JOURNAL OF THE AMERICAN STATISTICAL ASSOCIATION
2020, VOL. 115, NO. 532, 1598–1619: Applications and Case Studies
https://doi.org/10.1080/01621459.2019.1686392

Testing the Predictability of U.S. Housing Price Index Returns Based on an IVX-AR Model

Bingduo Yang[a], Wei Long[b], Liang Peng[c], and Zongwu Cai[d]

[a]Lingnan (University) College, Sun Yat-Sen University, Guangzhou, Guangdong, China; [b]Department of Economics, Tulane University, New Orleans, LA; [c]Department of Risk Management and Insurance, Georgia State University, Atlanta, GA; [d]Department of Economics, University of Kansas, Lawrence, KS

ABSTRACT

We use ten common macroeconomic variables to test for the predictability of the quarterly growth rate of house price index (HPI) in the United States during 1975:Q1–2018:Q2. We extend the instrumental variable based Wald statistic (IVX-KMS) proposed by Kostakis, Magdalinos, and Stamatogiannis to a new instrumental variable based Wald statistic (IVX-AR) which accounts for serial correlation and heteroscedasticity in the error terms of the linear predictive regression model. Simulation results show that the proposed IVX-AR exhibits excellent size control regardless of the degree of serial correlation in the error terms and the persistency in the predictive variables, while IVX-KMS displays severe size distortions. The empirical results indicate that the percentage of residential fixed investment in GDP is fairly a robust predictor of the growth rate of HPI. However, other macroeconomic variables' strong predictive ability detected by IVX-KMS is likely to be driven by the highly correlated error terms in the predictive regressions and thus becomes insignificant when the proposed IVX-AR method is implemented. Supplementary materials for this article, including a standardized description of the materials available for reproducing the work, are available as an online supplement.

ARTICLE HISTORY
Received January 2019
Accepted October 2019

KEYWORDS
Hypothesis test; Predictive regression model; Serially correlated errors; Stock return predictability

1. Introduction

An empirical observation about households in the United States is that they tend to concentrate their family holdings and assets on residential houses. Indeed, based on the Federal Reserve's 2017 *Survey of Consumer Finances*, it reveals that 63.7% of families own a primary residence with a median value of $185,000, which accounts for more than 60% of the net wealth of the median household. Due to the widespread increases in home prices after the global financial crisis, ownership rates and median and mean values of other residential property, which includes residences such as second homes and timeshares, also rose during the period of 2013–2017. It is well known that abnormal fluctuations in housing prices substantially impact the value of assets held by families and therefore, affect their decisions on consumption and investment. To closely track housing prices and predict the pattern of its movement, investors and policymakers usually rely on prior experience and some publicly available economic information. The relation between housing prices and various macroeconomic indicators has been extensively studied over the past decade, especially after the global financial crisis. For example, Del Negro and Otrok (2007) investigate how the expansionary monetary policy contributes to the increases in housing prices. Observing the path of the historical housing price, Shiller (2007) points out that the federal funds rate and the percentage of residential fixed investment in GDP can be useful indicators to predict housing bubbles. In a recent study, Kallberg, Liu, and Pasquariello (2014) examine the relation between the housing price in some U.S. cities and a large set of macroeconomic variables such as unemployment rate and per capita disposable income. Categorizing these variables as underlying systematic real and financial factors, they argue that the comovements among housing markets can be largely explained by changes in these factors.

In this article, we explore the possibility of using macroeconomic variables to predict the growth rate of the house price index. Testing predictability of returns of assets such as housing prices via lagged economic and financial variables with an unknown degree of persistence has been a cornerstone research topic in empirical finance. The classical predictive linear regression commonly employed in the literature is often built upon the following simultaneously structural linear model:

$$y_t = \alpha + \beta x_{t-1} + u_t \quad \text{and} \quad x_t = \pi x_{t-1} + e_t, \quad 1 \le t \le T, \quad (1)$$

where $\{(u_t, e_t)\}_{t=1}^{T}$ is a martingale difference sequence, and u_t and e_t may be correlated. The predictive variable x_{t-1} is assumed to follow an AR(1) process in (1) although the extension to a higher order autoregressive process such as AR(k) with $k > 1$ is straightforward. The above predictive regression model has been widely used in empirical economics and finance, for example, Granger causality testing, efficient market hypothesis, linear rational expectations hypothesis, asset pricing theories, and performance of a mutual fund.

Since the predictive variables x_{t-1} above (e.g., unemployment rate, interest rate, etc.) often exhibit a high degree of persistence (Welch and Goyal 2008), much like a nonstationary process, standard t-test result obtained from the ordinary

CONTACT Bingduo Yang bdyang2006@sina.com Lingnan (University) College, Sun Yat-Sen University, Guangzhou, Guangdong 510275, China.
Color versions of one or more of the figures in the article can be found online at www.tandfonline.com/r/JASA.
Supplementary materials for this article are available online. Please go to www.tandfonline.com/r/JASA.
These materials were reviewed for reproducibility.

least squares (OLS) method is no longer valid. The correlation coefficient between u_t and e_t is usually nonzero and negative in many real applications (see Table 4 in Campbell and Yogo (2006) and a real example in Cai and Wang (2014)), which causes the so-called "embedded endogeneity." In an early study, Stambaugh (1999) pointed out that this endogeneity will lead to a biased estimate of the slope coefficient β in finite samples and over-rejection of the null hypothesis of no predictive ability. This over-rejection can be interpreted as a tendency toward finding spurious predictive ability in a predictive regression model. Some bias-corrected based inferences have been proposed in the literature, such as the first order bias correction (Stambaugh 1999), the second-order bias correction (Amihud and Hurvich 2004), and the conservative bias-adjusted estimator (Lewellen 2004).

An alternative framework for inference is to assume that the predictive variable follows a near unit root process with the AR(1) coefficient $\pi = 1 - c/T$ where c is a constant and T is the sample size. The asymptotic properties for the coefficient of the near unit root predictive variable have been well developed in the previous studies (see Elliott and Stock 1994; Campbell and Yogo 2006; Jansson and Moreira 2006; Hjalmarsson 2011; Cai and Wang 2014). However, the constant term c in the coefficient cannot be consistently estimated in a limiting distribution, and thus the critical value for a predictability test cannot be correctly obtained.

During the last two decades, there have been many efforts to developing unified inferences that are robust to the persistent characteristics of the predictive variables and the embedded endogeneity arising from the two error terms in (1). The pioneer one is the Q-test which is based on the Bonferroni approach (Campbell and Yogo 2006). However, as pointed out by Phillips and Lee (2013), this approach might lead to severe undersizing, and the testing power is negligible.

The second line of the unified approach is to employ a weighted empirical likelihood method. For example, Zhu, Cai, and Peng (2014) propose a robust empirical likelihood inference for β, which always has a chi-squared limit regardless of the degree of persistence in predictive variables. To include the case of a stationary predicted variable on the left hand side and nonstationary predictive variables on the right hand side, Liu et al. (2019) construct a unified empirical likelihood inference to test the predictability by including the difference of the predictive variable into a simple linear predictive regression.

The third line of the unified approach is the instrumental variable estimation method (IVX) proposed by Magdalinos and Phillips (2009). At the cost of sacrificing the rate of convergence in the case of near unit root, the fundamental idea of this method is to construct instrumental variables that exhibit a lower degree of persistence than that of predictive variables to eliminate the endogeneity problem and robustify inference to the uncertainty over the precise nature of the integration. To test the general restrictions on predictive variables, Kostakis, Magdalinos, and Stamatogiannis (2015) showed that the IVX-based Wald test (IVX-KMS hereafter) yields a standard chi-squared distribution. In their article, the IVX-based test is robust to regressors' degree of persistence, ranging from purely stationary to purely nonstationary processes, in the presence of endogenous and heteroscedastic errors.[1] Meanwhile, Phillips and Lee (2016) show that this test remains valid for regressors with local unit roots in the explosive direction and mildly explosive roots. To reduce the bias in the predictive variable's slope coefficient, Demetrescu and Rodrigues (2016) proposed a residual-augmented IVX test statistic in a predictive regression context which is analog to that of Amihud and Hurvich (2004). They argued that their method is comparable in small samples under distinctive conditions such as strong persistence, endogeneity, non-Gaussian innovations, and heterogeneity. Also, Lee (2016) extended IVX filtering to quantile regression by correcting the distortions arising from highly persistent multivariate predictors while preserving the discriminatory power.

To test the predictability of the growth rate in housing price, one may apply one of the above unified methods with no need to concern the persistency of the predicting variables. Using Equation (1) to regress the growth rate of housing prices on some macroeconomic variables, one could obtain the estimated u_t and plot its autocorrelation coefficients. Results in Section 4 suggest that the assumption of uncorrelated error terms is questionable. Besides serial correlation, we also find evidence of persistent volatility in housing prices. To filter the linear dependence and the persistent volatility in u_t, we take a further step to fit an AR(p)-GARCH(m, n) model. The analysis with details given in Section 4 calls for an alternative predictability test for the growth rate of housing price, which should not only unify various degrees of persistency in predicting variables but also allow the error terms in the linear predictive regression to be serially correlated and heteroscedastic.

In this study, we propose a unified IVX-AR Wald statistic that accounts for such serial correlation in the error terms of the linear predictive regression model, which generalizes the IVX method in Kostakis, Magdalinos, and Stamatogiannis (2015) and Phillips and Lee (2016). Simulation results show that, if the error terms in the predictive regression are indeed serially correlated, IVX-KMS by Kostakis, Magdalinos, and Stamatogiannis (2015) suffers severe size distortions especially when the error terms are highly correlated, and the predictive variables are highly persistent. On the other hand, the proposed IVX-AR corrects the size distortions arising from serially correlated errors even in the presence of strong persistence and endogeneity. When the error terms are not serially correlated, these two methods exhibit similar performance.

Indeed, our motivation for this study is to implement the proposed method to test the predictive ability of ten widely used macroeconomic variables on the quarterly growth rate of the housing price in the United States during 1975:Q1–2018:Q2. We first consider the univariate case in which the growth rate of housing price is regressed on the ten regressors one by one, and the individual Wald tests are conducted under the null hypothesis that the individual regressor exhibits no predictive ability. By the IVX-AR method, only the percentage of residential fixed investment in GDP and the civilian unemployment rate

[1]The main assumption in Kostakis, Magdalinos, and Stamatogiannis (2015) is to assume that the error terms in the predictive model are a martingale difference sequence. However, many empirical studies have cast doubt on this assumption. See Ball and Kothari (1989), Case and Shiller (1989), and Getmansky, Lo, and Makarov (2004) as examples.

are found to have significant predictive ability with respect to the growth rate of housing prices. However, using IVX-KMS by Kostakis, Magdalinos, and Stamatogiannis (2015), we find that all ten regressors exhibit significant predictive ability on the growth rate of housing prices. In the multivariate case, we consider five combinations of the ten regressors and find that three of them are jointly significant, including the "kitchen sink" combination which accommodates all ten variables. This finding differs from that by Kostakis, Magdalinos, and Stamatogiannis (2015), who showed all five combinations are jointly significant. We apply the same procedures to a subperiod ranging from 2000:Q1 to 2018:Q2 and still find contradictory results. Given the main difference between IVX-AR and IVX-KMS is that the former accounts for the serial correlation in the error terms, we attribute the predictive ability falsely detected by IVX-KMS to the existence of high order of serial correlation in the error terms.

The remaining structure of the article is as follows. Section 2 introduces the IVX-AR based predictive regression model with an AR(p)+GARCH(m, n) process in the error terms. In the same section, we develop the estimation procedures and two hypothesis tests. Section 3 reports the finite-sample simulation results. Section 4 applies the proposed model to investigate the predictive ability of macroeconomic variables on the growth rate of housing prices. The final section draws conclusions, and the technical proofs are provided in the Appendix (supplementary materials).

2. Econometric Modeling

2.1. Model Setup and Assumptions

We consider the following linear predictive regression model with autocorrelation and conditional heteroscedasticity in error terms:

$$y_t = \alpha + x_{t-1}^\top \beta + u_t, \quad 1 \le t \le T, \tag{2}$$

where y_t is a predicted variable, for example, the growth rate of house price indexes or other asset returns, $x_{t-1} = (x_{t-1,1}, \dots, x_{t-1,d})^\top$ is a vector of d-dimensional predictive variables, for example, unemployment rate and interest rate, α is a constant, β is a vector of the corresponding slope coefficients of x_{t-1}, T is the sample size, and $A^\top$ denotes the transpose of a vector/matrix of A. In addition, we assume that the data $\{x_0, x_{-1}, \dots, x_{-q}\}$ are observed, and the error terms u_t are governed by an AR(q)+GARCH(m, n) process

$$u_t = \phi_1 u_{t-1} + \phi_2 u_{t-2} + \cdots + \phi_q u_{t-q} + v_t, \tag{3}$$

$$v_t = \sigma_t \epsilon_t, \qquad \sigma_t^2 = \omega_0 + \sum_{i=1}^{m} a_i v_{t-i}^2 + \sum_{j=1}^{n} b_j \sigma_{t-j}^2, \tag{4}$$

where $\phi = (\phi_1, \phi_2, \dots, \phi_q)^\top$ is a vector of coefficients for the AR(q). This idea of taking the AR error structure into account has been extensively studied; see Xiao, Linton, Carroll and Mammen (2003) and Liu, Chen, and Yao (2010) for nonparametric regression models; Hall and Yao (2003) for parametric regression models; Hill, Li, and Peng (2016) for a near unit root process. However, to the best of our knowledge, it is the first time to combine this idea with the IVX method to provide a unified predictability test allowing various persistency of predicting variables and correlated and heteroscedastic errors in the linear predictive regression.

Innovation $\{\epsilon_t\}$ is a sequence of the independent and identically distributed (iid) random variables with mean 0 and variance 1. In the literature, ϵ_t is usually assumed to follow a standard normal or standardized Student's t distribution or the generalized error distribution. The conditional volatility σ_t^2 follows a GARCH(m, n) model with parameters satisfying $\omega_0 > 0$, $a_i \ge 0$, $b_j \ge 0$, and $\sum_{i=1}^{\max(m,n)} (a_i + b_i) < 1$, in which the constraint on $a_i + b_i$ implies the unconditional variance of v_t is finite, whereas its conditional variance σ_t^2 evolves, see Tsay (2010). Specifically, model (4) reduces to an ARCH(m) model if $n = 0$. The IVX-KMS method by Kostakis, Magdalinos, and Stamatogiannis (2015) originally specifies the same GARCH(m, n) process as in Equation (4) and we use the same setting for conditional heteroscedasticity here.

