.025	0.031	0.039	0.025	0.032	0.042	0.001	0.003	0.006	0.000	0.001	0.005
400	0.080	0.080	0.091	0.085	0.091	0.098	0.026	0.035	0.047	0.030	0.036	0.046	0.001	0.003	0.007	0.000	0.002	0.005

Table 4: Power ($c_i = 2$) of Diebold-Mariano style test, $2DM_{ih}$

This table reports power for test at level 10%, 5%, and 1% of the $2DM_{ih}$ statistic with $n = 20$ ($\delta = 0.4$, $\theta_1 = \theta_2 = 0, \rho = 0$), various integration orders d_i and ϑ_i and length of in-sample (T in the column) and out-of-sample (m in the row) windows. Panel A (B) contains results for MAC estimator bandwidth $\lfloor m^{3/5} \rfloor$ ($\lfloor m^{4/5} \rfloor$). Results are based on 20,000 simulations.

	Significance level = 10%						*Significance level = 5%*						*Significance level = 1%*					
	$d_i = 0.4$			$d_i = 1$			$d_i = 0.4$			$d_i = 1$			$d_i = 0.4$			$d_i = 1$		
T\m	50	100	200	50	100	200	50	100	200	50	100	200	50	100	200	50	100	200
Panel A: MAC bandwidth = $\lfloor m^{3/5} \rfloor$																		
$\vartheta_i = 0.4$:																		
50	0.667	0.609	0.551	0.744	0.698	0.636	0.655	0.598	0.538	0.734	0.687	0.624	0.647	0.589	0.528	0.726	0.679	0.615
100	0.723	0.681	0.625	0.790	0.764	0.717	0.712	0.670	0.614	0.782	0.755	0.707	0.704	0.661	0.605	0.776	0.747	0.699
200	0.753	0.734	0.682	0.807	0.800	0.766	0.743	0.724	0.671	0.798	0.792	0.756	0.736	0.717	0.662	0.792	0.784	0.750
400	0.754	0.752	0.730	0.822	0.812	0.796	0.745	0.743	0.720	0.815	0.804	0.787	0.737	0.737	0.711	0.809	0.797	0.780
$\vartheta_i = 0.7$:																		
50	0.679	0.642	0.591	0.689	0.640	0.575	0.667	0.630	0.576	0.679	0.629	0.563	0.659	0.621	0.565	0.671	0.619	0.553
100	0.728	0.708	0.666	0.743	0.705	0.651	0.717	0.698	0.654	0.733	0.694	0.640	0.710	0.689	0.644	0.724	0.686	0.630
200	0.758	0.750	0.716	0.767	0.754	0.715	0.748	0.740	0.704	0.758	0.746	0.705	0.741	0.731	0.695	0.750	0.736	0.697
400	0.766	0.773	0.755	0.775	0.775	0.755	0.757	0.764	0.745	0.766	0.765	0.744	0.751	0.756	0.736	0.760	0.758	0.736
$\vartheta_i = 1$:																		
50	0.728	0.729	0.722	0.660	0.608	0.535	0.719	0.718	0.711	0.650	0.595	0.523	0.710	0.711	0.703	0.640	0.585	0.514
100	0.765	0.769	0.751	0.717	0.682	0.624	0.756	0.759	0.742	0.707	0.670	0.613	0.749	0.754	0.735	0.699	0.662	0.603
200	0.788	0.800	0.775	0.743	0.722	0.686	0.781	0.792	0.766	0.734	0.713	0.674	0.775	0.786	0.760	0.727	0.705	0.665
400	0.805	0.811	0.795	0.757	0.750	0.728	0.798	0.804	0.787	0.748	0.739	0.717	0.792	0.799	0.782	0.739	0.731	0.709
Panel B: MAC bandwidth = $\lfloor m^{4/5} \rfloor$																		
$\vartheta_i = 0.4$:																		
50	0.666	0.609	0.551	0.744	0.698	0.636	0.655	0.597	0.538	0.734	0.687	0.624	0.646	0.589	0.528	0.726	0.679	0.615
100	0.722	0.680	0.625	0.789	0.764	0.717	0.711	0.670	0.613	0.781	0.755	0.707	0.704	0.661	0.605	0.776	0.747	0.699
200	0.752	0.734	0.682	0.807	0.800	0.765	0.743	0.724	0.671	0.798	0.792	0.756	0.736	0.717	0.662	0.792	0.784	0.750
400	0.753	0.751	0.730	0.823	0.811	0.796	0.744	0.743	0.720	0.815	0.804	0.786	0.737	0.736	0.711	0.809	0.797	0.780
$\vartheta_i = 0.7$:																		
50	0.678	0.642	0.591	0.688	0.640	0.575	0.667	0.630	0.576	0.679	0.629	0.563	0.658	0.621	0.565	0.671	0.619	0.553
100	0.727	0.708	0.666	0.742	0.704	0.651	0.717	0.697	0.654	0.732	0.694	0.640	0.710	0.689	0.644	0.724	0.685	0.630
200	0.757	0.749	0.715	0.766	0.754	0.715	0.748	0.740	0.704	0.758	0.746	0.705	0.741	0.731	0.695	0.750	0.736	0.697
400	0.766	0.772	0.755	0.774	0.774	0.755	0.757	0.763	0.744	0.766	0.765	0.744	0.751	0.756	0.736	0.760	0.758	0.736
$\vartheta_i = 1$:																		
50	0.727	0.728	0.721	0.659	0.607	0.535	0.719	0.718	0.711	0.649	0.595	0.523	0.710	0.711	0.703	0.640	0.585	0.514
100	0.765	0.768	0.750	0.716	0.681	0.624	0.756	0.759	0.742	0.707	0.669	0.613	0.749	0.753	0.735	0.699	0.662	0.603
200	0.788	0.800	0.774	0.742	0.722	0.686	0.780	0.792	0.766	0.733	0.713	0.674	0.775	0.786	0.760	0.727	0.705	0.665
400	0.804	0.811	0.795	0.756	0.749	0.727	0.798	0.804	0.787	0.747	0.739	0.716	0.792	0.799	0.782	0.739	0.731	0.709

Table 5: Power ($c_i = 2$) of Diebold-Mariano style test, $\widetilde{DM}_{ih}$

This table reports power for test at level 10%, 5%, and 1% of the $\widetilde{DM}_{ih}$ statistic with $n = 20$ ($\delta = 0.4$, $\theta_1 = \theta_2 = 0, \rho = 0$), various integration orders d_i and ϑ_i and length of in-sample (T in the column) and out-of-sample (m in the row) windows. Panel A (B) contains results for MAC estimator bandwidth $\lfloor m^{3/5} \rfloor$ ($\lfloor m^{4/5} \rfloor$). Results are based on 20,000 simulations.

	Significance level = 10%						*Significance level = 5%*						*Significance level = 1%*					
	$d_i = 0.4$			$d_i = 1$			$d_i = 0.4$			$d_i = 1$			$d_i = 0.4$			$d_i = 1$		
T\m	50	100	200	50	100	200	50	100	200	50	100	200	50	100	200	50	100	200
Panel A: MAC bandwidth = $\lfloor m^{3/5} \rfloor$																		
$\vartheta_i = 0.4$:																		
50	0.647	0.589	0.528	0.726	0.679	0.615	0.645	0.588	0.527	0.725	0.677	0.613	0.644	0.587	0.526	0.724	0.677	0.613
100	0.704	0.661	0.604	0.776	0.746	0.698	0.702	0.660	0.603	0.775	0.746	0.697	0.702	0.660	0.602	0.774	0.746	0.696
200	0.736	0.716	0.662	0.791	0.784	0.749	0.735	0.715	0.660	0.790	0.782	0.748	0.735	0.714	0.659	0.790	0.782	0.748
400	0.737	0.736	0.711	0.809	0.797	0.779	0.736	0.735	0.709	0.808	0.796	0.778	0.736	0.735	0.709	0.808	0.795	0.778
$\vartheta_i = 0.7$:																		
50	0.659	0.621	0.564	0.671	0.619	0.553	0.656	0.619	0.563	0.669	0.617	0.552	0.656	0.619	0.562	0.668	0.617	0.551
100	0.710	0.688	0.644	0.724	0.685	0.629	0.708	0.687	0.642	0.723	0.684	0.628	0.708	0.687	0.642	0.722	0.683	0.628
200	0.741	0.730	0.695	0.750	0.735	0.696	0.740	0.729	0.693	0.749	0.734	0.695	0.740	0.729	0.693	0.748	0.734	0.695
400	0.750	0.756	0.736	0.759	0.757	0.736	0.749	0.755	0.735	0.758	0.756	0.734	0.749	0.754	0.734	0.758	0.756	0.734
$\vartheta_i = 1$:																		
50	0.710	0.711	0.702	0.641	0.586	0.514	0.708	0.709	0.701	0.639	0.583	0.512	0.708	0.709	0.700	0.638	0.582	0.511
100	0.748	0.753	0.734	0.699	0.661	0.603	0.747	0.752	0.733	0.697	0.660	0.602	0.747	0.751	0.733	0.697	0.660	0.602
200	0.774	0.786	0.759	0.727	0.704	0.665	0.774	0.785	0.759	0.725	0.703	0.664	0.773	0.785	0.759	0.725	0.703	0.663
400	0.791	0.798	0.782	0.739	0.731	0.709	0.791	0.798	0.781	0.738	0.730	0.707	0.790	0.798	0.781	0.738	0.729	0.707
Panel B: MAC bandwidth = $\lfloor m^{4/5} \rfloor$																		
$\vartheta_i = 0.4$:																		
50	0.647	0.589	0.528	0.726	0.679	0.615	0.645	0.588	0.526	0.725	0.677	0.613	0.644	0.587	0.526	0.724	0.677	0.613
100	0.703	0.661	0.604	0.776	0.746	0.698	0.702	0.660	0.603	0.775	0.746	0.697	0.702	0.660	0.602	0.774	0.746	0.696
200	0.736	0.716	0.662	0.791	0.784	0.749	0.735	0.715	0.660	0.790	0.782	0.748	0.735	0.714	0.659	0.790	0.782	0.748
400	0.737	0.736	0.710	0.809	0.797	0.779	0.736	0.735	0.709	0.808	0.796	0.779	0.736	0.735	0.709	0.808	0.795	0.778
$\vartheta_i = 0.7$:																		
50	0.659	0.621	0.565	0.671	0.619	0.553	0.656	0.619	0.563	0.669	0.617	0.552	0.656	0.619	0.562	0.668	0.617	0.551
100	0.709	0.688	0.644	0.724	0.685	0.629	0.708	0.687	0.642	0.723	0.684	0.628	0.708	0.687	0.642	0.722	0.683	0.628
200	0.741	0.730	0.695	0.750	0.735	0.696	0.740	0.729	0.693	0.749	0.734	0.695	0.740	0.729	0.693	0.748	0.734	0.695
400	0.750	0.756	0.736	0.759	0.757	0.735	0.749	0.755	0.735	0.758	0.756	0.734	0.749	0.754	0.734	0.758	0.756	0.734
$\vartheta_i = 1$:																		
50	0.710	0.711	0.702	0.640	0.585	0.514	0.708	0.709	0.701	0.639	0.583	0.512	0.708	0.709	0.700	0.638	0.582	0.511
100	0.748	0.753	0.734	0.698	0.661	0.603	0.747	0.752	0.733	0.697	0.660	0.602	0.747	0.751	0.733	0.697	0.660	0.602
200	0.774	0.786	0.759	0.727	0.704	0.665	0.773	0.785	0.759	0.725	0.703	0.664	0.773	0.785	0.759	0.725	0.703	0.663
400	0.791	0.798	0.782	0.739	0.730	0.709	0.791	0.798	0.781	0.738	0.730	0.707	0.790	0.798	0.781	0.738	0.729	0.707

Table 6: Power ($c_i = 2$) of encompassing style test, ϱ_{ih}

This table reports power for test at level 10%, 5%, and 1% of the ϱ_{ih} statistic with $n = 20$ ($\delta = 0.4, \theta_1 = \theta_2 = 0, \rho = 0$), various integration orders d_i and ϑ_i and length of in-sample (T in the column) and out-of-sample (m in the row) windows. Panel A (B) contains results for MAC estimator bandwidth $\lfloor m^{3/5} \rfloor$ ($\lfloor m^{4/5} \rfloor$). Results are based on 20,000 simulations.