Moreover, the predictive variables x_t are assumed to follow a vector autoregressive process as

$$x_t = \Pi_x x_{t-1} + e_t, \quad e_t = \sum_{\ell=0}^{\infty} \psi_\ell \varepsilon_{t-\ell}, \tag{5}$$

where $\Pi_x = I_d + C/T^{\eta_x}$ for some matrix $C = \mathrm{diag}(c_1, \dots, c_d)$ and some $0 \le \eta_x \le 1$, the notation I_d stands for the identity matrix of order d, $\{e_t = \sum_{\ell=0}^{\infty} \psi_\ell \varepsilon_{t-\ell}\}$ is a strictly stationary process, and v_t and ε_t may be correlated. By construction, the predictive variable belongs to one of the following persistence classes (Kostakis, Magdalinos, and Stamatogiannis 2015; Phillips and Lee 2016):

(I) Stationarity ($c_j < 0$ for $j = 1, \dots, d$ and $\eta_x = 0$);
(II) Moderate deviation from a unit root ($c_j < 0$ for $j = 1, \dots, d$ and $0 < \eta_x < 1$);
(III) Near unit root ($c_j < 0$ for $j = 1, \dots, d$ and $\eta_x = 1$);
(IV) Unit root ($c_j = 0$ for $j = 1, \dots, d$);
(V) Local unit root on the explosive side ($c_j > 0$ for $j = 1, \dots, d$ and $\eta_x = 1$);
(VI) Mildly explosive root ($c_j > 0$ for $j = 1, \dots, d$ and $0 < \eta_x < 1$).

The classes (I)–(VI) accommodate various degrees of persistence in predictive variables varying from a stationary one to a mildly explosive root. The model specification is completed in the Appendix (supplementary materials) with some additional formal assumptions.

2.2. Estimation Procedures

Define $\underline{y}_t = y_t - \bar{y}$, $\underline{x}_{t-1} = x_{t-1} - \bar{x}$, $\underline{u}_t = u_t - \bar{u}$ and $\underline{v}_t = v_t - \bar{v}$, where $\bar{y} = \frac{1}{T}\sum_{t=1}^{T} y_t$, $\bar{x} = \frac{1}{T}\sum_{t=1}^{T} x_{t-1}$, $\bar{u} = \frac{1}{T}\sum_{t=1}^{T} u_t$ and $\bar{v} = \frac{1}{T}\sum_{t=1}^{T} v_t$, then the demeaned version of models (2) and (3) are reduced to

$$\underline{y}_t = \underline{x}_{t-1}^\top \beta + \underline{u}_t, \quad 1 \le t \le T, \tag{6}$$

$$\underline{u}_t = \phi_1 \underline{u}_{t-1} + \phi_2 \underline{u}_{t-2} + \cdots + \phi_q \underline{u}_{t-q} + \underline{v}_t. \tag{7}$$

By construction, we rewrite Equation (6) as

$$\underline{y}_t - \sum_{j=1}^{q} \phi_j \underline{y}_{t-j} = (\underline{x}_{t-1} - \sum_{j=1}^{q} \phi_j \underline{x}_{t-j-1})^{\top} \beta + \underline{v}_t, \quad (8)$$

where $\underline{x}_{t-j-1} = (\underline{x}_{t-j-1,1}, \ldots, \underline{x}_{t-j-1,d})^{\top}$ for $j = 1, \ldots, q$. Equation (8) can be regarded as a predictive regression model with the predicted variable $\underline{y}_t - \sum_{j=1}^{q} \phi_j \underline{y}_{t-j}$ and the predictive variable $\underline{x}_{t-1} - \sum_{j=1}^{q} \phi_j \underline{x}_{t-j-1}$. A detailed generalization of the method in Kostakis, Magdalinos, and Stamatogiannis (2015) is as follows.

For a given vector of coefficients ϕ, we can construct an instrumental variable $\tilde{z}_{\phi,t-1}$ analog to the predictive variable $x_{t-1} - \sum_{j=1}^{q} \phi_j x_{t-j-1}$ as

$$\tilde{z}_{\phi,t-1} = \tilde{z}_{t-1} - \sum_{j=1}^{q} \phi_j \tilde{z}_{t-j-1},$$

where $\tilde{z}_t = \Pi_z \tilde{z}_{t-1} + \Delta x_t$ with $\tilde{z}_0 = 0$, $\Delta x_t = x_t - x_{t-1} = e_t + (C/T^{\eta_x}) x_{t-1}$ and $\Pi_z = I_d + C_z/T^{\eta_z}$ for some matrix C_z and $0 < \eta_z < 1$. As suggested by Kostakis, Magdalinos, and Stamatogiannis (2015), we use $C_z = -I_d$ and $\eta_z = 0.95$, which give quite good performance in the empirical implementation of the testing procedure. By construction, $\tilde{z}_t$ is less persistent when the predictive variable x_{t-1} is either unit root or near unit root.

To facilitate the estimation procedure, we define

$$\underline{\mathbf{x}}_{t-2,q} = (\underline{x}_{t-2}, \ldots, \underline{x}_{t-q-1}),\ \tilde{\mathbf{z}}_{t-2,q} = (\tilde{z}_{t-2}, \ldots, \tilde{z}_{t-q-1}),\ \underline{\mathbf{y}}_{t-1,q} = (\underline{y}_{t-1}, \ldots, \underline{y}_{t-q})^{\top},$$

$$\underline{\mathbf{u}}_{t-1,q} = (\underline{u}_{t-1}, \ldots, \underline{u}_{t-q})^{\top}, \text{ and } \mathbf{e}_{t-1,q} = (e_{t-1}, \ldots, e_{t-q})^{\top}.$$

Then $(\hat{\beta}^{\top}, \hat{\phi}^{\top})^{\top}$ satisfy

$$\sum_{t=1}^{T} (\tilde{z}_{t-1} - \tilde{\mathbf{z}}_{t-2,q}\hat{\phi}) \left[\underline{y}_t - \underline{\mathbf{y}}_{t-1,q}^{\top} \hat{\phi} - (\underline{x}_{t-1} - \underline{\mathbf{x}}_{t-2,q}\hat{\phi})^{\top} \hat{\beta} \right] = 0, \quad (9)$$

$$\sum_{t=1}^{T} (\underline{\mathbf{y}}_{t-1,q} - \underline{\mathbf{x}}_{t-2,q}^{\top} \hat{\beta}) \left[\underline{y}_t - \underline{\mathbf{y}}_{t-1,q}^{\top} \hat{\phi} - (\underline{x}_{t-1} - \underline{\mathbf{x}}_{t-2,q}\hat{\phi})^{\top} \hat{\beta} \right] = 0. \quad (10)$$

When $q = 1$, we propose a novel three-step algorithm to calculate $(\hat{\beta}^{\top}, \hat{\phi}_1)^{\top}$ as below.

(i) For a given value ϕ_1, we do IVX-KMS based regression which regresses $\underline{y}_t - \phi_1 \underline{y}_{t-1}$ on $\underline{x}_{t-1} - \phi_1 \underline{x}_{t-2}$, and obtain residuals $\hat{\underline{v}}_t(\phi_1)$, that is, $\hat{\underline{v}}_t(\phi_1) = \underline{y}_t - \phi_1 \underline{y}_{t-1} - (\underline{x}_{t-1} - \phi_1 \underline{x}_{t-2})^{\top} \hat{\beta}$ with $\hat{\beta}$ being the estimator from estimating Equation (9).

(ii) Find the estimator $\hat{\phi}_1^*$ by solving

$$\hat{\phi}_1^* = \operatorname{argmin}_{\phi_1} \sum_{t=1}^{T} \hat{\underline{v}}_t^2(\phi_1).$$

(iii) Obtain the coefficient estimators $\hat{\beta}^*$ by doing IVX-KMS based regression which regresses $\underline{y}_t - \hat{\phi}_1^* \underline{y}_{t-1}$ on $\underline{x}_{t-1} - \hat{\phi}_1^* \underline{x}_{t-2}$, that is, $\hat{\beta}^*$ is the solution of Equation (9) with ϕ_1 replaced by $\hat{\phi}_1^*$.

When $q > 1$, it is nontrivial to find a vector of estimators $\hat{\phi}$ in the step (ii) above. Here, we propose to obtain the estimators $\hat{\beta}$ and $\hat{\phi}$ by iteratively implementing the following steps:

(a) For a given vector $\hat{\phi}^{(0)}$, we calculate $\hat{\beta}^{(1)}$ by estimating Equation (9),

(b) For a given vector $\hat{\beta}^{(1)}$, we calculate $\hat{\phi}^{(1)}$ by estimating Equation (10) and update $\hat{\phi}^{(0)}$ to $\hat{\phi}^{(1)}$.

The initial value $\hat{\phi}^{(0)}$ can be obtained by regressing $\check{u}_t$ on $\check{u}_{t-1}, \ldots, \check{u}_{t-q}$ with $\check{u}_t$ being the OLS residuals from the model (2). Our simulation results suggest that this iterative procedure achieves convergence quickly.

2.3. Testing the Predictability of Regressors

The asymptotic mixed normality property of the IVX-AR estimator given in Theorem A of the Appendix (supplementary materials) suggests that the inference for predictability can be conducted by a standard Wald test. Specifically, we consider the multiple joint hypotheses with p constraints

$$H_0 : R\beta = r \qquad \text{versus} \qquad H_1 : R\beta \neq r,$$

where R is a $p \times d$ matrix, and r is a $p \times 1$ vector. Then, the IVX-AR based Wald statistic for testing the null hypothesis $H_0 : R\beta = r$ takes the form of

$$W_\beta = (R\hat{\beta} - r)^{\top} \hat{Q}_{R,\phi}^{-1} (R\hat{\beta} - r), \quad (11)$$

where

$$\hat{Q}_{R,\phi} = R \left[\sum_{t=1}^{T} (\tilde{z}_{t-1} - \tilde{\mathbf{z}}_{t-2,q}\hat{\phi})(\underline{x}_{t-1} - \underline{\mathbf{x}}_{t-2,q}\hat{\phi})^{\top} \right]^{-1} \hat{M}_\phi \left[\sum_{t=1}^{T} (\underline{x}_{t-1} - \underline{\mathbf{x}}_{t-2,q}\hat{\phi})(\tilde{z}_{t-1} - \tilde{\mathbf{z}}_{t-2,q}\hat{\phi})^{\top} \right]^{-1} R^{\top},$$

$$\hat{M}_\phi = T\hat{\Upsilon}_T - T\bar{z}_{T-1}\bar{z}_{T-1}^{\top} \check{\sigma}_{FM,\check{\phi}},$$

$$\hat{\Upsilon}_T = \frac{1}{T} \sum_{t=1}^{T} (\tilde{z}_{t-1} - \tilde{\mathbf{z}}_{t-2,q}\hat{\phi})(\tilde{z}_{t-1} - \tilde{\mathbf{z}}_{t-2,q}\hat{\phi})^{\top} \check{v}_t^2,$$

$$\check{\sigma}_{FM,\phi} = \check{\sigma}_v^2 - \check{\Omega}_{ve,\phi} \check{\Omega}_{ee,\phi}^{-1} \check{\Omega}_{ve,\phi}^{\top}, \quad \text{and} \quad \bar{z}_{T-1} = \frac{1}{T} \sum_{t=1}^{T} \tilde{z}_{t-1}$$

with $\check{v}_t$, $\check{\sigma}_v^2$, $\check{\Omega}_{ve,\phi}$, and $\check{\Omega}_{ee,\phi}$ being defined in the Appendix (supplementary materials).

Clearly, $\hat{\Upsilon}_T = \frac{1}{T} \sum_{t=1}^{T} (\tilde{z}_{t-1} - \tilde{\mathbf{z}}_{t-2,q}\hat{\phi})(\tilde{z}_{t-1} - \tilde{\mathbf{z}}_{t-2,q}\hat{\phi})^{\top} \check{v}_t^2$ is a White-type (1980) heteroscedasticity consistent estimator for $\Upsilon_T = E\left[(\tilde{z}_{t-1} - \tilde{\mathbf{z}}_{t-2,q}\phi)(\tilde{z}_{t-1} - \tilde{\mathbf{z}}_{t-2,q}\phi)^{\top} v_t^2\right]$. When the residuals v_t are homoscedastic, we can use $\hat{\Upsilon}_T^* = \frac{\check{\sigma}_v^2}{T} \sum_{t=1}^{T} (\tilde{z}_{t-1} - \tilde{\mathbf{z}}_{t-2,q}\hat{\phi})(\tilde{z}_{t-1} - \tilde{\mathbf{z}}_{t-2,q}\hat{\phi})^{\top}$ as a consistent estimator for $\Upsilon_T^* = \sigma_v^2 E(\tilde{z}_{t-1} - \tilde{\mathbf{z}}_{t-2,q}\phi)(\tilde{z}_{t-1} - \tilde{\mathbf{z}}_{t-2,q}\phi)^{\top}$. Following Kostakis, Magdalinos, and Stamatogiannis (2015), we use $T\bar{z}_{T-1}\bar{z}_{T-1}^{\top}\check{\sigma}_{FM,\phi}$ as a correction that removes the finite-sample distortion of the IVX-AR based estimators arising from the estimation for the intercept in the model (2).

Remark 1. It should be pointed out that Phillips and Magdalinos (2009) show that IVX is biased and provide a bias-corrected IVX estimation by directly estimating the asymptotic bias, which accounts for both series dependence and cross-series dependence, see their Equation (22). In this article, we employ the AR error structure to directly formulate the IVX estimation without any ad hoc bias-corrected estimation. Specifically, models (2) and (3) can be reformulated as a predictive regression model with the martingale structure in the residuals as

$$y_t - \sum_{j=1}^{q} \phi_j y_{t-j} = \alpha \left(1 - \sum_{j=1}^{q} \phi_j\right) + \left(x_{t-1} - \sum_{j=1}^{q} \phi_j x_{t-j-1}\right)^{\intercal} \beta + v_t.$$

The proposed IVX-AR estimation, which is based on the above expression with the martingale structure of $\{v_t\}$, is different from the bias-corrected IVX estimation by Phillips and Magdalinos (2009), in which they first derive the IVX estimator $\tilde{\beta}$ based on model (2), and then correct its asymptotic bias.

Simulation results, which are available from the authors upon request, show that the bias-corrected IVX estimation (Phillips and Magdalinos 2009) slightly improves the size control compared with IVX-KMS. However, it still suffers severe size distortion.

The theorem below follows Theorem A in the Appendix (supplementary materials):

Theorem 1. Assume that p is the number of constraints in the null hypothesis H_0. Under regularity conditions in Theorem A of the Appendix (supplementary materials), as $T \to \infty$,

$$W_\beta \Rightarrow \chi^2(p),$$

whenever the regressor x_{t-1} belongs to any class of (I)–(VI).

Theorem 1 holds regardless of the predictive variable x_{t-1} being stationary, moderate deviation from a unit root, near unit root, unit root, local unit root on the explosive side, or mildly explosive root. When the sample size is large, the estimation in the error terms as in Equations (3) and (4) does not show any effect on the asymptotic null distribution of the Wald statistic defined in (11) for testing the joint multiple hypothesis H_0.