	Significance level = 10%						*Significance level = 5%*						*Significance level = 1%*					
	$d_i = 0.4$			$d_i = 1$			$d_i = 0.4$			$d_i = 1$			$d_i = 0.4$			$d_i = 1$		
T\m	50	100	200	50	100	200	50	100	200	50	100	200	50	100	200	50	100	200
							Panel A: MAC bandwidth = $\lfloor m^{3/5} \rfloor$											
$\vartheta_i = 0.4:$																		
50	0.515	0.518	0.499	0.575	0.636	0.626	0.302	0.339	0.349	0.348	0.444	0.466	0.064	0.093	0.119	0.076	0.137	0.183
100	0.564	0.587	0.565	0.607	0.687	0.688	0.337	0.396	0.408	0.371	0.485	0.523	0.070	0.114	0.145	0.080	0.151	0.207
200	0.593	0.628	0.611	0.625	0.709	0.724	0.361	0.431	0.457	0.378	0.499	0.556	0.072	0.125	0.170	0.082	0.157	0.225
400	0.592	0.654	0.660	0.630	0.713	0.756	0.360	0.459	0.494	0.381	0.507	0.589	0.076	0.136	0.187	0.080	0.158	0.244
$\vartheta_i = 0.7:$																		
50	0.495	0.514	0.498	0.538	0.567	0.549	0.288	0.333	0.342	0.324	0.381	0.395	0.063	0.093	0.118	0.064	0.107	0.138
100	0.553	0.593	0.594	0.577	0.624	0.609	0.331	0.397	0.428	0.349	0.430	0.449	0.068	0.115	0.152	0.071	0.124	0.165
200	0.588	0.644	0.651	0.595	0.659	0.660	0.352	0.438	0.482	0.358	0.461	0.493	0.073	0.131	0.182	0.075	0.140	0.186
400	0.602	0.675	0.690	0.610	0.680	0.693	0.367	0.470	0.519	0.365	0.478	0.530	0.079	0.139	0.196	0.074	0.145	0.203
$\vartheta_i = 1:$																		
50	0.478	0.537	0.586	0.515	0.520	0.490	0.277	0.352	0.423	0.302	0.344	0.343	0.061	0.106	0.160	0.062	0.094	0.113
100	0.550	0.645	0.709	0.563	0.594	0.571	0.329	0.443	0.532	0.334	0.404	0.414	0.075	0.136	0.217	0.069	0.113	0.146
200	0.597	0.701	0.783	0.585	0.624	0.622	0.358	0.493	0.609	0.356	0.428	0.458	0.080	0.163	0.255	0.073	0.124	0.165
400	0.609	0.730	0.824	0.591	0.651	0.663	0.376	0.517	0.653	0.356	0.452	0.496	0.088	0.170	0.290	0.077	0.133	0.186
							Panel B: MAC bandwidth = $\lfloor m^{4/5} \rfloor$											
$\vartheta_i = 0.4:$																		
50	0.482	0.502	0.489	0.538	0.616	0.616	0.214	0.289	0.325	0.250	0.384	0.433	0.004	0.030	0.074	0.007	0.049	0.118
100	0.532	0.570	0.558	0.571	0.671	0.680	0.240	0.343	0.380	0.264	0.421	0.490	0.005	0.039	0.091	0.007	0.053	0.137
200	0.561	0.611	0.606	0.591	0.693	0.714	0.257	0.375	0.430	0.272	0.436	0.522	0.005	0.042	0.105	0.006	0.057	0.150
400	0.559	0.643	0.656	0.590	0.694	0.748	0.259	0.404	0.466	0.267	0.441	0.559	0.005	0.048	0.119	0.006	0.057	0.162
$\vartheta_i = 0.7:$																		
50	0.462	0.494	0.487	0.506	0.550	0.538	0.207	0.288	0.318	0.231	0.324	0.365	0.004	0.031	0.070	0.005	0.036	0.086
100	0.524	0.578	0.585	0.544	0.608	0.602	0.237	0.342	0.397	0.247	0.370	0.422	0.006	0.040	0.098	0.005	0.044	0.105
200	0.558	0.628	0.644	0.565	0.644	0.652	0.252	0.382	0.453	0.254	0.402	0.463	0.006	0.046	0.119	0.006	0.047	0.120
400	0.571	0.661	0.680	0.577	0.662	0.686	0.265	0.413	0.490	0.258	0.417	0.498	0.005	0.049	0.126	0.005	0.049	0.130
$\vartheta_i = 1:$																		
50	0.441	0.511	0.573	0.480	0.502	0.481	0.190	0.300	0.390	0.214	0.292	0.315	0.006	0.041	0.106	0.004	0.029	0.067
100	0.510	0.622	0.696	0.531	0.578	0.563	0.232	0.377	0.494	0.238	0.348	0.386	0.007	0.052	0.145	0.005	0.037	0.089
200	0.558	0.676	0.772	0.556	0.611	0.613	0.251	0.425	0.575	0.256	0.374	0.430	0.007	0.064	0.174	0.006	0.040	0.104
400	0.571	0.708	0.814	0.559	0.637	0.655	0.271	0.445	0.618	0.260	0.394	0.466	0.008	0.068	0.199	0.005	0.046	0.119

Table 7: Power ($c_i = 5$) of Diebold-Mariano style test, $2DM_{ih}$

This table reports power for test at level 10%, 5%, and 1% of the $2DM_{ih}$ statistic with $n = 20$ ($\delta = 0.4$, $\theta_1 = \theta_2 = 0, \rho = 0$), various integration orders d_i and ϑ_i and length of in-sample (T in the column) and out-of-sample (m in the row) windows. Panel A (B) contains results for MAC estimator bandwidth $\lfloor m^{3/5} \rfloor$ ($\lfloor m^{4/5} \rfloor$). Results are based on 20,000 simulations.

	Significance level = 10%						*Significance level = 5%*						*Significance level = 1%*					
	$d_i = 0.4$			$d_i = 1$			$d_i = 0.4$			$d_i = 1$			$d_i = 0.4$			$d_i = 1$		
T\m	50	100	200	50	100	200	50	100	200	50	100	200	50	100	200	50	100	200
							Panel A: MAC bandwidth = $\lfloor m^{3/5} \rfloor$											
$\vartheta_i = 0.4$:																		
50	0.979	0.975	0.965	0.989	0.985	0.978	0.978	0.973	0.963	0.989	0.984	0.977	0.977	0.972	0.962	0.988	0.984	0.976
100	0.986	0.987	0.984	0.995	0.995	0.992	0.985	0.986	0.983	0.994	0.995	0.991	0.984	0.986	0.982	0.994	0.994	0.991
200	0.989	0.989	0.990	0.996	0.996	0.996	0.989	0.989	0.989	0.996	0.996	0.996	0.988	0.988	0.988	0.995	0.996	0.996
400	0.990	0.990	0.991	0.996	0.997	0.997	0.989	0.989	0.990	0.996	0.997	0.997	0.988	0.988	0.990	0.996	0.997	0.997
$\vartheta_i = 0.7$:																		
50	0.981	0.982	0.982	0.982	0.976	0.969	0.980	0.980	0.980	0.981	0.975	0.967	0.979	0.979	0.978	0.980	0.975	0.965
100	0.987	0.989	0.991	0.989	0.989	0.986	0.986	0.988	0.990	0.989	0.988	0.985	0.985	0.987	0.989	0.988	0.988	0.984
200	0.990	0.991	0.993	0.992	0.992	0.993	0.990	0.991	0.992	0.992	0.992	0.992	0.989	0.990	0.992	0.991	0.991	0.991
400	0.989	0.993	0.993	0.992	0.993	0.994	0.988	0.992	0.993	0.991	0.993	0.993	0.987	0.992	0.993	0.991	0.992	0.993
$\vartheta_i = 1$:																		
50	0.985	0.988	0.988	0.977	0.973	0.964	0.984	0.987	0.988	0.975	0.971	0.962	0.982	0.986	0.987	0.975	0.969	0.960
100	0.991	0.994	0.992	0.985	0.987	0.984	0.991	0.993	0.992	0.984	0.986	0.983	0.990	0.993	0.991	0.983	0.985	0.982
200	0.994	0.994	0.994	0.989	0.989	0.990	0.994	0.994	0.994	0.988	0.988	0.989	0.993	0.993	0.993	0.988	0.988	0.989
400	0.994	0.996	0.995	0.988	0.990	0.991	0.994	0.996	0.994	0.987	0.989	0.991	0.994	0.995	0.994	0.986	0.988	0.990
							Panel B: MAC bandwidth = $\lfloor m^{4/5} \rfloor$											
$\vartheta_i = 0.4$:																		
50	0.979	0.974	0.965	0.989	0.985	0.979	0.978	0.973	0.963	0.989	0.984	0.977	0.977	0.972	0.962	0.988	0.984	0.976
100	0.986	0.987	0.984	0.995	0.995	0.992	0.985	0.986	0.983	0.994	0.995	0.991	0.984	0.986	0.982	0.994	0.994	0.991
200	0.989	0.989	0.990	0.996	0.996	0.996	0.989	0.989	0.989	0.996	0.996	0.996	0.988	0.988	0.988	0.995	0.996	0.996
400	0.990	0.989	0.991	0.996	0.997	0.997	0.989	0.989	0.990	0.996	0.997	0.997	0.988	0.988	0.990	0.996	0.997	0.997
$\vartheta_i = 0.7$:																		
50	0.981	0.982	0.982	0.982	0.977	0.969	0.980	0.980	0.980	0.981	0.975	0.967	0.979	0.979	0.978	0.980	0.975	0.965
100	0.987	0.989	0.991	0.989	0.989	0.986	0.986	0.988	0.990	0.988	0.988	0.985	0.985	0.987	0.989	0.988	0.988	0.984
200	0.990	0.991	0.993	0.992	0.992	0.993	0.990	0.991	0.992	0.992	0.992	0.992	0.989	0.990	0.992	0.991	0.991	0.991
400	0.989	0.993	0.994	0.992	0.993	0.994	0.988	0.992	0.993	0.991	0.993	0.993	0.987	0.992	0.993	0.991	0.992	0.993
$\vartheta_i = 1$:																		
50	0.985	0.988	0.988	0.977	0.973	0.964	0.983	0.987	0.988	0.975	0.971	0.962	0.982	0.986	0.987	0.975	0.969	0.960
100	0.991	0.994	0.992	0.985	0.987	0.984	0.991	0.993	0.992	0.984	0.986	0.983	0.990	0.993	0.991	0.983	0.985	0.982
200	0.994	0.994	0.994	0.989	0.989	0.990	0.994	0.994	0.994	0.988	0.988	0.989	0.993	0.993	0.993	0.988	0.988	0.989
400	0.994	0.996	0.995	0.988	0.990	0.991	0.994	0.996	0.994	0.987	0.989	0.991	0.994	0.995	0.994	0.986	0.988	0.990

Table 8: Power ($c_i = 5$) of Diebold-Mariano style test, $\widetilde{DM}_{ih}$

This table reports power for test at level 10%, 5%, and 1% of the $\widetilde{DM}_{ih}$ statistic with $n = 20$ ($\delta = 0.4$, $\theta_1 = \theta_2 = 0, \rho = 0$), various integration orders d_i and ϑ_i and length of in-sample (T in the column) and out-of-sample (m in the row) windows. Panel A (B) contains results for MAC estimator bandwidth $\lfloor m^{3/5} \rfloor$ ($\lfloor m^{4/5} \rfloor$). Results are based on 20,000 simulations.

	Significance level = 10%						*Significance level = 5%*						*Significance level = 1%*					
	$d_i = 0.4$			$d_i = 1$			$d_i = 0.4$			$d_i = 1$			$d_i = 0.4$			$d_i = 1$		
T\m	50	100	200	50	100	200	50	100	200	50	100	200	50	100	200	50	100	200
Panel A: MAC bandwidth = $\lfloor m^{3/5} \rfloor$																		
$\vartheta_i = 0.4$:																		
50	0.976	0.971	0.961	0.988	0.984	0.976	0.976	0.971	0.961	0.988	0.984	0.976	0.976	0.971	0.961	0.988	0.983	0.976
100	0.984	0.985	0.982	0.994	0.994	0.991	0.983	0.985	0.981	0.994	0.994	0.991	0.983	0.985	0.981	0.994	0.994	0.991
200	0.988	0.988	0.988	0.995	0.996	0.996	0.988	0.988	0.988	0.995	0.996	0.996	0.988	0.988	0.988	0.995	0.996	0.996
400	0.988	0.988	0.989	0.996	0.997	0.997	0.988	0.988	0.989	0.996	0.997	0.997	0.987	0.988	0.989	0.996	0.997	0.997
$\vartheta_i = 0.7$:																		
50	0.979	0.978	0.978	0.980	0.974	0.966	0.979	0.978	0.978	0.980	0.974	0.965	0.979	0.978	0.978	0.980	0.974	0.965
100	0.985	0.987	0.989	0.988	0.988	0.984	0.985	0.987	0.989	0.988	0.988	0.984	0.985	0.987	0.989	0.988	0.988	0.984
200	0.989	0.990	0.992	0.991	0.991	0.991	0.989	0.990	0.992	0.991	0.991	0.991	0.989	0.990	0.992	0.991	0.991	0.991
400	0.987	0.992	0.993	0.991	0.992	0.993	0.987	0.992	0.993	0.991	0.992	0.993	0.987	0.992	0.993	0.991	0.992	0.993
$\vartheta_i = 1$:																		
50	0.982	0.986	0.987	0.975	0.969	0.960	0.982	0.986	0.987	0.974	0.969	0.960	0.982	0.986	0.987	0.974	0.969	0.960
100	0.990	0.993	0.991	0.983	0.985	0.981	0.990	0.993	0.991	0.983	0.985	0.981	0.990	0.993	0.991	0.983	0.985	0.981
200	0.993	0.993	0.993	0.988	0.988	0.988	0.993	0.993	0.993	0.988	0.988	0.988	0.993	0.993	0.993	0.988	0.988	0.988
400	0.994	0.995	0.994	0.986	0.988	0.990	0.994	0.995	0.994	0.986	0.988	0.990	0.994	0.995	0.994	0.986	0.988	0.990
Panel B: MAC bandwidth = $\lfloor m^{4/5} \rfloor$																		
$\vartheta_i = 0.4$:																		
50	0.976	0.971	0.961	0.988	0.984	0.976	0.976	0.971	0.961	0.988	0.984	0.976	0.976	0.971	0.961	0.988	0.983	0.976
100	0.984	0.985	0.982	0.994	0.994	0.991	0.983	0.985	0.981	0.994	0.994	0.991	0.983	0.985	0.981	0.994	0.994	0.991
200	0.988	0.988	0.988	0.995	0.996	0.996	0.988	0.988	0.988	0.995	0.996	0.996	0.988	0.988	0.988	0.995	0.996	0.996
400	0.988	0.988	0.989	0.996	0.997	0.997	0.988	0.988	0.989	0.996	0.997	0.997	0.987	0.988	0.989	0.996	0.997	0.997
$\vartheta_i = 0.7$:																		
50	0.979	0.978	0.978	0.980	0.974	0.965	0.979	0.978	0.978	0.980	0.974	0.965	0.979	0.978	0.978	0.980	0.974	0.965
100	0.985	0.987	0.989	0.988	0.988	0.984	0.985	0.987	0.989	0.988	0.988	0.984	0.985	0.987	0.989	0.988	0.988	0.984
200	0.989	0.990	0.992	0.991	0.991	0.991	0.989	0.990	0.992	0.991	0.991	0.991	0.989	0.990	0.992	0.991	0.991	0.991
400	0.987	0.992	0.993	0.991	0.992	0.993	0.987	0.992	0.993	0.991	0.992	0.993	0.987	0.992	0.993	0.991	0.992	0.993
$\vartheta_i = 1$:																		
50	0.982	0.986	0.987	0.974	0.969	0.960	0.982	0.986	0.987	0.974	0.969	0.960	0.982	0.986	0.987	0.974	0.969	0.960
100	0.990	0.993	0.991	0.983	0.985	0.981	0.990	0.993	0.991	0.983	0.985	0.981	0.990	0.993	0.991	0.983	0.985	0.981
200	0.993	0.993	0.993	0.988	0.988	0.988	0.993	0.993	0.993	0.988	0.988	0.988	0.993	0.993	0.993	0.988	0.988	0.988
400	0.994	0.995	0.994	0.986	0.988	0.990	0.994	0.995	0.994	0.986	0.988	0.990	0.994	0.995	0.994	0.986	0.988	0.990

Table 9: Power ($c_i = 5$) of encompassing style test, ϱ_{ih}

This table reports power for test at level 10%, 5%, and 1% of the ϱ_{ih} statistic with $n = 20$ ($\delta = 0.4$, $\theta_1 = \theta_2 = 0, \rho = 0$), various integration orders d_i and ϑ_i and length of in-sample (T in the column) and out-of-sample (m in the row) windows. Panel A (B) contains results for MAC estimator bandwidth $\lfloor m^{3/5} \rfloor$ ($\lfloor m^{4/5} \rfloor$). Results are based on 20,000 simulations.

	Significance level = 10%						*Significance level = 5%*						*Significance level = 1%*					
	$d_i = 0.4$			$d_i = 1$			$d_i = 0.4$			$d_i = 1$			$d_i = 0.4$			$d_i = 1$		
T\m	50	100	200	50	100	200	50	100	200	50	100	200	50	100	200	50	100	200
Panel A: MAC bandwidth = $\lfloor m^{3/5} \rfloor$																		
$\vartheta_i = 0.4$:																		
50	0.874	0.961	0.979	0.801	0.956	0.993	0.679	0.874	0.938	0.575	0.852	0.967	0.209	0.480	0.706	0.159	0.450	0.758
100	0.893	0.977	0.992	0.801	0.959	0.996	0.700	0.902	0.967	0.574	0.860	0.975	0.220	0.518	0.767	0.149	0.451	0.782
200	0.899	0.981	0.996	0.806	0.960	0.997	0.717	0.912	0.978	0.577	0.860	0.979	0.230	0.537	0.789	0.153	0.452	0.787
400	0.901	0.985	0.997	0.799	0.964	0.997	0.719	0.920	0.982	0.567	0.866	0.979	0.231	0.547	0.805	0.147	0.456	0.796
$\vartheta_i = 0.7$:																		
50	0.785	0.882	0.917	0.860	0.970	0.992	0.573	0.758	0.846	0.652	0.882	0.967	0.167	0.379	0.588	0.198	0.489	0.753
100	0.829	0.932	0.971	0.865	0.976	0.996	0.620	0.820	0.920	0.664	0.898	0.978	0.180	0.421	0.670	0.195	0.510	0.786
200	0.847	0.951	0.989	0.866	0.975	0.997	0.636	0.846	0.953	0.665	0.900	0.982	0.185	0.450	0.717	0.196	0.517	0.798
400	0.849	0.958	0.993	0.871	0.977	0.997	0.643	0.861	0.965	0.666	0.905	0.982	0.186	0.464	0.744	0.199	0.519	0.800
$\vartheta_i = 1$:																		
50	0.518	0.592	0.681	0.874	0.963	0.978	0.333	0.445	0.558	0.679	0.874	0.941	0.083	0.183	0.320	0.207	0.480	0.701
100	0.579	0.675	0.783	0.894	0.978	0.994	0.373	0.502	0.650	0.701	0.908	0.972	0.089	0.193	0.365	0.223	0.521	0.771
200	0.601	0.719	0.855	0.897	0.981	0.997	0.388	0.543	0.723	0.711	0.914	0.979	0.089	0.208	0.391	0.226	0.536	0.794
400	0.618	0.751	0.887	0.898	0.983	0.998	0.398	0.566	0.765	0.716	0.916	0.982	0.093	0.220	0.420	0.228	0.548	0.803
Panel B: MAC bandwidth = $\lfloor m^{4/5} \rfloor$																		
$\vartheta_i = 0.4$:																		
50	0.840	0.953	0.977	0.751	0.942	0.990	0.506	0.820	0.926	0.401	0.778	0.952	0.011	0.205	0.559	0.008	0.174	0.603
100	0.860	0.971	0.991	0.751	0.948	0.995	0.532	0.857	0.959	0.388	0.789	0.962	0.011	0.222	0.626	0.006	0.175	0.622
200	0.868	0.976	0.996	0.756	0.949	0.995	0.548	0.865	0.972	0.391	0.789	0.968	0.011	0.236	0.653	0.006	0.174	0.630
400	0.871	0.980	0.996	0.750	0.951	0.996	0.551	0.879	0.974	0.389	0.796	0.968	0.011	0.241	0.671	0.006	0.173	0.635
$\vartheta_i = 0.7$:																		
50	0.739	0.868	0.911	0.819	0.961	0.991	0.411	0.693	0.825	0.482	0.823	0.954	0.008	0.156	0.455	0.009	0.202	0.605
100	0.786	0.921	0.968	0.828	0.969	0.995	0.450	0.758	0.902	0.488	0.845	0.967	0.009	0.174	0.530	0.010	0.214	0.644
200	0.805	0.941	0.988	0.832	0.968	0.997	0.464	0.785	0.938	0.488	0.846	0.974	0.008	0.187	0.572	0.009	0.219	0.657
400	0.815	0.950	0.991	0.833	0.971	0.996	0.474	0.801	0.954	0.489	0.857	0.974	0.010	0.192	0.594	0.009	0.220	0.666
$\vartheta_i = 1$:																		
50	0.461	0.563	0.663	0.840	0.955	0.977	0.215	0.378	0.520	0.506	0.818	0.928	0.004	0.069	0.237	0.010	0.204	0.552
100	0.519	0.636	0.765	0.860	0.972	0.993	0.241	0.423	0.610	0.531	0.862	0.964	0.004	0.071	0.263	0.012	0.227	0.631
200	0.538	0.681	0.839	0.865	0.976	0.996	0.249	0.454	0.682	0.541	0.866	0.974	0.004	0.073	0.274	0.011	0.233	0.658
400	0.553	0.712	0.872	0.868	0.977	0.997	0.254	0.478	0.720	0.547	0.874	0.975	0.004	0.077	0.298	0.011	0.241	0.670