2.4. Testing the Serial Correlation in the Error Terms

To see whether there exists autocorrelation in the error terms in model (2), we consider the multiple joint hypotheses

$$H_0^{(1)}: \phi_1 = \cdots = \phi_q = 0 \qquad \text{versus} \qquad H_1^{(1)}: \phi_j \neq 0, \quad \text{for some} \quad j \in \{1, \ldots, q\}.$$

Then, the Wald statistic for testing the null hypothesis $H_0^{(1)}$: $\phi_1 = \cdots = \phi_q = 0$ takes the form of

$$W_\phi = T \hat{\phi}^{\intercal} \check{\Sigma}_{uu,q} \check{V}_{uu,v}^{-1} \check{\Sigma}_{uu,q} \hat{\phi}, \tag{12}$$

where $\check{\Sigma}_{uu,q} = \frac{1}{T}\sum_{t=1}^{T} \check{\mathbf{u}}_{t-1,q} \check{\mathbf{u}}_{t-1,q}^{\intercal}$, $\check{V}_{uu,v} = \frac{1}{T}\sum_{t=1}^{T} \check{\mathbf{u}}_{t-1,q} \check{\mathbf{u}}_{t-1,q}^{\intercal} \check{v}_t^2$, and $\check{\mathbf{u}}_{t-1,q} = (\check{u}_{t-1}, \ldots, \check{u}_{t-q})^{\intercal}$ with $\check{u}_t$ defined earlier.

Again the theorem below follows Theorem A in the Appendix (supplementary materials).

Theorem 2. Assume that q is the dimension of ϕ in the null hypothesis $H_0^{(1)}$, that is, the order of AR model in Equation (3). Under the regularity conditions in Theorem A of the Appendix (supplementary materials), as $T \to \infty$,

$$W_\phi \Rightarrow \chi^2(q),$$

whenever the regressor x_{t-1} belongs to any class of (I)–(VI).

When the sample size is large, the estimation for α and β in Equation (2) does not show any effect on the asymptotic null distribution of the Wald statistic defined in Equation (12) for testing the multiple joint hypotheses $H_0^{(1)}$. Meanwhile, in an application we can choose the autoregressive order q by using either the Akaike information criterion (AIC) or the Bayesian information criterion (BIC).

3. Numerical Studies

3.1. Univariate Case

In this section, we evaluate the finite-sample performance of the Wald test using simulated datasets. To begin with, we consider the univariate case and assume that y_t is generated by

$$y_t = \alpha + \beta x_{t-1} + u_t, \text{ with } u_t = \phi_1 u_{t-1} + v_t,$$

$$x_t = \Pi x_{t-1} + e_t, \text{ with } e_t = \psi_1 e_{t-1} + \varepsilon_t,$$

where $1 \le t \le T$ and $v_t = \sigma_t \epsilon_t$ follows a GARCH(1,1) process such that $\sigma_t^2 = \omega + a_1 \epsilon_{t-1}^2 + b_1 \sigma_{t-1}^2$ with the GARCH parameter $\omega = 0.0001$, $a_1 = 0.04$, and $b_1 = 0.95$. $(\varepsilon_t, \epsilon_t)^{\intercal}$ is a random sample from a bivariate Gaussian copula $C(F_1(\epsilon_t), F_2(\varepsilon_t); \theta)$ with the copula parameter θ bounded between -1 and 1, and the two marginal distributions $F_1(\cdot)$ and $F_2(\cdot)$ follow the Student's t distribution.[2] Because nonnormality and heavy tails are common in most financial data, we assume both marginal distributions' degrees of freedom equal to 5.

Our target is to check whether the lagged variable x_{t-1} with an unknown degree of persistence can be used to predict y_t when u_t is serially correlated. For this purpose, we implement the Wald test under the null hypothesis $H_0 : \beta = 0$ versus the alternative $H_1 : \beta \neq 0$ with the 5% nominal level. We run 10,000 simulations and present the results with the sample size $T \in \{100, 200, 500\}$, $\phi_1 \in \{-0.9, -0.7, 0, 0.7, 0.9\}$, $\Pi \in \{0.2, 0.8, 0.95, 1, 1.005, 1.01\}$ and $(\psi_1, \theta) \in \{(0, 0), (0.2, 0.4)\}$. Here, the last two values of Π correspond to the scenario that regressor is mildly explosive, as discussed by Phillips and Lee (2016).[3] The remained four values of Π correspond to the four degrees of persistence (stationarity, moderate deviation from a unit root, nearly unit root, and unit root) exhibited by x_t, as defined in Section 2. $\theta = 0.4$ is an empirically relevant value, as will be shown in Table 1 in Section 4. Besides, to investigate the effect of the information criterion on choosing q, the order of autocorrelation in u_t, in each simulation, we document the performance of the proposed IVX-AR method using both AIC and BIC.[4]

[2] For copula models and their applications, we refer interest readers to the books by Joe (1997).

[3] We thank an anonymous reviewer for suggesting this.

[4] We thank an anonymous reviewer for suggesting this comparison.

Table 1. Results of univariate predictive regressions.

	$\hat{\beta}_{\text{OLS}}$	t_{OLS}	$\hat{\beta}_{\text{IVX-KMS}}$	$W_{\beta,\text{IVX-KMS}}$	$\hat{\beta}_{\text{IVX-AR}}$	$W_{\beta,\text{IVX-AR}}$	θ	q	W_{ϕ}
Panel 1: 1975:Q1–2018:Q2									
CPI	−0.0001	−5.211***	−0.0001	36.711***	−0.0001	1.803	−0.1685	5	134.042***
DEF	−0.0002	−5.147***	−0.0002	25.153***	−0.0002	1.356	0.0872	5	132.551***
GDP	0.0012	5.238***	0.0012	29.561***	−0.0002	0.816	0.2084	1	17.863**
INC	0.0016	2.898**	0.0027	18.622***	0.0000	0.016	0.1086	5	116.503***
IND	−0.0002	−3.427***	−0.0001	5.940**	0.0001	0.294	0.1329	5	146.373***
INT	0.0006	2.368***	0.0006	6.121**	−0.0001	0.090	0.1795	5	140.269***
INV	0.0070	8.750***	0.0078	92.486***	0.0070	13.785***	0.3888	1	17.556***
MOG	0.0006	1.988***	0.0006	4.350**	−0.0002	0.080	−0.0146	5	145.687***
RES	−0.0012	−2.526**	−0.0015	8.886***	0.0011	0.955	−0.2263	5	131.437***
UNE	−0.0018	−3.176***	−0.0030	20.885***	−0.0037	6.934***	−0.2659	1	32.737***
Panel 2: 2000:Q1–2018:Q2									
CPI	−0.0002	−2.262***	−0.0003	11.304***	−0.0002	0.728	−0.1685	1	27.873***
DEF	−0.0003	−1.802*	−0.0005	3.459*	−0.0003	0.361	0.0872	1	29.446***
GDP	0.0021	3.737***	0.0018	6.841***	0.0002	0.133	0.2084	1	13.251***
INC	0.0026	2.694***	0.0024	5.942**	0.0001	0.008	0.1086	1	22.537***
IND	0.0002	0.628	0.0004	1.322	0.0007	1.028	0.1329	1	33.358***
INT	0.0009	1.086	0.0012	0.863	0.0011	0.235	0.1795	1	32.353***
INV	0.0059	5.287***	0.0075	38.998***	0.0087	7.897***	0.3888	1	21.725***
MOG	0.0013	0.964	0.0018	0.811	0.0088	4.399**	−0.0146	1	32.704***
RES	−0.0010	−1.480	−0.0016	3.149*	0.0007	0.175	−0.2263	1	29.832***
UNE	−0.0040	−4.725***	−0.0043	17.103***	−0.0038	2.771*	−0.2659	1	20.049***

NOTE: This table documents the results of univariate predictive regressions for observations from the full sample period (Panel 1: 1975:Q1–2018:Q2) and the subperiod (Panel 2: 2000:Q1–2018:Q2). For both panels, the dependent variable is the quarterly growth rate of housing prices in the United States. *CPI* = quarterly consumer price index for all urban consumers: all items less shelter (Index 1982–1984 = 100). DEF = quarterly implicit price deflator of the gross domestic product (Index 2012 = 100). GDP = quarterly percent change of the gross domestic product from the preceding period. INC = quarterly percent change of real disposable personal income from quarter one year ago. IND = quarterly industrial production index (Index 2012 = 100). INT = quarterly effective federal funds rate. INV = quarterly percent of residential fixed investment in the gross domestic product. MOG = quarterly 30-year mortgage rate. RES = quarterly growth of the total reserve balances maintained with Federal Reserve banks. UNE = quarterly civilian unemployment rate. $\hat{\beta}_{\text{OLS}}$ refers to the least square estimate of each regression, and t_{OLS} represents the corresponding t-statistic. $\hat{\beta}_{\text{IVX-KMS}}$ indicates each univariate regression's slope coefficient estimated by Kostakis, Magdalinos, and Stamatogiannis (2015), and $W_{\beta,\text{IVX-KMS}}$ represents the corresponding Wald statistic for the null that the slope coefficient equals to zero. $\hat{\beta}_{\text{IVX-AR}}$ indicates each univariate regression's slope coefficient estimated by IVX-AR discussed in Section 2, and $W_{\beta,\text{IVX-AR}}$ represents the corresponding Wald statistic for the null that the slope coefficient equals to zero. θ refers to the correlation coefficient between the residuals of Equations (2) and (4). q denotes the optimal order of AR for the residuals of Equation (2) selected by the Bayesian information criterion (BIC). W_{ϕ} denotes the Wald statistic defined in Equation (12) and tests the null that $H_0: \phi_1 = \cdots = \phi_q = 0$ versus $H_1: \phi_j \neq 0$ for some $j \in \{1, 2, \ldots, q\}$. *, **, and ***, respectively, indicate rejection of the null at 10%, 5%, and 1%.

Panel 1 in Table 2 displays the empirical sizes in the case of no correlation ($\theta = 0$) between ϵ_t and ε_t and no autocorrelation ($\psi_1 = 0$) in e_t. When u_t is indeed serially correlated (i.e., $\phi_1 \in \{-0.9, -0.7, 0.7, 0.9\}$), it is clear that the Wald statistics calculated by IVX-AR(1) show excellent size control over all values of Π and sample size T. For IVX-KMS which ignores the serial correlation in u_t, its Wald statistics tend to be severely undersized when ϕ_1 is negative and x_t is highly persistent. On the other hand, the Wald statistics of IVX-KMS become remarkably oversized when ϕ_1 switches to be positive. Increasing the sample size T from 100 to 500 does not mitigate the size distortions. However, if x_t deviates from the mildly explosive and unit root process and becomes stationary (i.e., when $\Pi = 0.8$ or 0.2), distortions in empirical sizes by IVX-KMS appear to be smaller. When $\phi_1 = 0$ (i.e., no serial correlation in u_t), the performance of IVX-AR(1) and IVX-KMS is quite similar, and both empirical sizes are close to the 5% nominal level, which confirms the robustness of IVX-KMS to various degrees of conditional heteroscedasticity. This similarity should be expected because the main difference between IVX-AR and IVX-KMS is that the former accounts for the serial correlation in u_t. When comparing the performance of IVX-AR(1) using AIC with the counterpart using BIC, we do not find a statistically meaningful difference between the two methods, even though the former exhibits slightly better size control than the latter in most cases.

We then examine the power of the two methods' test statistics and plot their powers in Figure 1. To save space, here we only present the scenario that $\phi_1 = 0.7$ and $T = 200$. The six panels in Figure 1 correspond to the cases when Π, respectively, equals to 0.2, 0.8, 0.95, 1, 1.005, and 1.01. For power comparison, in each panel, we assume the true values of β, which is represented by the horizontal axis, equal to $\frac{2j}{T}$ where $j \in \{0, 1, 2, \ldots, 39, 40\}$. Therefore, $j = 0$ refers to the size of the test. Because of the severe oversizing in the Wald statistics by IVX-KMS, its power curve (blue-dotted) tends to be higher than that of IVX-AR(1) using AIC (black-solid) and BIC (red-dashed) when β is small, and Π is large, as shown by the last five panels in Figure 1. When the degree of persistence becomes quite low ($\Pi = 0.2$), the size distortions in the Wald statistic by IVX-KMS become much smaller, as suggested by Panel 1 in Table 2, and the dominance of IVX-AR's power curve becomes more obvious, as displayed by the first panel in Figure 1. In all six panels, we do not observe an obvious difference in power by IVX-AR(1) using AIC and BIC.

Next, we alternatively consider the case that $(\psi_1, \theta) = (0.2, 0.4)$, while keeping all other settings the same. Panel 2 in Table 2 documents the comparison of the empirical sizes by both methods, and Figure 2 plots their Wald statistics' power. In Table 2, most results from Panel 2 are similar to those from Panel

Table 2. Finite-sample sizes for univariate regression when u_t follows an AR(1) process.

	$\phi_1 = -0.9$			$\phi_1 = -0.7$			$\phi_1 = 0$			$\phi_1 = 0.7$			$\phi_1 = 0.9$		
Π	$W^{AIC}_{IVX\text{-}AR(1)}$	$W^{BIC}_{IVX\text{-}AR(1)}$	$W_{IVX\text{-}KMS}$	$W^{AIC}_{IVX\text{-}AR(1)}$	$W^{BIC}_{IVX\text{-}AR(1)}$	$W_{IVX\text{-}KMS}$	$W^{AIC}_{IVX\text{-}AR(1)}$	$W^{BIC}_{IVX\text{-}AR(1)}$	$W_{IVX\text{-}KMS}$	$W^{AIC}_{IVX\text{-}AR(1)}$	$W^{BIC}_{IVX\text{-}AR(1)}$	$W_{IVX\text{-}KMS}$	$W^{AIC}_{IVX\text{-}AR(1)}$	$W^{BIC}_{IVX\text{-}AR(1)}$	$W_{IVX\text{-}KMS}$
Panel 1: $\theta = 0, \psi_1 = 0$															
								$T = 100$							
0.2	0.064	0.064	0.029	0.049	0.051	0.033	0.063	0.063	0.057	0.051	0.053	0.092	0.056	0.055	0.095
0.8	0.050	0.049	0.001	0.050	0.049	0.002	0.048	0.045	0.055	0.059	0.059	0.293	0.061	0.062	0.419
0.95	0.043	0.042	0.000	0.051	0.047	0.000	0.047	0.046	0.056	0.062	0.063	0.365	0.062	0.060	0.564
1	0.061	0.063	0.000	0.057	0.059	0.001	0.050	0.049	0.054	0.047	0.044	0.371	0.051	0.049	0.569
1.005	0.050	0.056	0.000	0.051	0.057	0.000	0.047	0.052	0.045	0.051	0.054	0.376	0.073	0.078	0.580
1.01	0.046	0.055	0.000	0.046	0.050	0.000	0.045	0.048	0.042	0.053	0.052	0.377	0.062	0.066	0.591
								$T = 200$							
0.2	0.051	0.054	0.026	0.055	0.055	0.032	0.056	0.057	0.057	0.056	0.054	0.091	0.046	0.046	0.109
0.8	0.050	0.049	0.001	0.042	0.043	0.001	0.057	0.057	0.054	0.059	0.059	0.302	0.053	0.055	0.428
0.95	0.054	0.057	0.000	0.055	0.057	0.000	0.053	0.057	0.048	0.050	0.052	0.379	0.051	0.047	0.578
1	0.050	0.051	0.000	0.049	0.045	0.000	0.063	0.066	0.052	0.051	0.049	0.390	0.049	0.053	0.601
1.005	0.047	0.049	0.000	0.055	0.059	0.000	0.052	0.053	0.046	0.051	0.051	0.411	0.054	0.056	0.618
1.01	0.048	0.049	0.000	0.046	0.048	0.000	0.054	0.055	0.055	0.058	0.062	0.388	0.054	0.055	0.613
								$T = 500$							
0.2	0.051	0.052	0.022	0.054	0.053	0.027	0.054	0.054	0.038	0.045	0.041	0.079	0.050	0.055	0.096
0.8	0.052	0.040	0.001	0.057	0.057	0.001	0.045	0.044	0.047	0.049	0.047	0.301	0.063	0.067	0.410
0.95	0.051	0.054	0.000	0.055	0.054	0.001	0.053	0.055	0.053	0.059	0.059	0.364	0.050	0.049	0.590
1	0.059	0.059	0.000	0.046	0.046	0.000	0.045	0.045	0.054	0.059	0.062	0.400	0.051	0.056	0.623
1.005	0.045	0.048	0.000	0.056	0.057	0.000	0.055	0.055	0.051	0.054	0.053	0.417	0.061	0.067	0.665
1.01	0.040	0.038	0.000	0.050	0.054	0.000	0.053	0.056	0.055	0.055	0.057	0.405	0.047	0.056	0.635
Panel 2: $\theta = 0.4, \psi_1 = 0.2$															
								$T = 100$							
0.2	0.050	0.054	0.131	0.053	0.053	0.170	0.049	0.047	0.056	0.049	0.048	0.476	0.057	0.059	0.416
0.8	0.058	0.060	0.005	0.057	0.055	0.008	0.051	0.049	0.053	0.051	0.049	0.528	0.055	0.055	0.630
0.95	0.049	0.048	0.000	0.058	0.058	0.002	0.055	0.057	0.053	0.052	0.055	0.448	0.061	0.069	0.669
1	0.059	0.062	0.000	0.057	0.056	0.002	0.049	0.049	0.062	0.059	0.065	0.366	0.052	0.049	0.624
1.005	0.051	0.058	0.000	0.061	0.062	0.003	0.068	0.071	0.066	0.066	0.067	0.368	0.058	0.061	0.605
1.01	0.046	0.051	0.000	0.058	0.063	0.000	0.060	0.060	0.065	0.055	0.064	0.365	0.051	0.052	0.593

(Continued)

Table 2. Continued

	$\phi_1 = -0.9$			$\phi_1 = -0.7$			$\phi_1 = 0$			$\phi_1 = 0.7$			$\phi_1 = 0.9$		
Π	$W^{\mathrm{AIC}}_{\mathrm{IVX\text{-}AR(1)}}$	$W^{\mathrm{BIC}}_{\mathrm{IVX\text{-}AR(1)}}$	$W_{\mathrm{IVX\text{-}KMS}}$	$W^{\mathrm{AIC}}_{\mathrm{IVX\text{-}AR(1)}}$	$W^{\mathrm{BIC}}_{\mathrm{IVX\text{-}AR(1)}}$	$W_{\mathrm{IVX\text{-}KMS}}$	$W^{\mathrm{AIC}}_{\mathrm{IVX\text{-}AR(1)}}$	$W^{\mathrm{BIC}}_{\mathrm{IVX\text{-}AR(1)}}$	$W_{\mathrm{IVX\text{-}KMS}}$	$W^{\mathrm{AIC}}_{\mathrm{IVX\text{-}AR(1)}}$	$W^{\mathrm{BIC}}_{\mathrm{IVX\text{-}AR(1)}}$	$W_{\mathrm{IVX\text{-}KMS}}$	$W^{\mathrm{AIC}}_{\mathrm{IVX\text{-}AR(1)}}$	$W^{\mathrm{BIC}}_{\mathrm{IVX\text{-}AR(1)}}$	$W_{\mathrm{IVX\text{-}KMS}}$
Panel 2: $\theta = 0.4, \psi_1 = 0.2$															
								$T = 200$							
0.2	0.042	0.044	0.133	0.045	0.045	0.321	0.047	0.047	0.053	0.051	0.047	0.710	0.057	0.053	0.585
0.8	0.055	0.059	0.001	0.053	0.053	0.013	0.042	0.039	0.051	0.053	0.057	0.684	0.057	0.060	0.766
0.95	0.066	0.065	0.000	0.053	0.053	0.002	0.045	0.045	0.055	0.045	0.043	0.523	0.054	0.049	0.743
1	0.049	0.045	0.000	0.051	0.052	0.000	0.055	0.057	0.045	0.056	0.058	0.388	0.051	0.051	0.664
1.005	0.052	0.047	0.000	0.050	0.054	0.000	0.050	0.046	0.046	0.052	0.054	0.368	0.050	0.045	0.651
1.01	0.051	0.054	0.000	0.046	0.048	0.000	0.052	0.053	0.050	0.045	0.042	0.376	0.053	0.059	0.653
								$T = 500$							
0.2	0.063	0.062	0.405	0.055	0.054	0.702	0.051	0.050	0.055	0.049	0.046	0.956	0.051	0.053	0.884
0.8	0.049	0.047	0.002	0.048	0.047	0.036	0.056	0.055	0.049	0.056	0.055	0.909	0.052	0.047	0.935
0.95	0.049	0.053	0.000	0.049	0.049	0.002	0.050	0.052	0.052	0.058	0.057	0.691	0.050	0.048	0.878
1	0.051	0.054	0.000	0.047	0.051	0.000	0.049	0.045	0.051	0.051	0.047	0.390	0.043	0.043	0.696
1.005	0.054	0.055	0.000	0.059	0.061	0.000	0.052	0.053	0.052	0.063	0.057	0.381	0.050	0.056	0.630
1.01	0.050	0.051	0.000	0.053	0.054	0.000	0.041	0.041	0.045	0.050	0.056	0.405	0.054	0.063	0.668

NOTE: This table documents the empirical sizes for univariate regression, testing the null hypothesis $H_0 : \beta = 0$ versus the alternative $H_1 : \beta \neq 0$ under the 5% nominal size. $W^{\mathrm{AIC}}_{\mathrm{IVX\text{-}AR(1)}}$ and $W^{\mathrm{BIC}}_{\mathrm{IVX\text{-}AR(1)}}$ refer to the rejection rate for the Wald statistic calculated by the method proposed in Section 2 when q is, respectively, selected by AIC and BIC. $W_{\mathrm{IVX\text{-}KMS}}$ refers to the rejection rate for the Wald statistic calculated by the method proposed by Kostakis, Magdalinos, and Stamatogiannis (2015). Panel 1 refers to the case that $\theta = 0$ and $\psi_1 = 0$, while Panel 2 refers to the case that $\theta = 0.4$ and $\psi_1 = 0.2$. For both panels, u_t in Equation (2) follows an AR(1) process with the autoregressive coefficient $\phi_1 \in \{-0.9, -0.7, 0, 0.7, 0.9\}$, v_t in Equation (4) follows a GARCH(1,1) process with the GARCH parameters $\omega = 0.0001$, $a_1 = 0.04$, and $b_1 = 0.95$ and $\Pi \in \{0.2, 0.8, 0.95, 1, 1.005, 1.01\}$. The average rejection rates are calculated through 10,000 repetitions with sample size $T \in \{100, 200, 500\}$.

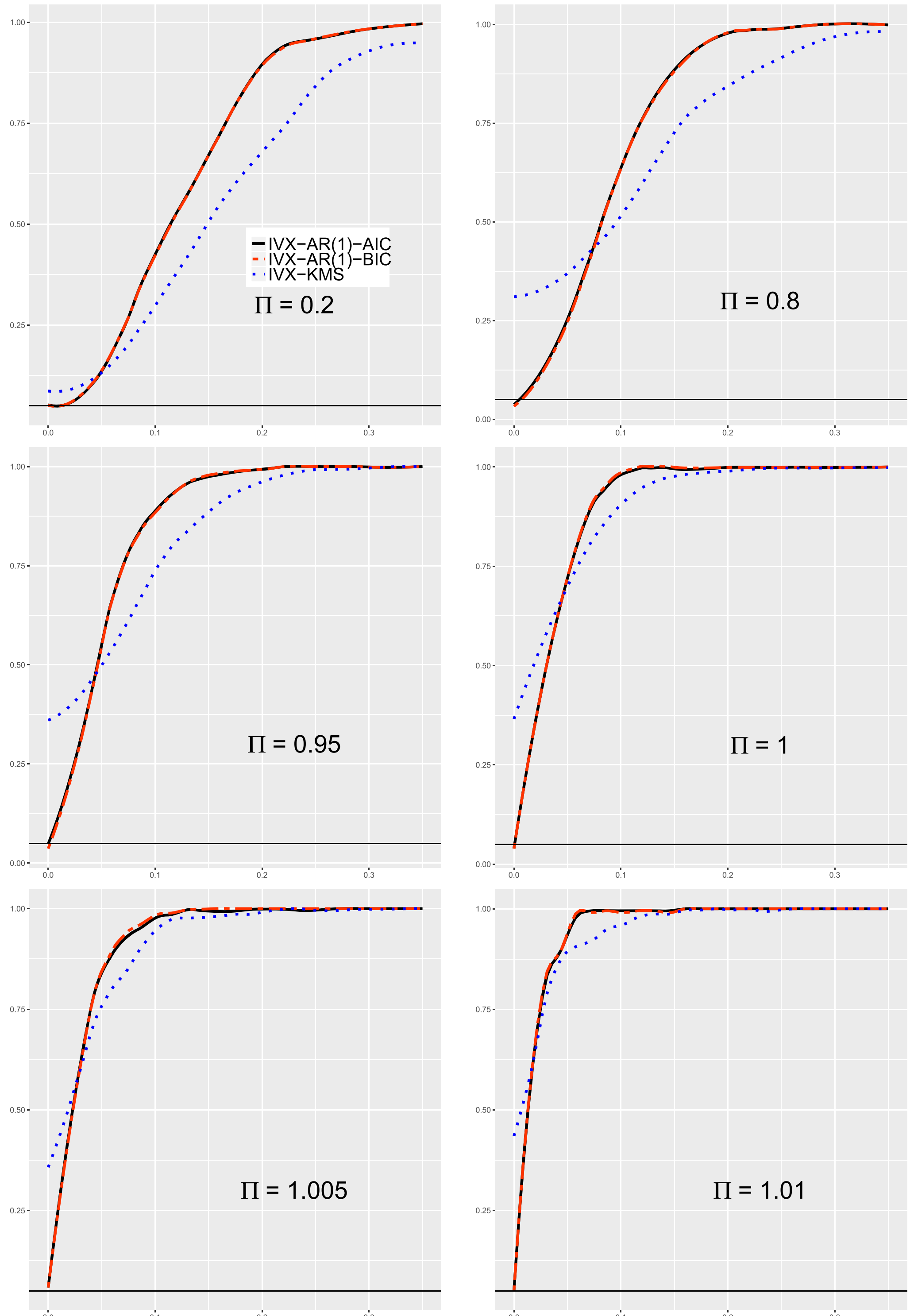


Figure 1. Power plots with $\theta = 0$ and $\psi_1 = 0$. NOTE: This figure displays the rejection rate for tests of the null hypothesis $\beta = 0$ versus the alternative $\beta \neq 0$ in Equation (2) as the true value of β (horizontal axis) increases under the univariate case. For each panel, $\beta = \frac{2j}{T}$ for $j \in \{0, 1, 2, \ldots, 39, 40\}$. The black-solid curve and the red-dashed curve, respectively, denote the rejection rate of the Wald statistic by the proposed IVX-AR method using AIC and BIC under the 5% nominal size (the horizontal solid line), while the blue-dotted curve denotes the rejection rate of the Wald statistic by the IVX-KMS method in Kostakis, Magdalinos, and Stamatogiannis (2015). The six panels correspond to the cases when $\Pi \in \{0.2, 0.8, 0.95, 1, 1.005, 1.01\}$. The autocorrelation coefficient for u_t is $\phi_1 = 0.7$ and the autocorrelation coefficient for e_t is $\psi_1 = 0$. The Gaussian copula parameter $\theta = 0$. v_t in Equation (4) follows a GARCH(1,1) process with the GARCH parameters $\omega = 0.0001$, $a_1 = 0.04$, and $b_1 = 0.95$. For each panel, the rejection rate is calculated with 10,000 repetitions and sample size $T = 200$.

1: the Wald statistics by IVX-AR display good size control over all combinations of Π and T, while the counterparts by IVX-KMS tend to be undersized when ϕ_1 is negative and oversized when ϕ_1 becomes positive. Increasing the sample size does not help to alleviate concerns on size distortions. When u_t is not serially correlated (i.e., $\phi_1 = 0$), the empirical sizes by both methods are close to the 5% nominal level. However, the size distortions are even larger in Panel 2 than in Panel 1, and we