Table 10: Size and power levels of aggregate panel encompassing style test, $\bar{\varrho}_h$

This table reports size and power for test at level 5% of the aggregate $\bar{\varrho}_h$ statistic with $n = 20$ ($\delta = 0.4$, $\theta_1 = \theta_2 = 0, \rho = 0$), various integration orders d_i and ϑ_i and length of in-sample (T in the column) and out-of-sample (m in the row) windows. Panel A (B) contains results for MAC estimator bandwidth $\lfloor m^{3/5} \rfloor$ ($\lfloor m^{4/5} \rfloor$). Results are based on 1,000 simulations.

	Size, significance level = 5%						*Power ($c_i = 2$), significance level = 5%*						*Power ($c_i = 5$), significance level = 5%*					
	$d_i = 0.4$			$d_i = 1$			$d_i = 0.4$			$d_i = 1$			$d_i = 0.4$			$d_i = 1$		
T\m	50	100	200	50	100	200	50	100	200	50	100	200	50	100	200	50	100	200
							Panel A: MAC bandwidth = $\lfloor m^{3/5} \rfloor$											
$\vartheta_i = 0.4$:																		
50	0.003	0.003	0.003	0.002	0.002	0.002	1.000	0.998	0.999	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
100	0.007	0.005	0.016	0.010	0.002	0.004	1.000	1.000	0.998	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
200	0.011	0.014	0.023	0.007	0.006	0.010	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
400	0.017	0.026	0.039	0.015	0.009	0.018	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
$\vartheta_i = 0.7$:																		
50	0.001	0.000	0.006	0.003	0.002	0.002	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
100	0.009	0.003	0.006	0.007	0.004	0.007	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
200	0.005	0.008	0.013	0.013	0.007	0.007	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
400	0.016	0.018	0.014	0.012	0.013	0.016	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
$\vartheta_i = 1$:																		
50	0.001	0.001	0.009	0.004	0.002	0.004	1.000	1.000	1.000	1.000	1.000	0.998	1.000	1.000	1.000	1.000	1.000	1.000
100	0.001	0.001	0.008	0.007	0.004	0.006	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
200	0.004	0.005	0.007	0.010	0.005	0.011	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
400	0.004	0.006	0.009	0.020	0.012	0.019	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
							Panel B: MAC bandwidth = $\lfloor m^{4/5} \rfloor$											
$\vartheta_i = 0.4$:																		
50	0.002	0.005	0.003	0.001	0.002	0.002	1.000	0.998	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
100	0.004	0.005	0.013	0.006	0.001	0.004	1.000	1.000	0.998	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
200	0.010	0.010	0.020	0.005	0.005	0.008	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
400	0.015	0.023	0.037	0.012	0.010	0.015	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
$\vartheta_i = 0.7$:																		
50	0.000	0.000	0.004	0.001	0.002	0.001	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
100	0.007	0.003	0.005	0.005	0.004	0.006	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
200	0.003	0.005	0.012	0.008	0.005	0.007	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
400	0.013	0.012	0.012	0.010	0.010	0.015	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
$\vartheta_i = 1$:																		
50	0.001	0.001	0.008	0.003	0.001	0.002	1.000	1.000	1.000	1.000	0.999	0.996	1.000	1.000	1.000	1.000	1.000	1.000
100	0.001	0.001	0.008	0.006	0.002	0.005	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
200	0.004	0.004	0.006	0.009	0.004	0.009	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
400	0.003	0.004	0.008	0.014	0.007	0.014	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000

B. Empirical results

Table 11: Parametric CSS estimates of the integration orders
This table reports the full-sample parametric conditional-sum-of-squares (CSS) estimation results for the integration orders of the indicators across countries. RV and EPU stand for realized volatility and economic policy uncertainty, respectively. The standard error of these estimates is 0.055. Superscripts ***, **, and * correspond to statistical significance at the 1%, 5%, and 10% levels, respectively, for a one-tailed hypothesis test against positive alternatives.

Country	RV	EPU
USA	0.686^{***}	0.651^{***}
Australia	0.612^{***}	0.615^{***}
Brazil	0.610^{***}	0.568^{***}
Canada	0.641^{***}	0.665^{***}
Germany	0.497^{***}	0.570^{***}
UK	0.628^{***}	0.642^{***}
France	0.595^{***}	0.577^{***}
Spain	0.594^{***}	0.587^{***}
Greece	0.470^{***}	0.668^{***}
Hong Kong	0.631^{***}	0.526^{***}
Ireland	0.590^{***}	0.486^{***}
Japan	0.510^{***}	0.721^{***}
Mexico	0.618^{***}	0.680^{***}
Sweden	0.664^{***}	0.690^{***}

Table 12: Estimates in contemporaneous model

This table reports estimates of the slope parameter, β_{i0}, and memory orders of the errors, d_{i0} and ϑ_{i0}, across countries. Estimations are performed by CSS based on (34) over the full sample, covering 2001-2017. Projections are carried out with $d^* = 1$. The last column (PV) reports estimates of the slope parameter from a regression similar to Pástor and Veronesi (2013), including a first-order autoregressive term. Robust standard errors are reported in parenthesis. Superscripts ***, **, and * correspond to statistical significance at the 1%, 5%, and 10% levels, respectively, for a one-tailed hypothesis test against positive alternatives.

Country	$\hat{\beta}_i$	$\hat{\vartheta}_i$	$\hat{d}_i$	PV
USA	0.028 (0.025)	0.489 (0.464)	0.372 (0.326)	0.216* (0.144)
Australia	0.043** (0.025)	0.483 (0.518)	0.521* (0.372)	0.239** (0.131)
Brazil	0.051 (0.040)	0.467 (0.638)	0.600 (0.592)	−0.047 (0.066)
Canada	0.097*** (0.025)	0.494 (0.427)	0.601** (0.355)	0.057 (0.177)
Germany	−0.125 (0.051)	0.366 (0.516)	0.545 (0.755)	−0.136 (0.063)
UK	0.015 (0.019)	0.554* (0.427)	0.425* (0.280)	−0.039 (0.075)
France	−0.076 (0.031)	0.491 (0.462)	0.383 (0.442)	0.055 (0.102)
Spain	−0.034 (0.034)	0.424 (0.613)	0.448 (0.466)	0.258*** (0.100)
Greece	0.118*** (0.042)	0.489 (0.537)	0.559 (0.624)	0.220** (0.104)
Hong Kong	−0.011 (0.036)	0.395 (0.716)	0.446 (0.498)	0.097 (0.093)
Ireland	0.070** (0.030)	0.319 (0.794)	0.383 (0.459)	0.017 (0.082)
Japan	0.090** (0.042)	0.595 (0.539)	0.455 (0.613)	0.320** (0.175)
Mexico	0.059** (0.031)	0.589 (0.526)	0.447 (0.448)	0.139 (0.139)
Sweden	0.022 (0.024)	0.434 (0.612)	0.505* (0.352)	0.138* (0.102)

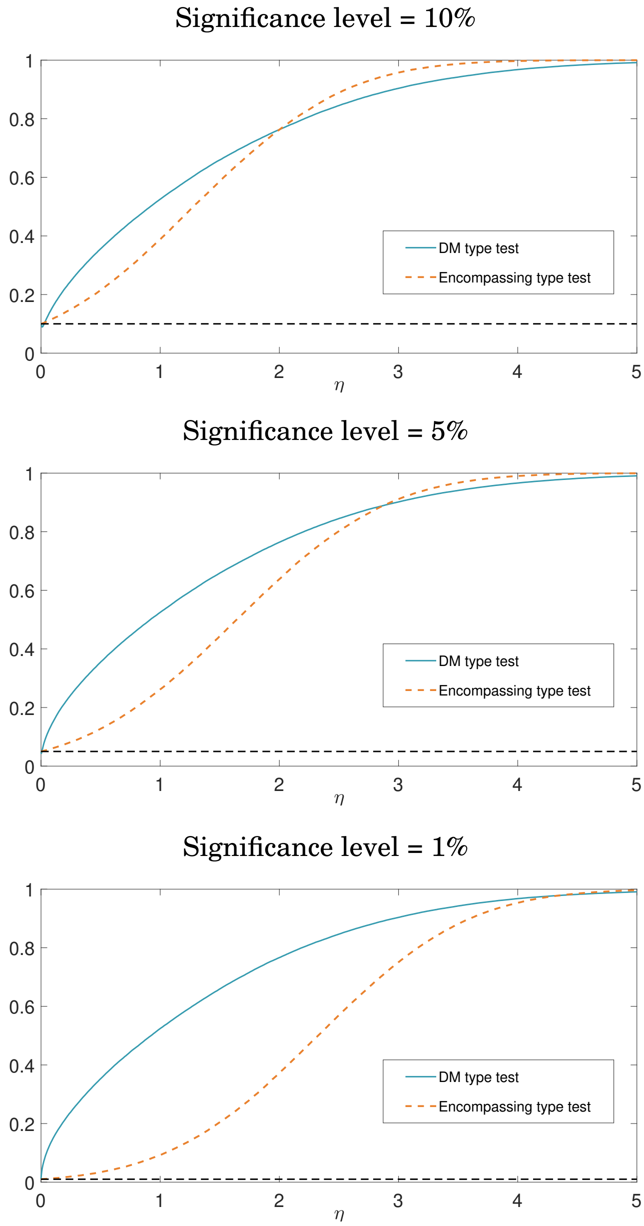
Table 13: Test of uninformativeness of economic policy uncertainty

This table reports ϱ_{ih} and $\overline{\varrho}_h$ statistics based on an expanding window estimation scheme of (34) and direct forecasts at $h = 1,2,3,6,9,12$ monthly horizons. The initial in-sample window is the 2001-2007 period, with first forecast generated for January 2008. Panel A performs estimations and forecasts under the assumption of the presence of an unobserved common factor structure, where projections are carried out with $d^* = 1$. Panel B assumes no unobserved common factor structure. Superscripts ***, **, and * correspond to statistical significance at the 1%, 5%, and 10% levels, respectively.

Country	$h=1$	$h=2$	$h=3$	$h=6$	$h=9$	$h=12$
	Panel A: Allowing for a common factor structure					
USA	-2.213	-1.471	-1.147	-0.161	-1.605	-0.450
Australia	-1.493	-1.184	0.861	-0.442	0.874	2.124**
Brazil	-1.425	-0.244	0.700	0.702	-1.179	1.037
Canada	0.621	1.242	-0.075	0.144	0.412	0.050
Germany	-0.698	0.310	0.991	-0.754	-3.691	-0.506
UK	2.206**	-0.365	0.469	-2.230	-1.046	-0.374
France	-1.452	-0.290	-0.886	0.874	-1.413	0.395
Spain	-1.415	-2.248	0.087	1.894**	-0.773	0.245
Greece	0.353	0.595	2.011*	1.457	-0.431	-0.561
Hong Kong	0.048	-0.825	0.119	-1.072	-1.251	-0.013
Ireland	-0.066	-0.169	-1.865	-0.284	0.079	1.141
Japan	-1.920	0.054	-0.888	0.193	-0.633	-0.746
Mexico	-0.969	-0.933	1.105	1.335*	0.567	0.154
Sweden	0.139	0.379	-1.556	0.093	-1.187	0.993
$\overline{\varrho}_h$	-2.214	-1.376	-0.020	0.467	-3.014	0.932
	Panel B: Assuming no common factor structure					
USA	-1.356	-0.341	-0.397	-1.663	0.596	-1.239
Australia	0.548	-0.775	1.184	-1.991	-0.039	1.161
Brazil	0.186	0.067	0.780	1.518	-0.847	-1.562
Canada	0.827	1.422*	1.370*	2.121**	1.806**	1.198
Germany	-0.326	0.650	-1.487	1.063	-1.197	1.293*
UK	-0.068	0.134	0.638	0.266	0.271	1.430*
France	-1.012	-1.088	-0.958	-0.297	-1.415	0.400
Spain	0.664	-0.174	1.709**	1.822**	-0.596	1.544*
Greece	1.394*	-2.006	1.576*	0.054	0.096	0.023
Hong Kong	0.236	-0.739	0.271	-0.080	0.852	0.391
Ireland	-1.170	-0.750	-1.210	-0.955	-0.127	0.423
Japan	1.907**	-0.136	1.491*	-1.592	0.602	-0.371
Mexico	-1.167	-1.126	1.376*	-0.162	0.283	1.619*
Sweden	-0.563	0.779	-0.608	-1.002	0.416	-0.472
$\overline{\varrho}_h$	0.027	-1.091	1.533*	-0.240	0.187	1.559*

B. Figures

Figure 1: Local asymptotic power curves for $\widetilde{DM}_{i1}$ and ϱ_{ih}



This figure depicts the local asymptotic power curves for $\widetilde{DM}_{i1}$ and ϱ_{ih} based on Theorem 4. The significance level used in each figure (10% (upper figure), 5% (middle figure), and 1% (lower figure)) is indicated by the horizontal dashed black line.

Figure 2: Stock market volatility and political uncertainty

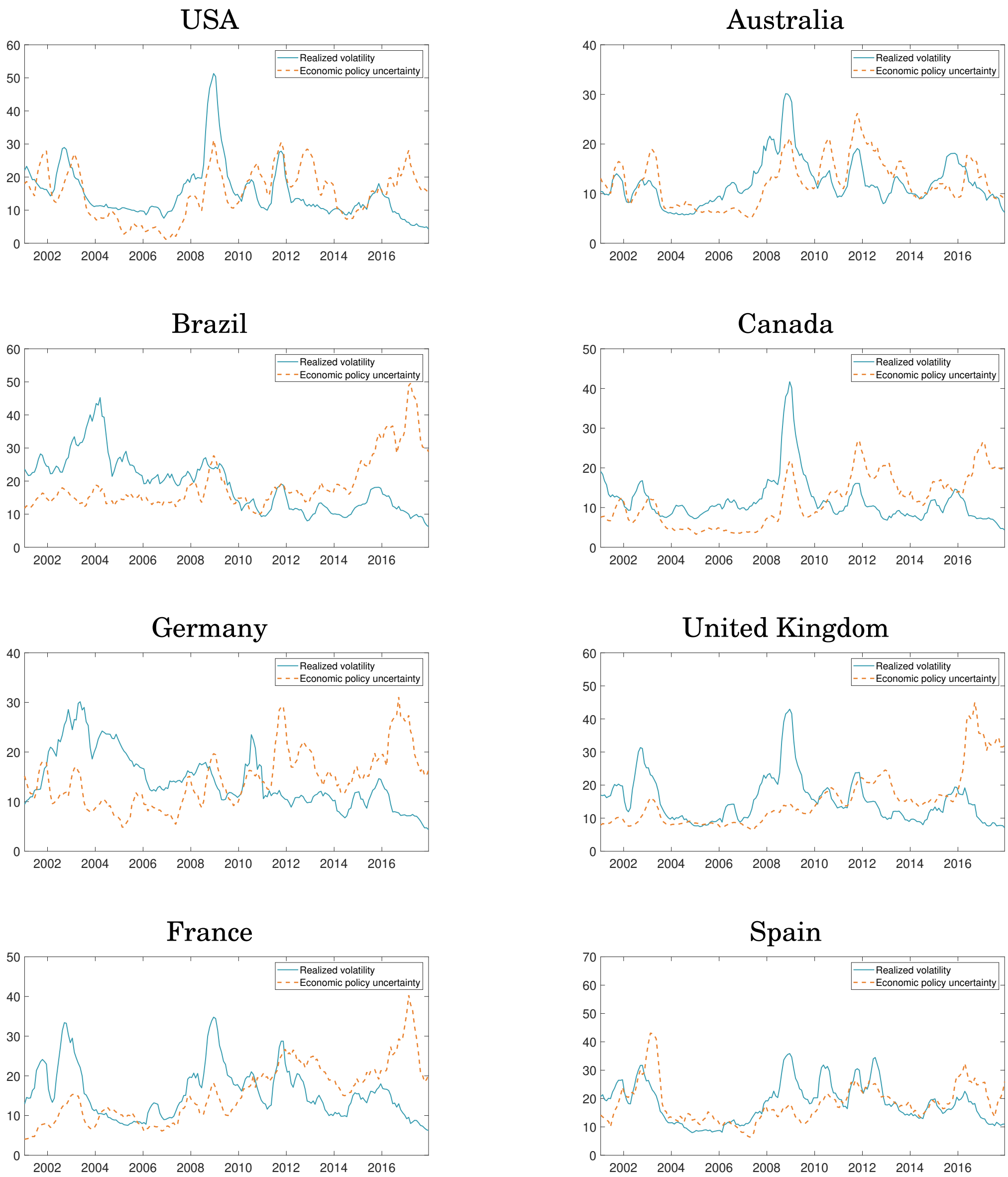
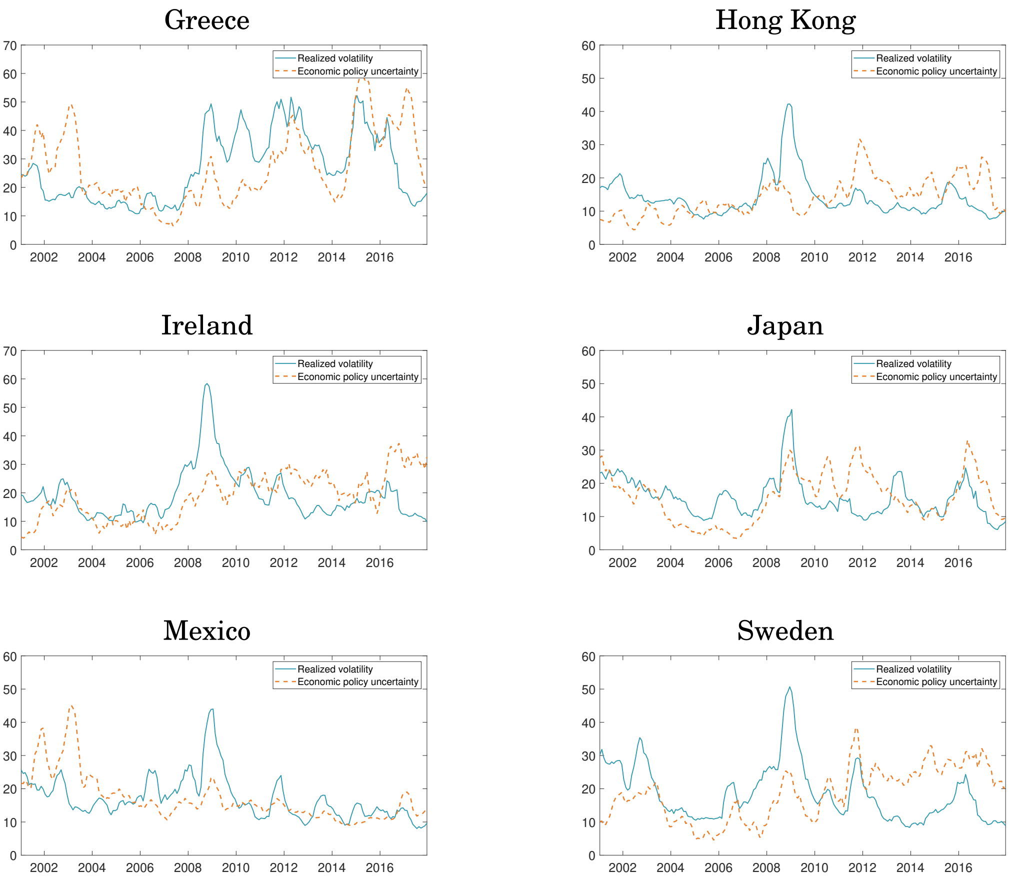


Figure 2 (Cont.): Stock market volatility and political uncertainty



This figure depicts international stock market volatility and economic political uncertainty. In each figure, the solid green line plots the realized volatility of each country's stock market index, and the dashed orange line plots the economic policy uncertainty (EPU) index of Baker et al. (2016). The EPU is scaled to the same mean and variance as realized volatility within each figure. Realized volatility is computed monthly from daily returns within the month. The figures depict each variable's six-month moving average between January 2001 and December 2017.

C. Proofs

We here prove the theoretical results presented in the main text, along with an auxiliary lemma.