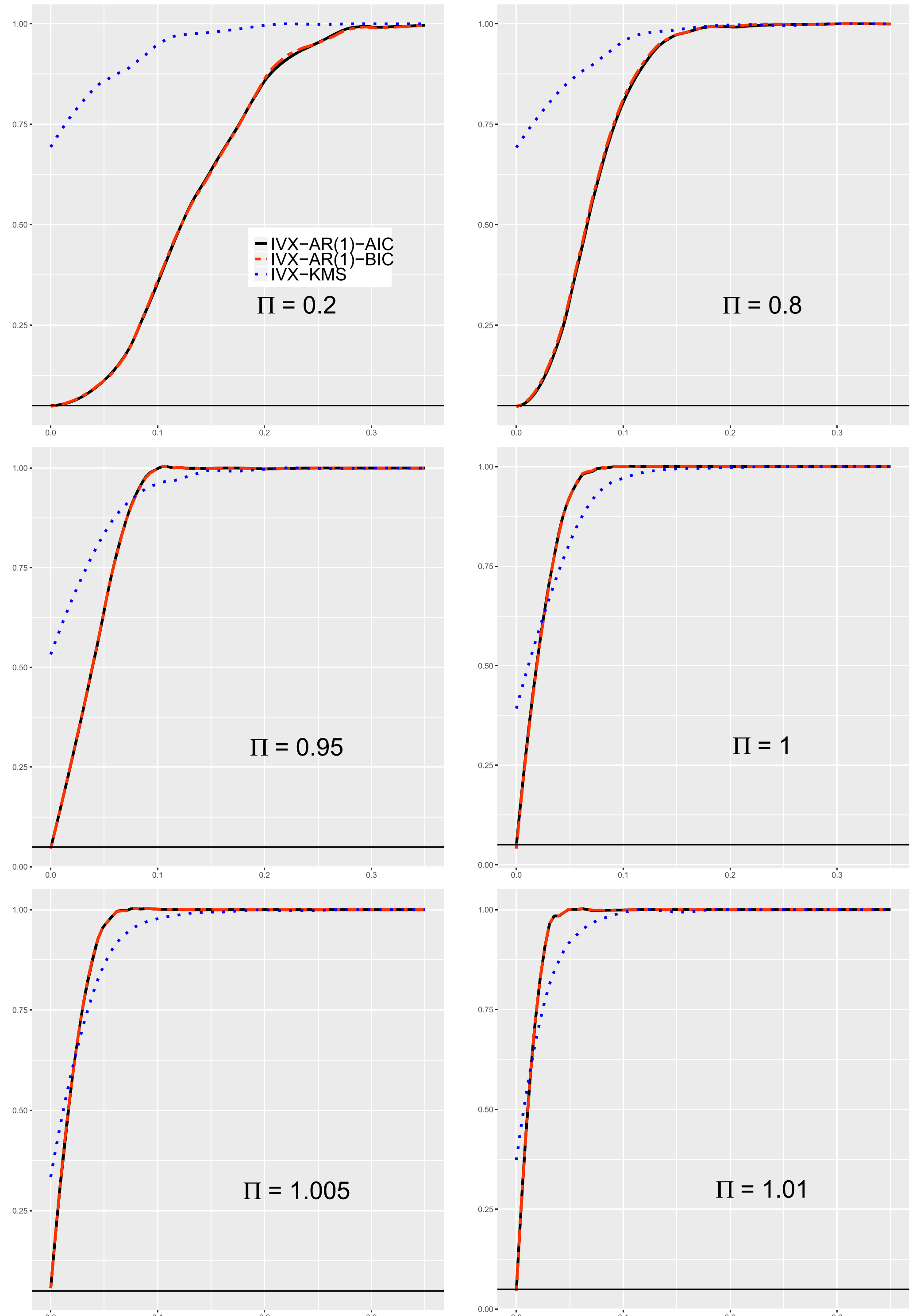


Figure 2. Power plots with $\theta = 0.4$ and $\psi_1 = 0.2$. NOTE: This figure displays the rejection rate for tests of the null hypothesis $\beta = 0$ versus the alternative $\beta \neq 0$ in Equation (2) as the true value of β (horizontal axis) increases under the univariate case. For each panel, $\beta = \frac{2j}{T}$ for $j \in \{0, 1, 2, \ldots, 39, 40\}$. The black-solid curve and the red-dashed curve, respectively, denote the rejection rate of the Wald statistic by the proposed IVX-AR method using AIC and BIC under the 5% nominal size (the horizontal solid line), while the blue-dotted curve denotes the rejection rate of the Wald statistic by the IVX-KMS method in Kostakis, Magdalinos, and Stamatogiannis (2015). The six panels correspond to the cases when $\Pi \in \{0.2, 0.8, 0.95, 1, 1.005, 1.01\}$. The autocorrelation coefficient for u_t is $\phi_1 = 0.7$ and the autocorrelation coefficient for e_t is $\psi_1 = 0.2$. The Gaussian copula parameter $\theta = 0.4$. v_t in Equation (4) follows a GARCH(1,1) process with the GARCH parameters $\omega = 0.0001$, $a_1 = 0.04$, and $b_1 = 0.95$. For each panel, the rejection rate is calculated with 10,000 repetitions and sample size $T = 200$.

could still observe obvious oversizing when ϕ_1 is positive. As in Panel 1, IVX-AR(1) using AIC and BIC still displays quite similar size control. Examining the six panels in Figure 2, which presents the case with $\phi_1 = 0.7$ and $T = 200$, we find consistent evidence of oversizing exhibited by the power curve of IVX-KMS. The power of the Wald statistics by IVX-AR(1) increases quickly as the true value of β increases. When x_t becomes nonstationary (i.e., $\Pi \geq 1$), as displayed by the last three panels in Figure 2, the power of the Wald statistic by IVX-AR(1) using AIC or BIC exhibits more obvious dominance over that by IVX-KMS. As in Figure 1, the power displayed by IVX-AR(1) using AIC and BIC is still quite similar.

For the univariate regression, we finally consider a more general case by evaluating the performance of IVX-AR(q) when q, the order of the autocorrelation in u_t, is greater than 1. As the empirical example in the next section will show, for certain macroeconomic variables such as the consumer price index and the effective federal funds rate, the optimal choice of the order is $q = 5$. To assure that the simulation setup for a higher order of q is relevant to the empirical results, we investigate the performance of IVX-AR(5) and compare it with that by IVX-KMS which completely ignores such high order of serial correlation in u_t. In other words, the simulation setup here is the same as before except that u_t now follows an AR(5) process and is defined as $u_t = \phi_1 u_{t-1} + \phi_2 u_{t-2} + \phi_3 u_{t-3} + \phi_4 u_{t-4} + \phi_5 u_{t-5} + v_t$, where $1 \le t \le T$. For the five autoregressive coefficients $(\phi_1, \phi_2, \phi_3, \phi_4, \phi_5)$, we consider five groups that are close to the results obtained in the empirical part: (i) (0.55, 0.18, −0.24, 0.30, −0.05); (ii) (0.42, −0.08, 0.33, 0.25, −0.07); (iii) (0.37, −0.10, 0.31, 0.27, 0.04); (iv) (0.24, −0.03, 0.45, 0.17, −0.09); (v) (0.16, 0.08, −0.25, 0.48, −0.05).

Table 3 shows that, by accounting for the serial correlation up to the 5th lag, the sizes of the Wald statistics by IVX-AR(5) are still very close to the 5% nominal size for all values of Π and T. On the other hand, the sizes of the Wald statistics by IVX-KMS are remarkably larger than the 5% nominal size in the presence of higher order serial correlation. This extreme oversizing does not mitigate as the sample size increases but becomes smaller when ϕ_1 decreases from 0.55 to 0.24. Similar to the scenario that $q = 1$, IVX-AR(5) using AIC and BIC still exhibit similar performance in terms of size control. The power plots in Figure 3 correspond to the scenario that the five autoregressive coefficients take the values in the group (i) and $T = 200$. As in Figure 2, Figure 3 illustrates that IVX-KMS is extremely oversized in all six panels, and the power of IVX-AR(5) using AIC or BIC increases rapidly as β grows, especially when $\Pi \ge 1$. In addition, compared with the power curves of IVX-AR(1) using AIC and BIC when $q = 1$, the difference between the two methods becomes slightly more salient when the order of autocorrelation is higher: the power displayed by IVX-AR(5) using AIC is relatively larger than BIC when $\Pi = 0.2$, while the power by IVX-AR(5) using BIC tends to be larger when $\Pi \ge 1$. However, the distance between the two curves is not substantial, and the performance between AIC and BIC is still quite similar.

3.2. Multivariate Case

In this part, we examine the finite-sample performance of the two methods' Wald tests for the multivariate regression model. We assume y_t is generated by

$$y_t = \alpha + x_{t-1}^{\mathsf{T}}\beta + u_t, \text{ with } u_t = \phi_1 u_{t-1} + v_t,$$

$$x_t = \Pi x_{t-1} + e_t, \text{ with } e_t = \psi_1 e_{t-1} + \varepsilon_t,$$

where $1 \le t \le T$, x_{t-1} is a 4×1 vector of four predictive variables and $\beta = (\beta_1, \beta_2, \beta_3, \beta_4)^{\mathsf{T}}$ are their associated slope coefficients. To render the examined setup empirically relevant, for the simulations of the multivariate case, we use values for Π's, ψ's, and ϕ_1 estimated from a multivariate predictive regression with the quarterly growth of HPI being the predicted variable and the quarterly implicit price deflator of GDP (DEF), quarterly effective federal funds rate (INT), quarterly percent change of real disposable personal income (INC), and quarterly percentage change of GDP (GDP) being the regressors. As will be shown in the next section, the four regressors display different degrees of persistence with $\Pi \in \{0.534, 0.674, 0.977, 0.998\}$, $(\psi_1, \psi_2, \psi_3, \psi_4) = (-0.116, -0.002, 0.174, 0.628)$ and $\phi_1 = 0.517$. For the correlation structure of the residuals $(\varepsilon_1, \varepsilon_2, \varepsilon_3, \varepsilon_4, \epsilon)$, for simplicity, we assume they are generated from a multi-normal distribution with mean $\mu = (0, 0, 0, 0, 0)$ and the empirically relevant covariance matrix

$$\begin{pmatrix} 0.0412 & 0.0347 & -0.0299 & 0.1088 & -0.0001 \\ 0.0347 & 0.7664 & 0.1072 & 0.9606 & -0.0008 \\ -0.0299 & 0.1072 & 1.4576 & 0.4638 & 0.0006 \\ 0.1088 & 0.9606 & 0.4638 & 11.4784 & 0.0011 \\ -0.0001 & -0.0008 & 0.0006 & 0.0011 & 0.0001 \end{pmatrix}.$$

Each simulation is repeated 10,000 times with the sample size $T \in \{100, 200, 500\}$. In the multivariate case, besides examining the performance of the individual Wald test on each slope coefficient, we additionally investigate the joint Wald test on β under different combinations of correlation structure (θ) and autocorrelation strength (ϕ_1) when the sample size T varies. Similar to the univariate case, we will also compare the performance of IVX-AR(1) using both AIC and BIC.

Table 4 documents the empirical sizes of both joint and individual Wald tests by IVX-AR(1) and IVX-KMS with 5% nominal level. W_{joint} denotes the rejection rate for the joint Wald test with the null hypothesis that the four regressors' slope coefficients simultaneously equal to zero (i.e., $\beta = (0, 0, 0, 0)^{\mathsf{T}}$). $W_{0.534}$, $W_{0.674}$, $W_{0.977}$, and $W_{0.998}$, respectively, denote the rejection rates for four individual Wald tests with the null hypothesis that one of the four regressors' slope coefficients equals to zero while letting the rest free. For example, $W_{0.534}$ refers to the rejection rate for the individual Wald test with the null hypothesis that $\beta_1 = 0$ while letting β_2, β_3, and β_4 free. From Table 4, we do not observe severe size distortion in the empirical sizes of the joint Wald test by IVX-AR(1) and all are close to the 5% nominal level. For the four individual Wald tests, we do not observe remarkable size distortions when IVX-AR(1) is applied. The performance of IVX-AR(1) using AIC and BIC is similar in terms of size control, even though the estimated empirical sizes using AIC are slightly closer to the 5% nominal size than BIC in many cases. However, we could observe similar patterns of size distortions in the empirical sizes of the Wald tests by IVX-KMS as in the univariate case, and increasing the sample size does not help to mitigate the size distortions. In summary, IVX-AR still outperforms IVX-KMS in multivariate regressions when u_t is indeed serially correlated.

Subsequently, we compare both methods' empirical power of the joint Wald test under the null that $\beta = (0, 0, 0, 0)^{\mathsf{T}}$ as the true value of one regressor's slope coefficient increases.

Table 3. Finite-sample sizes for univariate regression when u_t follows an AR(5) process.

	Group (i)			Group (ii)			Group (iii)			Group (iv)			Group (v)		
Π	$W^{AIC}_{IVX\text{-}AR(5)}$	$W^{BIC}_{IVX\text{-}AR(5)}$	$W_{IVX\text{-}KMS}$	$W^{AIC}_{IVX\text{-}AR(5)}$	$W^{BIC}_{IVX\text{-}AR(5)}$	$W_{IVX\text{-}KMS}$	$W^{AIC}_{IVX\text{-}AR(5)}$	$W^{BIC}_{IVX\text{-}AR(5)}$	$W_{IVX\text{-}KMS}$	$W^{AIC}_{IVX\text{-}AR(5)}$	$W^{BIC}_{IVX\text{-}AR(5)}$	$W_{IVX\text{-}KMS}$	$W^{AIC}_{IVX\text{-}AR(5)}$	$W^{BIC}_{IVX\text{-}AR(5)}$	$W_{IVX\text{-}KMS}$
							$T = 100$								
0.2	0.050	0.049	0.354	0.066	0.065	0.175	0.068	0.069	0.135	0.049	0.052	0.117	0.053	0.055	0.083
0.8	0.053	0.058	0.496	0.051	0.048	0.454	0.059	0.058	0.423	0.049	0.047	0.373	0.057	0.056	0.112
0.95	0.081	0.091	0.480	0.060	0.062	0.576	0.064	0.067	0.560	0.050	0.049	0.427	0.053	0.056	0.134
1	0.079	0.087	0.414	0.067	0.068	0.537	0.070	0.079	0.553	0.064	0.065	0.416	0.061	0.060	0.119
1.005	0.075	0.076	0.366	0.067	0.064	0.541	0.069	0.059	0.558	0.066	0.068	0.377	0.060	0.059	0.109
1.01	0.077	0.074	0.420	0.061	0.061	0.533	0.064	0.062	0.534	0.062	0.067	0.371	0.067	0.064	0.129
							$T = 200$								
0.2	0.053	0.059	0.585	0.060	0.061	0.280	0.076	0.082	0.199	0.056	0.058	0.179	0.050	0.050	0.092
0.8	0.052	0.054	0.552	0.050	0.047	0.587	0.063	0.063	0.511	0.051	0.053	0.505	0.066	0.065	0.155
0.95	0.079	0.089	0.552	0.045	0.041	0.664	0.058	0.057	0.657	0.060	0.060	0.562	0.052	0.053	0.169
1	0.079	0.094	0.453	0.069	0.069	0.599	0.070	0.073	0.678	0.068	0.067	0.463	0.051	0.053	0.125
1.005	0.052	0.052	0.425	0.066	0.064	0.598	0.058	0.067	0.633	0.051	0.053	0.444	0.058	0.058	0.130
1.01	0.055	0.058	0.416	0.070	0.076	0.603	0.064	0.068	0.644	0.056	0.059	0.466	0.053	0.055	0.146
							$T = 500$								
0.2	0.066	0.069	0.596	0.064	0.077	0.521	0.054	0.078	0.406	0.055	0.059	0.338	0.051	0.048	0.168
0.8	0.064	0.066	0.576	0.054	0.057	0.595	0.052	0.057	0.625	0.056	0.055	0.437	0.047	0.046	0.273
0.95	0.062	0.063	0.544	0.051	0.048	0.619	0.060	0.062	0.699	0.049	0.054	0.527	0.051	0.049	0.229
1	0.061	0.071	0.477	0.069	0.075	0.659	0.051	0.059	0.625	0.053	0.057	0.517	0.048	0.046	0.130
1.005	0.056	0.056	0.049	0.061	0.061	0.613	0.056	0.056	0.694	0.062	0.063	0.493	0.050	0.056	0.139
1.01	0.061	0.064	0.463	0.063	0.069	0.648	0.059	0.058	0.738	0.058	0.061	0.490	0.038	0.034	0.155

NOTE: This table documents the empirical sizes for univariate regression, testing the null hypothesis $H_0 : \beta = 0$ versus the alternative $H_1 : \beta \neq 0$ under the 5% nominal size. u_t in Equation (2) follows an AR(5) process with the five autoregressive coefficients $(\phi_1, \phi_2, \phi_3, \phi_4, \phi_5)$ are selected from five groups that are close to the results obtained in the empirical part: (i) $(0.55, 0.18, -0.24, 0.30, -0.05)$; (ii) $(0.42, -0.08, 0.33, 0.25, -0.07)$; (iii) $(0.37, -0.10, 0.31, 0.27, 0.04)$; (iv) $(0.24, -0.03, 0.45, 0.17, -0.09)$; (v) $(0.16, 0.08, -0.25, 0.48, -0.05)$. $W^{AIC}_{IVX\text{-}AR(5)}$ and $W^{BIC}_{IVX\text{-}AR(5)}$ refer to the rejection rate for the Wald statistic calculated by the method proposed in Section 2 when q is, respectively, selected by AIC and BIC. $W_{IVX\text{-}KMS}$ refers to the rejection rate for the Wald statistic calculated by the method proposed by Kostakis, Magdalinos, and Stamatogiannis (2015). v_t in Equation (4) follows a GARCH(1,1) process with the GARCH parameters $\omega = 0.0001, a_1 = 0.04$, and $b_1 = 0.95$. For all cases, $\Pi \in \{0.2, 0.8, 0.95, 1, 1.005, 1.01\}$, e_t in Equation (5) follows an AR(1) process with the autoregressive coefficient $\psi_1 = 0.2$, and the Gaussian copula parameter $\theta = 0.4$. The average rejection rates are calculated through 10,000 repetitions with sample size $T \in \{100, 200, 500\}$.

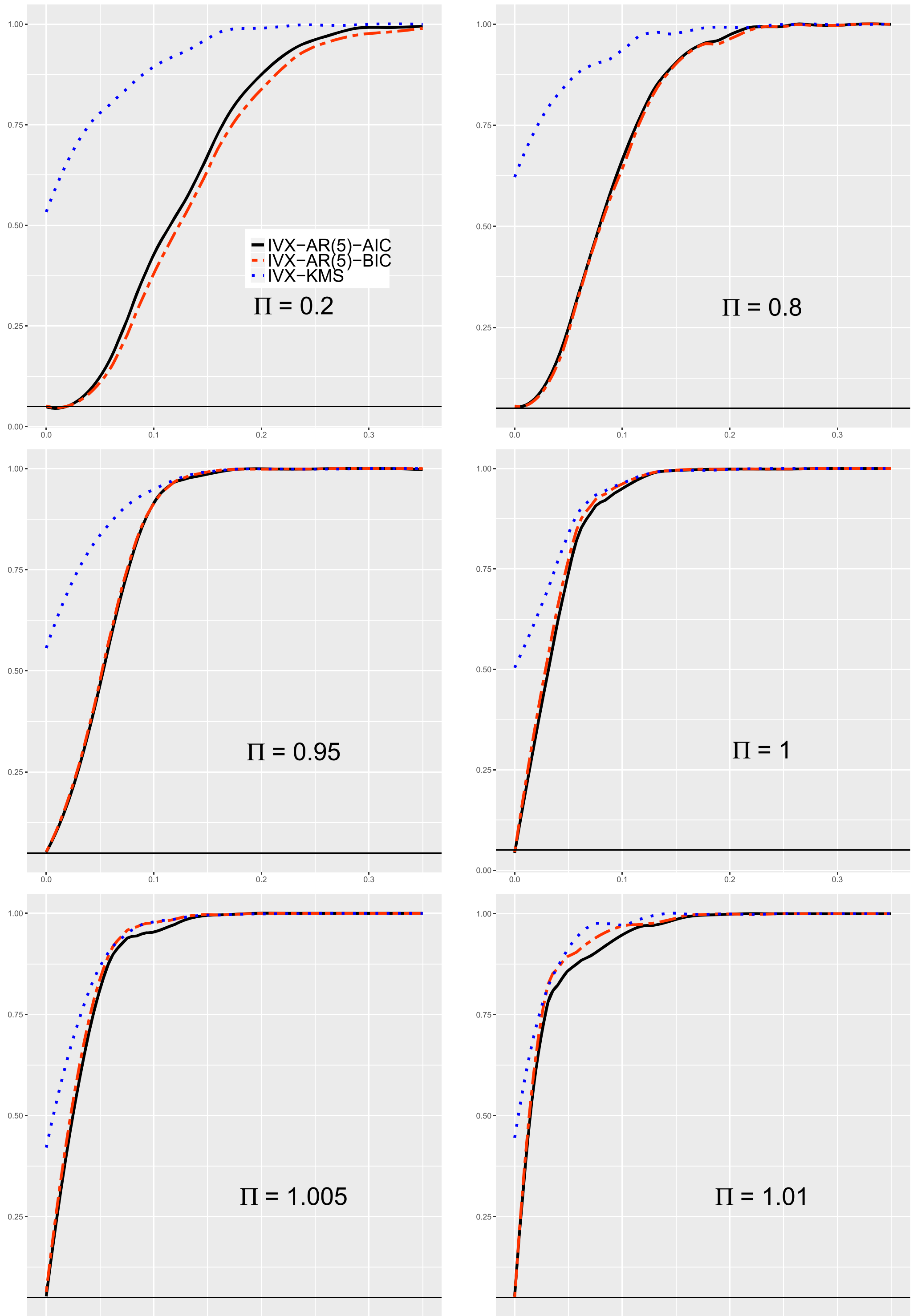


Figure 3. Power plots with u_t follows AR(5). NOTE: This figure displays the rejection rate for tests of the null hypothesis $\beta = 0$ versus the alternative $\beta \neq 0$ in Equation (2) as the true value of β (horizontal axis) increases under the univariate case. For each panel, $\beta = \frac{2j}{T}$ for $j \in \{0, 1, 2, \ldots, 39, 40\}$. The black-solid curve and the red-dashed curve, respectively, denote the rejection rate of the Wald statistic by the proposed IVX-AR method using AIC and BIC under the 5% nominal size (the horizontal solid line), while the blue-dotted curve denotes the rejection rate of the Wald statistic by the IVX-KMS method in Kostakis, Magdalinos, and Stamatogiannis (2015). The six panels correspond to the cases when $\Pi \in \{0.2, 0.8, 0.95, 1, 1.005, 1.01\}$. u_t in Equation (2) follows an AR(5) process with $(\phi_1, \phi_2, \phi_3, \phi_4, \phi_5) = (0.55, 0.18, -0.24, 0.30, -0.05)$. The autocorrelation coefficient for e_t is $\psi_1 = 0.2$. The Gaussian copula parameter $\theta = 0.4$. v_t in Equation (4) follows a GARCH(1,1) process with the GARCH parameters $\omega = 0.0001$, $a_1 = 0.04$, and $b_1 = 0.95$. For each panel, the rejection rate is calculated with 10,000 repetitions and sample size $T = 200$.

The four panels in Figure 4 refer to the multivariate regression's four regressors and their corresponding degrees of persistence when $T = 200$ and all parameters are set to be empirically relevant. For example, the first panel refers to the power of the joint Wald test by IVX-AR(1) using AIC and BIC and IVX-KMS under the alternative hypothesis that $\beta = \left(\frac{2j}{T}, 0, 0, 0\right)^{\intercal}$, where $j \in \{0, 1, 2, \ldots, 39, 40\}$. In other words, it implies the case when β_1, the slope coefficient associated to the regressor with a low degree of persistence ($\Pi = 0.534$), is nonzero while $\beta_2 = \beta_3 = \beta_4 = 0$. In the same manner, the next three panels correspond to the power of the joint Wald tests under the alternative hypotheses that $\beta = \left(0, \frac{2j}{T}, 0, 0\right)^\top$, $\beta = \left(0, 0, \frac{2j}{T}, 0\right)^\top$, and $\beta = \left(0, 0, 0, \frac{2j}{T}\right)^\top$. Figure 4 illustrates that the power of the Wald test by IVX-AR(1) increases quickly and the gap between the curve of IVX-KMS and the curve of IVX-AR(1) using AIC or BIC becomes narrower when the degree of persistence increases. For the Wald statistics by IVX-KMS, we can still observe obvious oversizing in all four panels when the slope coefficients indeed equal to zero. The power curve of IVX-AR(1) using AIC tracks the counterpart using BIC closely, and we do not observe obvious differences between the two curves.

4. Predictability of HPI Growth Rate

In this section, we implement the proposed IVX-AR method to test the predictive ability of some fundamental macroeconomic factors on the housing price in the United States. From the Federal Housing Finance Agency (FHFA), we collect the quarterly House Price Index (HPI, Index 1980:Q1 = 100) during 1975:Q1–2018:Q2. This index is based on all-transactions of properties and calculated through a modified version of the weighted-repeat sales by Case and Shiller (1989). Compared with the widely cited Standard and Poor/Case-Shiller Index, HPI covers more transactions and longer time interval, and thus can well represent the trend of the national-wide housing price. Based on HPI, we calculate the quarterly growth rate of the housing price and use this rate as the dependent variable. Figure 5 plots HPI and its growth rates during 1975:Q1–2018:Q2. The shaded areas denote the economic recessions defined by the National Bureau of Economic Research (NBER). Figure 5 shows that the housing price has kept a steady upward trend since 1975 and started its booming cycle after 1998 due to the housing price increase in large cities such as San Diego, Seattle, and Los Angeles on the West coast (Shiller 2007). The housing bubble collapsed during the 2007 subprime mortgage crisis. Thanks to a series of fiscal and monetary policies, the housing market began to gradually recover, and the housing price index reached its historical high in 2017.

The comovment between the housing price and macroeconomic variables has already been widely studied (e.g., Del Negro and Otrok 2007; Kallberg, Liu, and Pasquariello 2014). In this analysis, we consider ten macroeconomic variables collected from the Federal Reserve Economic Data (FRED), and all data are quarterly between 1975:Q1 and 2018:Q2.

CPI: Consumer price index with all items less shelter for all urban consumers (Index 1982–1984 = 100).
DEF: The implicit price deflator of the gross domestic product (Index 2012 = 100).
GDP: Percent change of the gross domestic product from the preceding period.
INC: Percent change of the real disposable personal income from the quarter one year ago.

Table 4. Finite-sample size for multivariate regression when $\boldsymbol{u_t}$ follows an AR(1) process.

	W_{joint}			$W_{0.534}$			$W_{0.674}$			$W_{0.977}$			$W_{0.998}$		
T	$\text{IVX}^{\text{AIC}}_{\text{AR}(1)}$	$\text{IVX}^{\text{BIC}}_{\text{AR}(1)}$	IVX_{KMS}	$\text{IVX}^{\text{AIC}}_{\text{AR}(1)}$	$\text{IVX}^{\text{BIC}}_{\text{AR}(1)}$	IVX_{KMS}	$\text{IVX}^{\text{AIC}}_{\text{AR}(1)}$	$\text{IVX}^{\text{BIC}}_{\text{AR}(1)}$	IVX_{KMS}	$\text{IVX}^{\text{AIC}}_{\text{AR}(1)}$	$\text{IVX}^{\text{BIC}}_{\text{AR}(1)}$	IVX_{KMS}	$\text{IVX}^{\text{AIC}}_{\text{AR}(1)}$	$\text{IVX}^{\text{BIC}}_{\text{AR}(1)}$	IVX_{KMS}
100	0.056	0.067	0.194	0.054	0.054	0.099	0.052	0.056	0.107	0.059	0.060	0.128	0.053	0.048	0.130
200	0.051	0.058	0.215	0.050	0.051	0.086	0.050	0.050	0.134	0.052	0.053	0.137	0.052	0.058	0.137
500	0.530	0.550	0.202	0.051	0.053	0.078	0.046	0.044	0.120	0.057	0.059	0.139	0.052	0.055	0.146

NOTE: This table documents the empirical sizes of five Wald tests for multivariate regression with four regressors that exhibit distinctive degrees of persistence. In this table, the parameters are set to be empirically relevant values with $\Pi = (0.534, 0.674, 0.977, 0.998)$, $\phi_1 = 0.517$ and $(\psi_1, \psi_2, \psi_3, \psi_4) = (-0.116, -0.002, 0.174, 0.628)$. $(\varepsilon_1, \varepsilon_2, \varepsilon_3, \varepsilon_4, \epsilon)$ are generated from a multi-normal distribution with mean $(0, 0, 0, 0, 0)$ and the empirically relevant covariance matrix defined in Section 3.2. v_t in Equation (4) follows a GARCH(1,1) process with the GARCH parameters $\omega = 0.0001$, $a_1 = 0.04$ and $b_1 = 0.95$. W_{joint} denotes the rejection rate for the joint Wald test with the null hypothesis that the four slope coefficients associated with the four regressors simultaneously equal zero (i.e., $\beta = (0, 0, 0, 0)^\top$) under the 5% nominal size. $W_{0.534}$ denotes the rejection rate for the Wald test with the null hypothesis that the coefficient of the first regressor (stationarity) equals zero while letting β_2, β_3, and β_4 free under the 5% nominal size. $W_{0.674}$ denotes the rejection rate for the Wald test with the null hypothesis that the coefficient of the second regressor (moderate deviation from unit root) equals zero while letting β_1, β_3, and β_4 free under the 5% nominal size. $W_{0.977}$ denotes the rejection rate for the Wald test with the null hypothesis that the coefficient of the third regressor (nearly unit root) equals zero while letting β_1, β_2, and β_4 free under the 5% nominal size. $W_{0.998}$ denotes the rejection rate for the Wald test with the null hypothesis that the coefficient of the fourth regressor (nearly unit root) equals zero while letting β_1, β_2, and β_3 free under the 5% nominal size. $W^{\text{AIC}}_{\text{IVX-AR}(1)}$ and $W^{\text{BIC}}_{\text{IVX-AR}(1)}$ refer to the rejection rate for the Wald statistic calculated by the method proposed in Section 2 when q is, respectively, selected by AIC and BIC. IVX_{KMS} refers to the rejection rate for the Wald statistic calculated by the method proposed by Kostakis, Magdalinos, and Stamatogiannis (2015). The average rejection rates are calculated through 10,000 repetitions with the sample size $T \in \{100, 200, 500\}$.