Proof of Theorem 1. Given the parameter consistency arguments derived in Ergemen (2019), $\hat{\varphi}_{it}^h$ is a consistent estimator for φ_{it}^h. First, it can be easily verified that

$$Var\left(\frac{1}{m}\sum_{t=1}^{m}\varphi_{it}^h\right) \propto O(1+m^{2\kappa_i-1}),$$

see, e.g., the proofs of Theorems 1 and 2 of Ergemen and Velasco (2017) for a similar treatment, so

$$Var\left(\frac{1}{m^{\kappa_i-1/2}}\frac{1}{m}\sum_{t=1}^{m}\varphi_{it}^h\right) \propto O(1).$$

Under Assumptions A-B, Ergemen (2019) establishes that the projected series incurs a projection error of size $O_p(n^{-1}+(nT)^{-1/2})$. Given that the test statistic is based on the projected series and has a $m^{1/2-\kappa_i}$ convergence rate, we need to account for the projection error that becomes of size

$$O_p\left(m^{1/2-\kappa_i}n^{-1}+m^{1/2-\kappa_i}n^{-1/2}T^{-1/2}\right) = o_p(1)$$

if $m^{1/2-\kappa_i}n^{-1}+m^{1/2-\kappa_i}n^{-1/2}T^{-1/2} \to 0$ as $(m,n,T)_j \to \infty$.

Next, we analyze the asymptotic behavior of the test statistic under the null in (19). Define $D_{it+h}(\tau_i) = \partial\tilde{y}_{it+h|t}^*(\tau_i)/\partial\tau_i$. Applying the Mean Value Theorem,

$$\tilde{y}_{it+h|t}^*(\hat{\tau}_{it}) = \tilde{y}_{it+h|t}^*(\tau_i) + D_{it+h}(\bar{\tau}_{it})(\hat{\tau}_{it}-\tau_i)$$

for some $\hat{\tau}_{it} \le \bar{\tau}_{it} \le \tau_i$, so, for fixed i,

$$
\begin{aligned}
\varphi_{it}^h &= [u_{it+h}(\tau_i) + D_{it+h}(\bar{\tau}_{it})(\hat{\tau}_{it} - \tau_i)]^2 - (u_{it+h}(\tau_i) - \overline{u_{ih}(\tau_i)})^2 \\
&= (u_{it+h}(\tau_i))^2 - (u_{it+h}(\tau_i) - \overline{u_{ih}(\tau_i)})^2 \\
&\qquad + 2u_{it+h}(\tau_i) D_{it+h}(\bar{\tau}_{it})(\hat{\tau}_{it} - \tau_i) + D_{it+h}(\bar{\tau}_{it})^2 (\hat{\tau}_{it} - \tau_i)^2 \\
&= \overline{u_{ih}(\tau_i)}(2u_{it+h}(\tau_i) - \overline{u_{ih}(\tau_i)}) \\
&\qquad + 2u_{it+h}(\tau_i) D_{it+h}(\bar{\tau}_{it})(\hat{\tau}_{it} - \tau_i) + D_{it+h}(\bar{\tau}_{it})^2 (\hat{\tau}_{it} - \tau_i)^2
\end{aligned}
\tag{C.1}
$$

where $\overline{u_{ih}(\tau_i)} = m^{-1} \sum_{t=h+1}^{m+h} u_{it}(\tau_i)$ and

$$
\hat{\tau}_{it} - \tau_i = (\hat{\tau}_{it} - \hat{\tau}_{i0}) + (\hat{\tau}_{i0} - \tau_i) = O_p(T^{-1/2})
$$

under Assumption C.2. Then,

$$
\begin{aligned}
\frac{1}{m} \sum_{t=1}^{m} \varphi_{it}^h &= \overline{u_{ih}(\tau_i)}^2 + \frac{2}{m} \sum_{t=1}^{m} u_{it+h}(\tau_i) D_{it+h}(\bar{\tau}_{it})(\hat{\tau}_{it} - \tau_i) + \frac{1}{m} \sum_{t=1}^{m} D_{it+h}(\bar{\tau}_{it})^2 (\hat{\tau}_{it} - \tau_i)^2 \\
&= \overline{u_{ih}(\tau_i)}^2 + O_p(T^{-1/2}) + O_p(T^{-1}) \\
&= \overline{u_{ih}(\tau_i)}^2 + O_p(T^{-1/2})
\end{aligned}
$$

by Lemma A under Assumptions A.1 and C. Then using (C.1),

$$
\begin{aligned}
&\frac{1}{m^{\kappa_i - 1/2}} \frac{1}{m} \sum_{t=j+1}^{m} \varphi_{it}^h \varphi_{it-j}^h \\
&= \frac{1}{m^{\kappa_i - 1/2}} \left(\frac{1}{m} \overline{u_{ih}(\tau_i)}^2 \sum_{t=j+1}^{m} (2u_{it+h}(\tau_i) - \overline{u_{ih}(\tau_i)})(2u_{it+h-j}(\tau_i) - \overline{u_{ih}(\tau_i)}) + O_p(T^{-1/2}) \right) \\
&= \frac{1}{m^{\kappa_i - 1/2}} \left(\frac{1}{m} \overline{u_{ih}(\tau_i)}^2 \left[\left(\sum_{t=j+1}^{m} 4u_{it+h}(\tau_i) u_{it+h-j}(\tau_i) \right) - 3m \overline{u_{ih}(\tau_i)}^2 \right] + O_p(T^{-1/2}) \right) \\
&= \frac{1}{m^{\kappa_i - 1/2}} \left(4\overline{u_{ih}(\tau_i)}^2 \left(\frac{1}{m} \sum_{t=j+1}^{m} u_{it+h}(\tau_i) u_{it+h-j}(\tau_i) \right) + O_p(m^{-2}) + O_p(T^{-1/2}) \right) \\
&= 4\overline{u_{ih}(\tau_i)}^2 \left(\frac{1}{m^{\kappa_i - 1/2}} \frac{1}{m} \sum_{t=j+1}^{m} u_{it+h}(\tau_i) u_{it+h-j}(\tau_i) \right) + O_p(m^{-\kappa_i - 3/2}) + O_p(m^{1/2 - \kappa_i} T^{-1/2}),
\end{aligned}
$$

reasoning as before. Hence,

$$
\hat{\omega}_{\varphi}^2 = 4\overline{u_{ih}(\tau_i)}^2 \hat{\omega}_{iu}^2 + O_p(m^{1/2 - \kappa_i} T^{-1/2}).
$$

So the test statistic can be written as

$$DM_{ih} = m^{1/2-\hat{\kappa}_i} \frac{\overline{u_{ih}(\tau_i)}^2 + O_p(T^{-1/2})}{\sqrt{4\overline{u_{ih}(\tau_i)}^2 \hat{\omega}_{iu}^2 + O_p(m^{1/2-\kappa_i}T^{-1/2})}}$$
$$= \frac{(m^{1/2-\hat{\kappa}_i}\overline{u_{ih}(\tau_i)})^2 + O_p(m^{1-2\kappa_i}T^{-1/2})}{\sqrt{4(m^{1/2-\hat{\kappa}_i}\overline{u_{ih}(\tau_i)})^2\hat{\omega}_{iu}^2 + O_p(m^{3/2-3\kappa_i}T^{-1/2})}}$$
$$= m^{1/2-\hat{\kappa}_i}\frac{\left|\overline{u_{ih}(\tau_i)}\right|}{2\hat{\omega}_{iu}} + O_p\left(m^{3/2-3\kappa_i}T^{-1/2}\right),$$

since $\kappa_i < 1/2$. A $\sqrt{T}$-consistent estimate for κ_i can be deduced based on Propositions 2-4 of Kruse et al. (2018) since κ_i depends on τ_i, for which Ergemen (2019) establishes $\sqrt{T}$-consistency under Assumptions A-B. Furthermore, for $\kappa_i \in (-1/2, 1/2)$, $\hat{\omega}_{iu}$ is a consistent estimator of ω_{iu}, cf. Robinson (2005) and Abadir et al. (2009). Thus, if $m^{3/2-3\kappa_i}T^{-1/2} \to 0$, in addition to the previously imposed condition $m^{1/2-\kappa_i}n^{-1} + m^{1/2-\kappa_i}n^{-1/2}T^{-1/2} \to 0$ as $(n,T,m)_j \to \infty$ to control for the projection error, i.e., under Assumption D, we have that

$$m^{1/2-\hat{\kappa}_i}\frac{\left|\overline{u_{ih}(\tau_i)}\right|}{2\hat{\omega}_{iu}} \Rightarrow \frac{|w_i|}{2},$$

where w_i is a standard normally distributed random variable, applying the Functional Central Limit Theorem. □

Proof of Corollary 1. The first result follows directly from Theorem 1. For the second result, the χ_1^2 distribution is obtained directly by applying the definition of φ_{it}^h and based on the arguments in the proof of Theorem 1 with $\hat{\omega}_{iu}^2$ a consistent estimator for the case of u_{it+h}. □

Proof of Theorem 2. In showing the result, we again argue that $\hat{\Xi}_{it}^h$ is consistent for Ξ_{it}^h, given parameter consistency arguments in Ergemen (2019) and we follow Breitung and Knüppel (2018)'s steps, making the necessary adjustments for long memory properties. First, we have that

$$\sum_{t=1}^{m}(\tilde{y}^*_{it+h|t}(\tau_i) - \overline{\tilde{y}^*_{ih}(\tau_i)})(\tilde{y}^*_{it+h}(\tau_i) - \overline{\tilde{y}^*_{ih}(\tau_i)}) = \sum_{t=1}^{m}\tilde{y}^*_{it+h|t}(\tau_{it})(u_{it+h}(\tau_i) - \overline{u_{ih}(\tau_i)}).$$

The estimation error $\tilde{y}^*_{it+h|t}(\tau_i)$ is correlated with $\overline{u_{ih}(\tau_i)}$. In order to tackle this issue, we decompose the forecast, explicitly showing the dependence on the memory estimates given in Assumption C, into a component $\tilde{y}^*_{it+h|t}(\hat{\tau}_{i0})$ that is independent of $\{u_{i1+h},\ldots,u_{im+h}\}$ and a remaining term and establish that the latter is asymptotically negligible. Applying a first-order expansion,

$$\tilde{y}^*_{it+h|t}(\hat{\tau}_{it}) = \tilde{y}^*_{it+h|t}(\hat{\tau}_{i0}) + D_{it+h}(\bar{\tau}_{it})(\hat{\tau}_{it} - \hat{\tau}_{i0})$$

where $\hat{\tau}_{it} \leq \bar{\tau}_{it} \leq \hat{\tau}_{i0}$. By Assumption C, $\tilde{y}^*_{it+h|t}(\hat{\tau}_{i0})$ is uncorrelated with $\{u_{i1+h},\ldots,u_{im+h}\}$. We then use the decomposition

$$\sum_{t=1}^{m}[\tilde{y}^*_{it+h|t}(\hat{\tau}_{i0}) + D_{it+h}(\bar{\tau}_{it})(\hat{\tau}_{it} - \hat{\tau}_{i0})](u_{it+h}(\tau_i) - \overline{u_{ih}(\tau_i)}) = A + B_1 + B_2,$$

where

$$A = \sum_{t=1}^{m} \tilde{y}^*_{it+h|t}(\hat{\tau}_{i0})(u_{it+h}(\tau_i) - \overline{u_{ih}(\tau_i)}),$$
$$B_1 = \sum_{t=1}^{m} D_{it+h}(\bar{\tau}_{it})(\hat{\tau}_{it} - \hat{\tau}_{i0})u_{it+h}(\tau_i),$$
$$B_2 = \overline{u_{ih}(\tau_i)} \sum_{t=1}^{m} D_{it+h}(\bar{\tau}_{it})(\hat{\tau}_{it} - \hat{\tau}_{i0}).$$