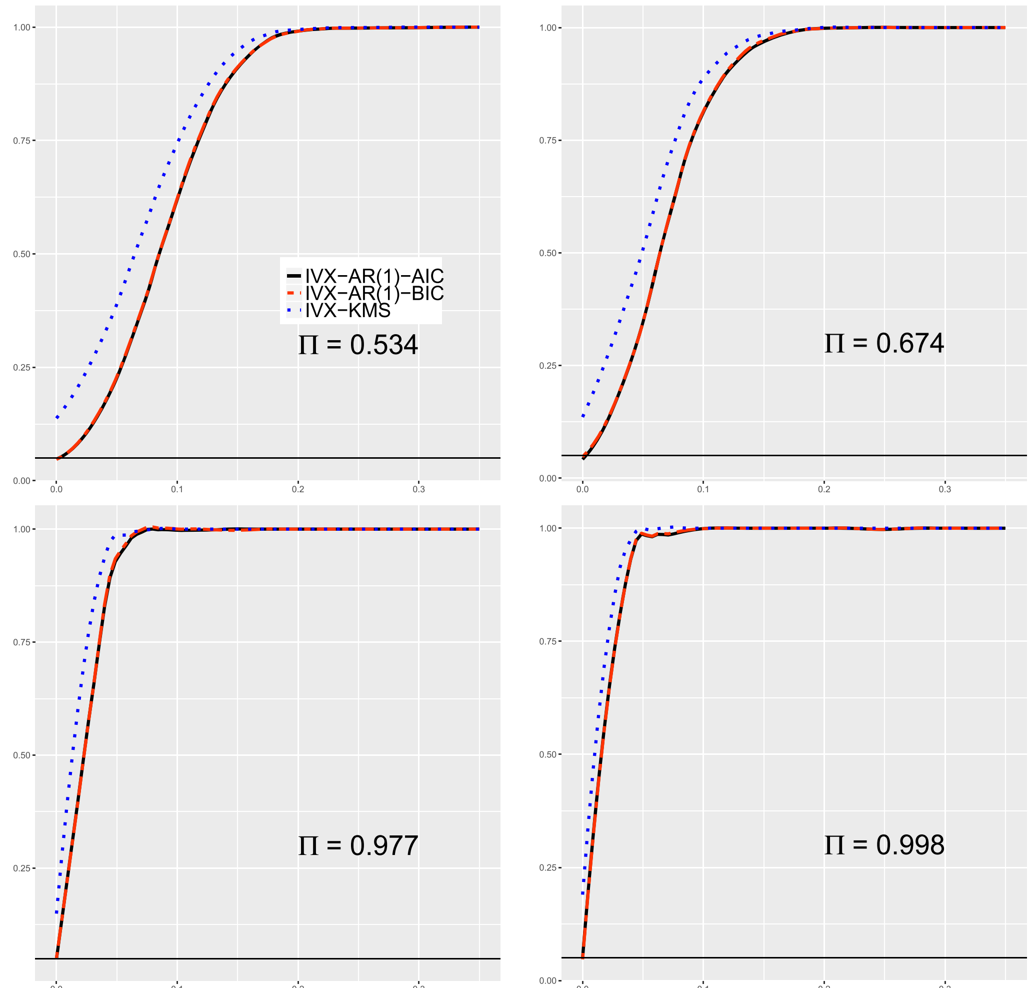


Figure 4. Power plots for multivariate regression with empirically relevant parameters. NOTE: This figure displays the rejection rate for tests of the null hypothesis $\beta = (0,0,0,0)^{\intercal}$, that is, all four coefficients in β equal to zero, as the true value for each regressor's slope coefficient (horizontal axis) increases. That is, for panel 1, $\beta = \left(\frac{2j}{T},0,0,0\right)^{\intercal}$, for panel 2, $\beta = \left(0,\frac{2j}{T},0,0\right)^{\intercal}$, for panel 3, $\beta = \left(0,0,\frac{2j}{T},0\right)^{\intercal}$, and for panel 4, $\beta = \left(0,0,0,\frac{2j}{T}\right)^{\intercal}$, where $j \in \{0,1,2,\ldots,39,40\}$. The joint Wald statistics are, respectively, calculated by the proposed IVX-AR(1) using AIC and BIC and IVX-KMS proposed by Kostakis, Magdalinos, and Stamatogiannis (2015). Values for the autocorrelation coefficients and the correlation structure of the residuals Σ are empirically relevant and obtained from the predictive regression with the quarterly returns of the HPI being the regressand and the quarterly implicit price deflator of GDP, quarterly effective federal funds rate, quarterly percent change of real disposable personal income and quarterly percentage change of GDP being the regressors. Specifically, in this figure, $\phi_1 = 0.517$, $\Pi = \mathrm{diag}(0.998, 0.977, 0.674, 0.534)$, and $(\psi_1, \psi_2, \psi_3, \psi_4) = (0.628, 0.174, -0.002, -0.116)$. The four panels correspond to the four regressors, with the black-solid curve and the red-dashed curve, respectively, denote the rejection rate of the Wald statistic by the proposed IVX-AR method using AIC and BIC under the 5% nominal size (the horizontal solid line), while the blue-dotted curve denotes the rejection rate of the Wald statistic by the IVX-KMS method. For each panel, the rejection rate is calculated with 10,000 repetitions and sample size $T = 200$.

IND: The industrial production index (Index 2012 = 100). It is an economic indicator that measures real output for all facilities located in the United States manufacturing, mining, and electric, and gas utilities (excluding those in U.S. territories).

INT: The effective federal funds rate. It is the interest rate at which depository institutions trade federal funds (balances held at Federal Reserve Banks) with each other overnight.

INV: The shares of the residential fixed investment in the gross domestic product. Gross private domestic investment is a critical component of gross domestic product as it provides an indicator of the future productive capacity of the economy. Residential investment represents expenditures on residential structures and residential equipment that is owned by landlords and rented to tenants.

MOG: 30-year mortgage rate. It represents contract interest rates on commitments for fixed-rate first mortgages.

RES: The total reserve balances maintained with the Federal Reserve banks.

UNE: The civilian unemployment rate. It represents the number of unemployed as a percentage of the labor force.

4.1. Testing the Unit Root

As discussed at the beginning, the IVX-AR method is superior to the IVX-KMS method in that it corrects the size distortions of the Wald statistic caused by the serial correlation in the error terms. Therefore, if our target is to examine the predictive ability of the ten regressors with respect to the growth rate of HPI, it should be rational to begin with investigating whether these regressors exhibit high degree of persistence and whether serial correlation and heteroscedasticity indeed exist in the error terms of the predictive regression. Following Kostakis,

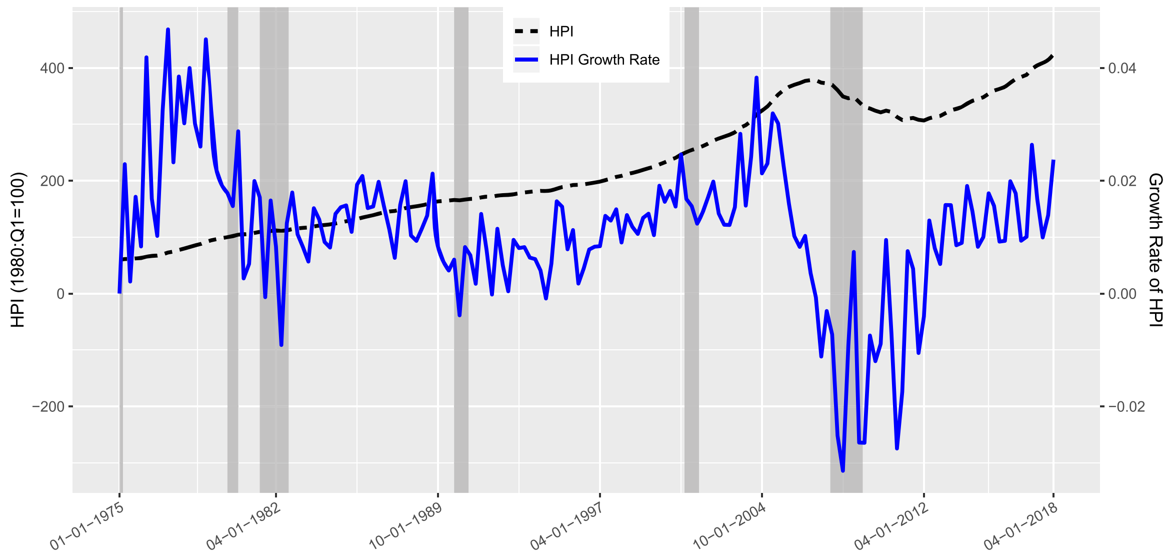


Figure 5. Housing price index and its quarterly growth rate. NOTE: This figure plots the quarterly housing price index (HPI, 1980:Q1 = 100) and its quarterly growth rate in the United States during 1975:Q1–2018:Q2. Recessions defined by the National Bureau of Economic Research (NBER) are shown as the shaded regions.

Table 5. Unit root tests for ten predictive variables.

	$\widehat{\Pi}_x$	ADF	DF-GLS	PP	KPSS
CPI	0.997	5.074	−1.927	−1.701	3.545***
DEF	0.998	2.032	−1.469	−2.216	3.523***
GDP	0.534	−2.292**	−2.469	−9.592***	1.960***
INC	0.674	−1.548	−1.839	−6.429***	0.266
IND	0.996	1.639	−2.241	−1.921	3.475***
INT	0.977	−1.151	−2.626	−2.945	2.475***
INV	0.984	−0.926	−3.259**	−2.282	0.619**
MOG	0.992	−0.875	−1.914	−2.727	2.770***
RES	1.001	0.498	−1.464	−1.280	1.514***
UNE	0.984	−0.942	−3.094**	−2.222	0.524**

NOTE: This table documents the results of the unit root tests for 10 fundamental macroeconomic variables. CPI = quarterly consumer price index for all urban consumers: all items less shelter (Index 1982–1984 = 100). DEF = quarterly implicit price deflator of the gross domestic product (Index 2012 = 100). GDP = quarterly percent change of the gross domestic product from the preceding period. INC = quarterly percent change of real disposable personal income from quarter one year ago. IND = quarterly industrial production index (Index 2012 = 100). INT = quarterly effective federal funds rate. INV = quarterly percent of residential fixed investment in the gross domestic product. MOG = quarterly 30-year mortgage rate. RES = quarterly growth of the total reserve balances maintained with Federal Reserve banks. UNE = quarterly civilian unemployment rate. All variables span from 1975:Q1 to 2018:Q2. $\widehat{\Pi}_x$ indicates the least square estimate of Π_x in the AR(1) process: $x_t = s + \Pi_x x_{t-1} + e_t$. ADF represents the statistic of the Augmented Dickey-Fuller test by Said and Dickey (1984) for the null that x_t has a unit root. DF-GLS represents the statistic from an ADF-type test by Elliott, Rothenberg, and Stock (1996) for the null that x_t has a unit root. PP refers to the statistic from the Phillips–Perron test by Phillips and Perron (1988) for the null that x_t has a unit root. KPSS refers to the statistic from Kwiatkowski, Phillips, Schmidt and Shin's (1992) unit root test for the null that x_t is stationary. For ADF and DF-GLS test statistics, the optimal length of lag is determined by the Bayesian information criterion (BIC). *, **, and ***, respectively, indicate rejection of the null at 10%, 5%, and 1%.

Magdalinos, and Stamatogiannis (2015), for each regressor, we first examine the least square estimate of the autoregressive parameter Π_x in $x_t = s + \Pi_x x_{t-1} + e_t$. As shown in the first column of Table 5, except GDP and INC, the least square estimates of the autoregressive coefficients for the rest regressors are close to 1, implying strong persistence within these series at a quarterly frequency. Then, we apply four widely used tests to formally check the existence of a unit root within these regressors: the augmented Dickey–Fuller (ADF) test (Said and Dickey 1984), the DF-GLS test (Elliott, Rothenberg, and Stock 1996), the Phillips–Perron (PP) test (Phillips and Perron 1988), and the Kwiatkowski, Phillips, Schmidt, and Shin (KPSS) test (Kwiatkowski et al. 1992). ADF, DF-GLS, and PP tests have the null hypothesis that a unit root exists, while the KPSS test has the null hypothesis of no unit root. For the ADF and DF-GLS test statistics, we determine the optimal length of lag through BIC. Table 5 shows that, for CPI, DEF, IND, INT, MOG, and RES, the four unit root tests exhibit consistent results and confirm the existence of a unit root. But for DEF, INC, INV, and UNE, the four tests give contradictory results. In summary, most of these variables are highly persistent with autoregressive coefficients very close to unity, which cast doubts on the validity of the classic t-test, as suggested by Cavanagh, Elliott, and Stock (1995). Under this circumstance, the IVX method is needed to check the reliability of these variables' predictive ability.

4.2. Testing the Serial Correlation and Heteroscedasticity

Next, we need to examine whether the error terms of the predictive regression, u_t, are serially correlated and exhibit evidence of conditional heteroscedasticity. For simplicity, we regress the growth rate of HPI on four regressors (CPI, DEF, INT, and RES) which are usually used to evaluate the monetary policy

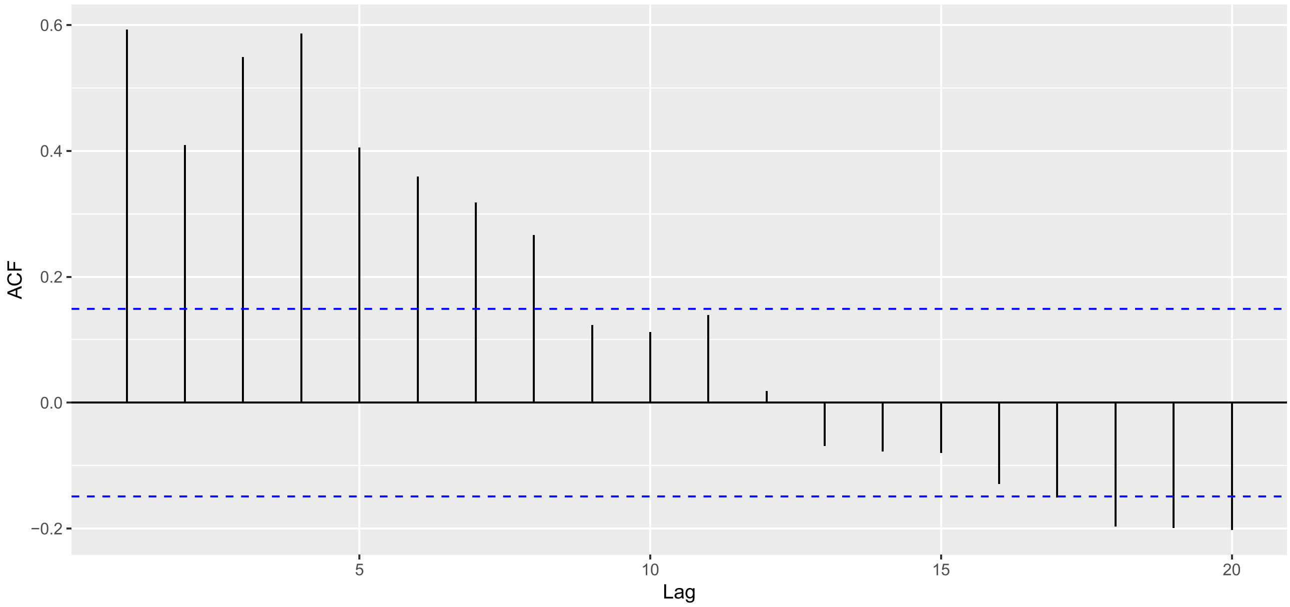


Figure 6. ACF plot of u_t in Model (2) by regressing the growth rate of HPI on the ten regressors. The two blue dashed lines denote the upper and lower bounds of the 95% confidence interval.

Table 6. Results of serial correlation tests for $\boldsymbol{u_t}$ in Equation (2).

Lag	Wald	LB	BP	BG
1	27.678***	61.855***	60.795***	1.299
2	48.649***	91.561***	89.822***	32.884**
3	47.696***	145.191***	141.919***	39.582***
4	126.944***	206.823***	201.439***	40.115***
5	125.417***	236.423***	229.854***	40.119***

NOTE: Tests of serial correlation for the estimates of u_t in Equation (2) by regressing the growth rate of HPI on four regressors: CPI, DEF, INT, and RES. Wald denotes the test for the null that all autocorrelation coefficients in u_t simultaneously equal zero. *LB* and *BP* denote the Ljung–Box test and the Box–Pierce test which examines the null hypothesis of independence in a given time series. BG denotes the Breusch–Godfrey test for the null that there is no serial correlation up to a specific order. For each test, the statistic is calculated for different numbers of lags (1, 2, 3, 4, and 5). *, **, and ***, respectively, indicate rejection of the null at 10%, 5%, and 1%.

Table 7. AIC, AICc, and BIC for AR(1)–AR(10).