Further, expanding A around τ_i with $\tilde{y}^*_{it+h|t}(\tau_i) = \overline{\tilde{y}^*_{it}(\tau_i)}$ yields

$$\begin{aligned} A &= (\hat{\tau}_{i0} - \tau_i)\sum_{t=1}^{m} D_{it+h}(\bar{\tau}_{i0})u_{it+h}(\tau_i) - (\hat{\tau}_{i0} - \tau_i)\overline{u_{ih}(\tau_i)} \sum_{t=1}^{m} D_{it+h}(\bar{\tau}_{i0}) \\ &= A_1 + A_2 \end{aligned}$$

with $\hat{\tau}_{i0} \leq \bar{\tau}_{i0} \leq \tau_i$. Since $\hat{\tau}_{i0}$ and $D_{it+h}(\bar{\tau}_{i0})$ are uncorrelated with $u_{it+h}(\tau_i)$, we have that $A_1 = O_p(T^{-1/2})O_p(m^{1/2})$ and $A_2 = O_p(T^{-1/2})O_p(m^{v_i - 1/2})O_p(m)$ by Assumption C.2, using arguments as in the proof of Lemma A and reasoning as in the proof of Theorem 1. Thus, $A = O_p(T^{-1/2}m^{v_i+1/2})$. Under the null in (27), $(\hat{\tau}_{it} - \hat{\tau}_{i0})$ and $D_{it+h}(\bar{\tau}_{it})$ are uncorrelated with $u_{it+h}(\tau_i)$. Then, by Assumptions C.2 and using similar arguments as in the proof of Lemma A, it follows that

$$\sum_{t=1}^{m}(\hat{\tau}_{it} - \hat{\tau}_{i0})^2 D_{it+h}(\bar{\tau}_{it})^2 u_{it+h}(\tau_i)^2 = O_p\left(\frac{m^{2v_i+2}}{T^2}\right).$$

Thus, $B_1 = O_p(m^{v_i+1}T^{-1})$. Since by Assumptions C.2 and Lemma A,

$$\sum_{t=1}^{m}(\hat{\tau}_{it}-\hat{\tau}_{i0})D_{it+h}(\bar{\tau}_{it}) = O_p\left(\frac{m^{3/2}}{T}\right),$$

$B_2 = O_p(m^{v_i-1/2})O_p(m^{3/2}T^{-1}) = O_p(m^{v_i+1}T^{-1})$. Therefore, we have

$$A + B_1 + B_2 = O_p(m^{v_i+1/2}T^{-1/2} + m^{v_i+1}T^{-1}),$$

and, thus,

$$T^{1/2}m^{-1/2-v_i}(A+B_1+B_2) = O_p(1) + O_p(m^{1/2}T^{-1/2}),$$

where the $O_p(1)$ term leading to the asymptotic normal distribution is

$$T^{1/2}m^{-1/2-v_i}A = \sqrt{T}(\hat{\tau}_{i0}-\tau_i)m^{1/2-v_i}\frac{1}{m}\sum_{t=1}^{m}D_{it+h}(\bar{\tau}_{i0})\left(u_{it+h}(\tau_i)-\overline{u_{ih}(\tau_i)}\right). \tag{C.2}$$

Next, we analyze

$$\sum_{t=1}^{m}(\tilde{y}^*_{it+h|t}(\hat{\tau}_i)-\overline{\tilde{y}^*_{ih}(\hat{\tau}_i)})^2(\tilde{y}^*_{it+h}(\tau_i)-\overline{\tilde{y}^*_{ih}(\tau_i)})^2$$
$$= \sum_{t=1}^{m}(\tilde{y}^*_{it+h|t}(\hat{\tau}_{it})-\overline{\tilde{y}^*_{ih}(\hat{\tau}_i)})^2(u_{it+h}(\tau_i)-\overline{u_{ih}(\tau_i)})^2.$$

Using the mean value expansions above,

$$\tilde{y}^*_{it+h|t}(\hat{\tau}_i)-\overline{\tilde{y}^*_{ih}(\hat{\tau}_i)} = \tilde{D}_{it+h}(\bar{\tau}_{i0})(\hat{\tau}_{i0}-\tau_i)+\tilde{\Psi}_{it+h}(\hat{\tau}_{it},\hat{\tau}_{i0}),$$

where

$$\tilde{D}_{it+h}(\bar{\tau}_{is}) = D_{it+h}(\bar{\tau}_{is}) - m^{-1}\sum_{t=1}^{m}D_{it+h}(\bar{\tau}_{is}),$$
$$\tilde{\Psi}_{it+h}(\hat{\tau}_{it},\hat{\tau}_{i0}) = D_{it+h}(\bar{\tau}_{it})(\hat{\tau}_{it}-\hat{\tau}_{i0}) - m^{-1}\sum_{t=1}^{m}D_{it+h}(\bar{\tau}_{it})(\hat{\tau}_{it}-\hat{\tau}_{i0}),$$

for $s = 0,1,\ldots$. We can then write

$$\sum_{t=1}^{m}(\tilde{y}^*_{it+h|t}(\hat{\tau}_i)-\overline{\tilde{y}^*_{ih}(\hat{\tau}_i)})^2(\tilde{y}^*_{it+h}(\tau_i)-\overline{\tilde{y}^*_{ih}(\tau_i)})^2 = C_0+C_1+C_2,$$

where

$$C_0 = (\hat{\tau}_{i0} - \tau_i)^2 \sum_{t=1}^{m} \tilde{D}_{it+h}(\bar{\tau}_{i0})^2 (u_{it+h}(\tau_i) - \overline{u_{ih}(\tau_i)})^2,$$

$$C_1 = \sum_{t=1}^{m} \tilde{\Psi}_{it+h}(\hat{\tau}_{it}, \hat{\tau}_{i0})^2 (u_{it+h}(\tau_i) - \overline{u_{ih}(\tau_i)})^2,$$

$$C_2 = 2(\hat{\tau}_{i0} - \tau_i) \sum_{t=1}^{m} \tilde{D}_{it+h}(\bar{\tau}_{i0}) \tilde{\Psi}_{it+h}(\hat{\tau}_{it}, \hat{\tau}_{i0})(u_{it+h}(\tau_i) - \overline{u_{ih}(\tau_i)})^2.$$

Proceeding as before, the leading term

$$C_0 = O_p(T^{-1} m^{2v_i+1}),$$

whereas

$$C_1 = O_p(T^{-2} m^{2v_i+2}),$$
$$C_2 = O_p(T^{-3/2} m^{2v_i+5/2}),$$

by Assumption C and since $m^{-1} \sum_{t=1}^{m} u_{it+h}(\tau_i) = O_p(m^{v_i - 1/2})$. Thus,

$$Tm^{-1-2v_i}(C_0 + C_1 + C_2)$$
$$= T(\hat{\tau}_{i0} - \tau_i)^2 m^{1-2v_i} \frac{1}{m^2} \sum_{t=1}^{m} \tilde{D}_{it+h}(\bar{\tau}_{i0})^2 (u_{it+h}(\tau_i) - \overline{u_{ih}(\tau_i)})^2 + O_p(m^{3/2} T^{-3/2}). \quad \text{(C.3)}$$

Further, as in the proof of Theorem 1, we note again that we work with the projected series. Given the convergence rate, we require that $m^{1/2-v_i}(n^{-1} + (nT)^{-1/2}) \to 0$ as $(n, T, m)_j \to \infty$. Then, from (C.2) and (C.3), if additionally $m/T \to 0$, i.e., invoking Assumption E, and applying the FCLT, we establish that

$$m^{1/2-\hat{v}_i} \frac{1}{\hat{\omega}_{\Xi} m} \sum_{t=1}^{m} \hat{\Xi}_{it}^{h} \xrightarrow{d} N(0,1)$$

since consistency of $\hat{v}_i$ can be shown, reasoning as for $\hat{\kappa}_i$ in the proof of Theorem 1. □

Proof of Theorem 3. Due to asymptotic approximate independence of the individual ϱ_{ih} test statistics for large n, see Appendix 1 in Ergemen (2019), the result in the Theorem follows from a standard CLT. □

Proof of Theorem 4. We again motivate the cases based on the true parameters, given the parameter consistency arguments discussed earlier. Under the local alternative in (30),

$$\tilde{y}^*_{it+1}(\tau_i) - \overline{\tilde{y}^*_{i1}(\tau_i)} = \tilde{v}^*_{1it+1}(d_i - d_{i0}) - \overline{\tilde{v}^*_{1it+1}(d_i - d_{i0})} + (c_i/\sqrt{m})(\tilde{x}^*_{it}(d_i) - \overline{\tilde{x}^*_{it}(d_i)}),$$

where the model prediction error is $e_{it+1|t}(\tau_i) = \tilde{v}^*_{1it+1}(d_i - d_{i0}) + O_p(T^{-1/2})$ under Assumption A.1. Reasoning as in the proof of Theorem 1 and further noting that the error term is now $\tilde{v}^*_{1it+1}(d_i - d_{i0})$ as $T \to \infty$, we have under Assumption A.1 and B.2 that

$$m^{1/2} \frac{\overline{\tilde{v}^*_{1it+1}(d_i - d_{i0})}}{\sigma_{v_1}} \xrightarrow{d} w_{1i}$$

by a standard CLT, and

$$m^{1/2} \frac{1}{\sigma_{v_1} \sigma_x m} \sum_{t=1}^{m} (\tilde{x}^*_{it}(d_i) - \overline{\tilde{x}^*_{it}(d_i)}) \tilde{v}^*_{1it+1}(d_i - d_{i0}) \xrightarrow{d} w_{2i}$$

for fixed i, where w_{1i} and w_{2i} are two independent standard normally distributed random variables. Then,

$$\begin{aligned}
\sum_{t=1}^{m} \varphi^1_{it} &= (m^{1/2} \overline{\tilde{v}^*_{1it+1}(d_i - d_{i0})})^2 - \frac{2c_i}{\sqrt{m}} \sum_{t=1}^{m} (\tilde{x}^*_{it}(d_i) - \overline{\tilde{x}^*_{it}(d_i)}) \tilde{v}^*_{1it+1}(d_i - d_{i0}) \\
&\quad - \frac{c_i^2}{m} \sum_{t=1}^{m} (\tilde{x}^*_{it}(d_i) - \overline{\tilde{x}^*_{it}(d_i)})^2 + O_p(mT^{-1}) \\
&\xrightarrow{d} \sigma^2_{v_1} w^2_{i1} - 2\sigma_{v_1} \sigma_x c_i w_{i2} - c_i^2 \sigma_x^2
\end{aligned}$$

as $m/T \to 0$. Accordingly,

$$\widetilde{DM}_{i1} \xrightarrow{d} w^2_{i1} - 2c_i \frac{\sigma_x}{\sigma_{v_1}} - c_i^2 \frac{\sigma_x^2}{\sigma^2_{v_1}},$$

since $\hat{\sigma}^2_{v_1} = m^{-1} \sum_{t=1}^{m} (\tilde{v}^*_{1it+1}(d_i - d_{i0}) - \overline{\tilde{v}^*_{1it+1}(d_i - d_{i0})} + (c_i/\sqrt{m}) \tilde{x}^*_{it}(d_i))^2 = \sigma^2_{v_1} + O_p(m^{-1/2})$ under Assumption A.1.

Next, using

$$\tilde{y}^*_{it+1|t}(\tau_i) = \frac{c_i}{\sqrt{m}} \tilde{x}^*_{it}(d_i) + O_p(T^{-1/2})$$

under Assumption A.1, we have as $m/\sqrt{T} \to 0$,

$$\sum_{t=1}^{m} \Xi_{it}^{1} = \sum_{t=1}^{m} \left[\tilde{v}_{1it+1}^{*}(d_i - d_{i0}) - \overline{\tilde{v}_{1it+1}^{*}(d_i - d_{i0})} + (c_i/\sqrt{m})\tilde{x}_{it}^{*}(d_i) \right] \tilde{y}_{it+1|t}^{*}(\tau_i)$$
$$\xrightarrow{d} c_i \sigma_{v_1} \sigma_x w_{i2} + c_i^2 \sigma_x^2,$$

reasoning as above. Furthermore,

$$m\omega_{\Xi}^2 = \sum_{t=1}^{m} (\Xi_{it}^{1})^2 = \frac{c_i}{m} \sum_{t=1}^{m} (\tilde{v}_{1it+1}^{*}(d_i - d_{i0}) - \overline{\tilde{v}_{1it+1}^{*}(d_i - d_{i0})})^2 (\tilde{x}_{it}^{*}(d_i) - \overline{\tilde{x}_{it}^{*}(d_i)})^2 + o_p(1)$$
$$\xrightarrow{p} c_i^2 \sigma_{v_1}^2 \sigma_x^2.$$

Thus,

$$\varrho_{i1} \xrightarrow{d} sign(c_i) w_{i2} + |c_i| \frac{\sigma_x}{\sigma_{v_1}}.$$

□

Lemma A. *Under Assumption A.1,*

$$\sup_i \left| \frac{1}{m} \sum_{t=1}^{m} u_{it+h}(\tau_i) D_{it+h}(\bar{\tau}_{it}) \right| = O_p(1),$$
$$\sup_i \left| \frac{1}{m} \sum_{t=1}^{m} D_{it+h}(\bar{\tau}_{it})^2 \right| = O_p(1).$$

Proof of Lemma A. Let d_{min} and ϑ_{min} denote $\min_i d_i$ and $\min_i \vartheta_i$, respectively. We first observe that

$$\tilde{x}_{it}^{*}(d_i) = \Delta_t^{d_i - \vartheta_{i0}} v_{2it} = \sum_{j=0}^{t-1} \pi_j (d_i - \vartheta_{i0}) v_{2it-j},$$
$$\tilde{x}_{it}^{*}(\vartheta_i) = \Delta_t^{\vartheta_i - \vartheta_{i0}} v_{2it} = \sum_{j=0}^{t-1} \pi_j (\vartheta_i - \vartheta_{i0}) v_{2it-j}.$$

We then analyze the dominating terms in $\frac{1}{m} \sum_{t=1}^{m} u_{it+h}(\tau_i) D_{it+h}(\bar{\tau}_{it})$, since the remaining terms are bounded above by the Cauchy-Schwarz inequality. The first term,

$$\begin{aligned}\sup_i \left| \frac{1}{m} \sum_{t=1}^{m} \tilde{x}_{it}^*(d_i)\dot{\tilde{x}}_{it}^*(d_i) \right| &= \sup_i \left| \frac{1}{m} \sum_{t=1}^{m} \sum_{j=0}^{t-1} \pi_j(d_i - \vartheta_{i0})\dot{\pi}_j(d_i - \vartheta_{i0}) \right| \\ &= O_p(1 + m^{2(\vartheta_{max} - d_{min})-1} \log m) \\ &= O_p(1)\end{aligned}$$

under Assumption A.1. Similarly,

$$\begin{aligned}\sup_i \left| \frac{1}{m} \sum_{t=1}^{m} \tilde{x}_{it}^*(\vartheta_i)\dot{\tilde{x}}_{it}^*(\vartheta_i) \right| &= \sup_i \left| \frac{1}{m} \sum_{t=1}^{m} \sum_{j=0}^{t-1} \pi_j(\vartheta_i - \vartheta_{i0})\dot{\pi}_j(\vartheta_i - \vartheta_{i0}) \right| \\ &= O_p(1 + m^{2(\vartheta_{max} - \vartheta_{min})-1} \log m) \\ &= O_p(1).\end{aligned}$$

The third term,

$$\begin{aligned}&\sup_i \left| \frac{1}{m} \sum_{t=1}^{m} \dot{\tilde{x}}_{it}^*(d_i) \left(\tilde{v}_{1it+h}^*(d_i - d_{i0}) - \rho_i' \tilde{v}_{2it+h}^*(\vartheta_i - \vartheta_{i0}) \right) \right| \\ &\qquad = \sup_i \left| \frac{1}{m} \sum_{t=1}^{m} \sum_{j=0}^{t-1} \dot{\pi}_j(d_i - \vartheta_{i0}) \left(\pi_j(d_i - d_{i0}) + \pi_j(\vartheta_i - \vartheta_{i0}) \right) \right| \\ &\qquad = O_p(1 + m^{\vartheta_{max} + d_{max} - 2d_{min} - 1} \log m + m^{2\vartheta_{max} - d_{min} - \vartheta_{min} - 1} \log m) \\ &\qquad = O_p(1)\end{aligned}$$

by Assumption A.1 and following similar reasoning as above. The fourth term,

$$\begin{aligned}&\sup_i \left| \frac{1}{m} \sum_{t=1}^{m} \dot{\tilde{x}}_{it}^*(\vartheta_i) \left(\tilde{v}_{1it+h}^*(d_i - d_{i0}) - \rho_i' \tilde{v}_{2it+h}^*(\vartheta_i - \vartheta_{i0}) \right) \right| \\ &\qquad = O_p(1 + m^{\vartheta_{max} + d_{max} - d_{min} - \vartheta_{min} - 1} \log m + m^{2(\vartheta_{max} - \vartheta_{min})-1} \log m) \\ &\qquad = O_p(1).\end{aligned}$$

To analyze the expression $\frac{1}{m} \sum_{t=1}^{m} D_{it+h}(\bar{\tau}_{it})^2$, we again study only the dominating terms. The first term

$$\sup_i \left| \frac{1}{m} \sum_{t=1}^{m} \dot{\tilde{x}}_{it}^{*}(d_i)^2 \right| = \sup_i \left| \frac{1}{m} \sum_{t=1}^{m} \sum_{j=0}^{t-1} \dot{\pi}_j^2 (d_i - \vartheta_{i0}) \right|$$
$$= O_p(1 + m^{2(\vartheta_{max} - d_{min}) - 1} \log^2 m)$$
$$= O_p(1)$$

under Assumption A.1. Similarly, the second term

$$\sup_i \left| \frac{1}{m} \sum_{t=1}^{m} \dot{\tilde{x}}_{it}^{*}(\vartheta_i)^2 \right| = O_p(1 + m^{2(\vartheta_{max} - \vartheta_{min}) - 1} \log^2 m)$$
$$= O_p(1).$$

Finally, the third term

$$\sup_i \left| \frac{1}{m} \sum_{t=1}^{m} \dot{\tilde{x}}_{it}^{*}(d_i) \dot{\tilde{x}}_{it}^{*}(\vartheta_i) \right| = O_p(1 + m^{\vartheta_{max} + d_{max} - \vartheta_{min} - d_{min} - 1} \log^2 m)$$
$$= O_p(1).$$

All remaining terms are of smaller asymptotic size, and thus the results in the lemma follow. □

Research Papers 2019

CREATES
Center for Research in Econometric Analysis of Time Series

2018-24: Tom Engsted and Thomas Q. Pedersen: Disappearing money illusion

2018-25: Erik Christian Montes Schütte: In Search of a Job: Forecasting Employment Growth in the US using Google Trends

2018-26: Maxime Morariu-Patrichi and Mikko Pakkanen: State-dependent Hawkes processes and their application to limit order book modelling

2018-27: Tue Gørgens and Allan H. Würtz: Threshold regression with endogeneity for short panels

2018-28: Mark Podolskij, Bezirgen Veliyev and Nakahiro Yoshida: Edgeworth expansion for Euler approximation of continuous diffusion processes

2018-29: Isabel Casas, Jiti Gao and Shangyu Xie: Modelling Time-Varying Income Elasticities of Health Care Expenditure for the OECD

2018-30: Yukai Yang and Luc Bauwens: State-Space Models on the Stiefel Manifold with A New Approach to Nonlinear Filtering

2018-31: Stan Hurn, Nicholas Johnson, Annastiina Silvennoinen and Timo Teräsvirta: Transition from the Taylor rule to the zero lower bound

2018-32: Sebastian Ankargren, Måns Unosson and Yukai Yang: A mixed-frequency Bayesian vector autoregression with a steady-state prior

2018-33: Carlos Vladimir Rodríguez-Caballero and Massimiliano Caporin: A multilevel factor approach for the analysis of CDS commonality and risk contribution

2018-34: James G. MacKinnon, Morten Ørregaard Nielsen, David Roodman and Matthew D. Webb: Fast and Wild: Bootstrap Inference in Stata Using boottest

2018-35: Sepideh Dolatabadim, Paresh Kumar Narayan, Morten Ørregaard Nielsen and Ke Xu: Economic significance of commodity return forecasts from the fractionally cointegrated VAR model

2018-36: Charlotte Christiansen, Niels S. Grønborg and Ole L. Nielsen: Mutual Fund Selection for Realistically Short Samples

2018-37: Niels S. Grønborg, Asger Lunde, Kasper V. Olesen and Harry Vander Elst: Realizing Correlations Across Asset Classes

2018-38: Riccardo Borghi, Eric Hillebrand, Jakob Mikkelsen and Giovanni Urga: The dynamics of factor loadings in the cross-section of returns

2019-01: Andrea Gatto and Francesco Busato: Defining, measuring and ranking energy vulnerability

2019-02: Federico Carlini and Paolo Santucci de Magistris: Resuscitating the co-fractional model of Granger (1986)

2019-03: Martin M. Andreasen and Mads Dang: Estimating the Price Markup in the New Keynesian Model

2019-04: Daniel Borup, Bent Jesper Christensen and Yunus Emre Ergemen: Assessing predictive accuracy in panel data models with long-range dependence