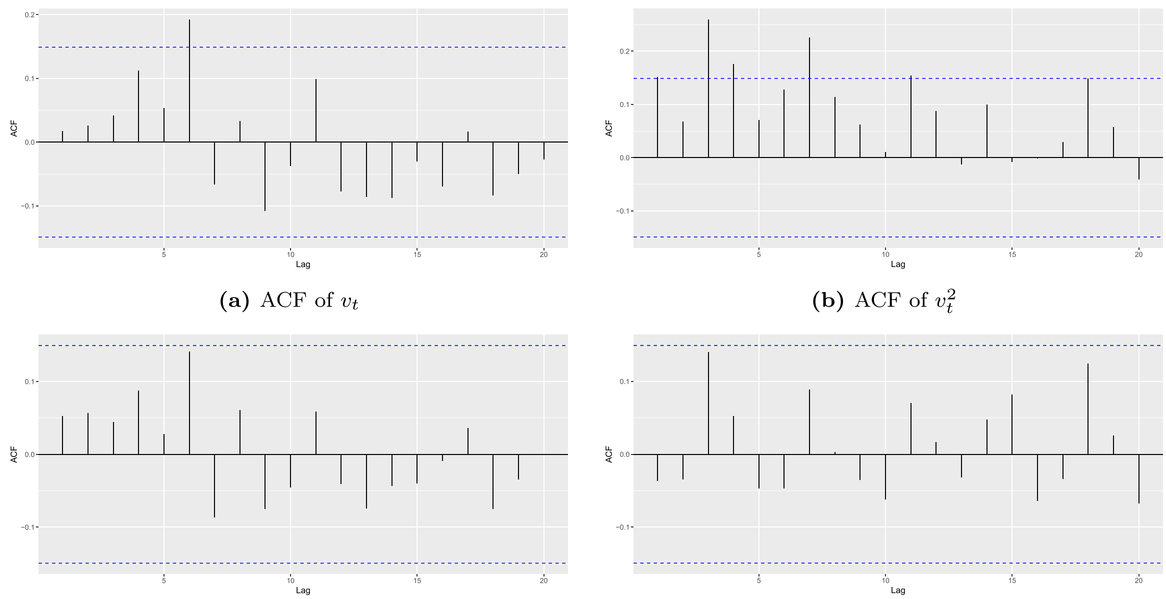
q	AIC	AICc	BIC
1	−1132.690	−1132.619	−1126.383
2	−1132.062	−1131.920	−1122.602
3	−1166.400	−1166.162	−1153.787
4	−1173.298	−1172.939	**−1157.532**
5	−1171.915	−1171.409	−1152.995
6	−1169.924	−1169.245	−1147.851
7	−1175.435	−1174.557	−1150.209
8	−1173.811	−1172.706	−1145.431
9	**−1187.549**	**−1186.191**	−1156.016
10	−1185.559	−1183.92	−1150.873

NOTE: This table displays the values of the Akaike information criterion (AIC), corrected Akaike information criterion (AICc) and Bayesian information criterion (BIC) for fitting $\hat{u}_t$ to AR(1)–AR(10). $\hat{u}_t$ is obtained by regressing the growth rate of HPI on CPI, DEF, INT, and RES. The optimal order is in bold.

and then draw the autocorrelation coefficient (ACF) plot of the estimated error terms in Figure 6. One can observe that the autocorrelation coefficients of the error terms are significantly different from zero until the 8th lag, suggesting the existence of serial correlation in u_t. We take one more step to check whether the error terms are serially correlated through four tests: the Wald test proposed in Theorem 3, the Ljung–Box test, the Box–Pierce test, and the Breusch–Godfrey test. Table 6 documents the test results and confirms that the error terms are indeed serially correlated. To eliminate such a serial correlation, we fit the estimated error terms $\hat{u}_t$ to an AR model and select the optimal order q by comparing the AIC, AICc, and BIC. Table 7 shows that the three criteria do not agree on the optimal order of AR. We ultimately choose $q = 4$ based on BIC, which prefers the more parsimonious model specification. By fitting $\hat{u}_t$ to AR(4), we obtain the error terms v_t, which are not serially correlated by both the serial correlation tests (results are available from the author upon request) and the ACF plot in Figure 7(a). However, the ACF plot of v_t^2, as displayed in Figure 7(b), suggests the existence of conditional heteroscedasticity. To capture such persistent volatility, we further fit the estimated $\hat{v}_t$ to a GARCH(1,1) model, as specified in Equation (4), and then examine the ACF plots of ϵ_t and ϵ_t^2. Panel(c) and Panel(d) in Figure 7 show that, after the AR(4)-GARCH(1,1) filtering, both ϵ_t and ϵ_t^2 become uncorrelated. In summary, the preliminary examinations confirm the existence of serial correlation and conditional heteroscedasticity, which motivate us to carry out the proposed IVX-AR method to investigate the predictive ability of these macroeconomic variables for the growth rate of HPI.

4.3. Univariate Regressions Based on Full Sample (1975:Q1–2018:Q2)

We firstly run univariate regressions to estimate the slope coefficients of the ten regressors and examine their respective predictive ability. In Table 1, $\widehat{\beta}_{\text{IVX-AR}}$ represents the estimate of the univariate regressions using IVX-AR and $W_{\beta,\text{IVX-AR}}$ denotes the corresponding Wald test statistic for the null that the slope coefficient in the predictive regression is zero. For



(a) ACF of v_t **(b)** ACF of v_t^2

(c) ACF of ϵ_t **(d)** ACF of ϵ_t^2

Figure 7. ACF plots for v_t, v_t^2, ϵ_t, and ϵ_t^2. NOTE: Panel(a) shows the ACF plot for v_t in Equation (3) by fitting an AR(4) model. Panel (b) shows the ACF plot for v_t^2. Panel (c) shows the ACF plot for ϵ_t in Equation (4) by fitting an GARCH(1,1) model. Panel (d) shows the ACF plot for ϵ_t^2. The two blue dashed lines denote the upper and lower bounds of the 95% confidence interval.

comparing purposes, in Table 1, we also report the OLS estimate ($\widehat{\beta}_{\text{OLS}}$) and its corresponding t-statistic (t_{OLS}), and the IVX-KMS estimate ($\widehat{\beta}_{\text{IVX-KMS}}$) and the corresponding Wald statistic ($W_{\beta,\text{IVX-KMS}}$). We additionally report the correlation coefficient θ between v_t and ϵ_t in Equations (3) and (4), the optimal choice of q determined by the Bayesian information criterion (BIC) in Equation (3), and the Wald statistics (W_ϕ) proposed in Theorem 3 which tests the null hypothesis that $H_0^{(1)} : \phi_1 = \cdots = \phi_q = 0$ versus $H_1^{(1)} : \phi_j \neq 0$ for some $j \in \{1, 2, \ldots, q\}$.

Panel 1 in Table 1 documents the results of the univariate predictive regressions based on observations between 1975:Q1 and 2018:Q2. Using OLS and IVX-KMS, we find that all ten macroeconomic variables have the strong predictive ability on the housing price because the null of zero slope coefficient can be rejected at the 5% significance level in all 10 univariate regressions. This is expected given the high order serial correlation exhibited by the residuals of the univariate regressions, as shown by the optimal choice of q and the significant Wald statistics (W_ϕ). On the contrary, the IVX-AR based regressions, which account for the serial correlation in the residuals, suggest that only INV and UNE preserve their significant predictive ability at the 1% level, while the null of zero slope coefficient cannot be rejected at any conventional significance levels for the rest regressors. In other words, even though the housing price is sensitive to various economic indicators and the monetary policy, the testing results by IVX-AR show that most economic factors alone are hardly useful to predict its growth. These results can be explained by the complexity in the housing market and the incomplete information set used by the univariate models, as suggested by Cochrane (2011). To mitigate the drawbacks of the univariate analysis, we next implement multivariate regressions to investigate the joint significance of multiple macroeconomic variables.

4.4. Multivariate Regressions Based on Full Sample (1975:Q1–2018:Q2)

In the multivariate analysis, we specifically consider the following five combinations:

1. INV + UNE;
2. INV + UNE + IND + GDP + INC;
3. CPI + DEF + INT + RES;
4. CPI + INT + MOG;
5. A "Kitchen Sink" which includes all ten aforementioned variables.

The first panel in Table 8 demonstrates the results of the multivariate predictive regressions for observations between 1975:Q1 and 2018:Q2. The first combination considers the joint significance of INV and UNE because Table 1 shows that both exhibit significant predictive ability by IVX-AR. In the second combination, besides INV and UNE, we additionally add the GDP growth (GDP), the real per capita disposable income (INC), and the industrial product index (IND), which are all main macroeconomic indicators and closely tracked by investors and policymakers. Kallberg, Liu, and Pasquariello (2014) categorized these variables as "underlying systematic real and financial factors" and attribute the comovement among housing markets to these factors. The third combination mainly concentrates on the monetary policy, which is measured by the inflation rate, the price deflator of GDP, the effective federal funds rate, and the total reserves. Del Negro and Otrok (2007) used similar variables and a VAR model to investigate how housing price responds to expansionary monetary policy. The fourth combination measures the cost of housing investment, which tends to be affected by low interest rates (Fu 2007 and Shiller 2007). In the last combination, we follow Welch and Goyal (2008) and consider a "kitchen sink" which considers the joint significance of all above variables. We report the results obtained by IVX-AR

Table 8. Results of predictive regressions with multiple regressors.

	(1) CPI	(2) DEF	(3) GDP	(4) INC	(5) IND	(6) INT	(7) INV	(8) MOG	(9) RES	(10) UNE	(11) Joint Wald	(12) q	(13) W_ϕ
Panel 1: 1975:Q1–2018:Q2													
Combination 1													
IVX-AR	–	–	–	–	–	–	0.0061***	–	–	−0.0010	13.561***	1	13.644***
IVX-KMS	–	–	–	–	–	–	0.0080***	–	–	0.0002	93.329***	–	–
Combination 2													
IVX-AR	–	–	−0.0003	−0.0005	−0.0002	–	0.0060**	–	–	−0.0014	16.537***	1	10.022***
IVX-KMS	–	–	0.0004	0.0000	−0.0001	–	0.0056***	–	–	−0.0009	106.216***	–	–
Combination 3													
IVX-AR	−0.0001	−0.0005	–	–	–	−0.0005	–	–	0.0017	–	4.093	4	128.583***
IVX-KMS	−0.0005	0.0009	–	–	–	−0.0009*	–	–	0.0000	–	24.469***	–	–
Combination 4													
IVX-AR	−0.0001	–	–	–	–	−0.0008	–	0.0003	–	–	4.748	1	26.777**
IVX-KMS	−0.0002***	–	–	–	–	0.0009	–	−0.0028***	–	–	22.105***	–	–
Combination 5													
IVX-AR	0.0002	−0.0014	−0.0003	−0.0005	0.0007	−0.0025***	0.0082***	0.0035**	0.0056***	−0.0030*	61.612***	1	3.395*
IVX-KMS	0.0003	−0.0015	−0.0001	0.0001	0.0004	−0.0024***	0.0080***	0.0029***	0.0050***	−0.0029***	219.138***	–	–
Panel 2: 2000:Q1–2018:Q2													
Combination 1													
IVX-AR	–	–	–	–	–	–	0.0069*	–	–	−0.0018	8.024**	1	13.919***
IVX-KMS	–	–	–	–	–	–	0.0049***	–	–	−0.0026**	38.794***	–	
Combination 2													
IVX-AR	–	–	−0.0002	0.0000	0.0005	–	0.0086**	–	–	−0.0001	8.402	1	9.113***
IVX-KMS	–	–	0.0019	0.0004	−0.0006	–	0.0013	–	–	−0.0042	47.212***	–	
Combination 3													
IVX-AR	−0.0004	−0.0001	–	–	–	0.0034	–	–	0.0050	–	2.646	1	23.650***
IVX-KMS	−0.0017***	0.0032***	–	–	–	0.0003	–	–	0.0014	–	14.334***	–	
Combination 4													
IVX-AR	0.0001	–	–	–	–	−0.0007	–	0.0091**	–	–	4.496	1	19.298***
IVX-KMS	−0.0007***	–	–	–	–	0.0035***	–	−0.0130***	–	–	17.164***	–	
Combination 5													
IVX-AR	0.0006	−0.0041***	0.0003	−0.0007	0.0014*	−0.0019	0.0124***	0.0036	0.0138***	−0.0045**	114.203***	2	8.396**
IVX-KMS	0.0005	−0.0040***	0.0005	−0.0008	0.0011*	−0.0015	0.0109***	−0.0002	0.0125***	−0.0051***	232.981***	–	

NOTE: This table documents the results of multivariate predictive regressions for observations from the full sample period (Panel 1: 1975:Q1–2018:Q2) and the subperiod (Panel 2: 2000:Q1–2018:Q2). For both panels, the dependent variable is the quarterly growth rate of housing prices in the United States. CPI = quarterly consumer price index for all urban consumers: all items less shelter (Index 1982–1984 = 100). DEF = quarterly implicit price deflator of the gross domestic product (Index 2012 = 100). GDP = quarterly percent change of the gross domestic product from the preceding period. INC = quarterly percent change of real disposable personal income from quarter one year ago. IND = quarterly industrial production index (Index 2012 = 100). INT = quarterly effective federal funds rate. INV = quarterly percent of residential fixed investment in the gross domestic product. MOG = quarterly 30-year mortgage rate. RES = quarterly growth of the total reserve balances maintained with Federal Reserve banks. UNE = quarterly civilian unemployment rate. IVX-AR denotes the proposed IVX-AR method. IVX-KMS denotes the IVX method by Kostakis, Magdalinos, and Stamatogiannis (2015). For each combination, values in columns (1)–(10) are the slope coefficients, respectively, estimated by IVX-AR and IVX-KMS. The significance of each estimate is determined by the statistic obtained from the Wald test under the null that the slope coefficient equals to zero. Column (11) contains the joint Wald statistic obtained by IVX-AR and IVX-KMS for the null that all slope coefficients simultaneously equal to zero. q in column (12) denotes the optimal order of AR for the residuals of Equation (2) selected by the Bayesian information criterion (BIC). W_ϕ in column (13) denotes the Wald statistic defined in Equation (12) and tests the null that $H_0 : \phi_1 = \cdots = \phi_q = 0$ versus $H_1 : \phi_j \neq 0$ for some $j \in \{1, 2, \ldots, q\}$. *, **, and ***, respectively, indicate rejection of the null at 10%, 5%, and 1%.

and IVX-KMS in Panel 1 of Table 8. For combinations 1 and 2, both methods give consistent inferences: the two combinations are jointly significant at the 1% level, and INV is still individually significant at the 1% level. However, UNE is no longer significant at any conventional level as it in the univariate analysis. The IVX-AR and IVX-KMS methods also give similar results in the "kitchen sink" regression, as displayed by combination 5 in Panel 1 of Table 8. In this combination, the ten variables are jointly significant at the 1% level, and the five variables—INT, INV, MOG, RES, and UNE—become individually significant at least at the 5% level. However, these two methods display contradictory results in combinations 3 and 4. In both cases, neither the joint Wald statistic nor each variable's individual Wald statistic is significant at any conventional level by IVX-AR. On the contrary, using IVX-KMS, we find that not only are both cases jointly significant at the 1% level but also some variables are individually significant (INT in combination 3, CPI and MOG in combination 4). This can be explained by the high order of autocorrelation ($q = 4$) in $\check{u}_t$, as displayed by column (12). The Wald statistics, W_ϕ, in column (13) further confirm the existence of the high order serial correlation in the two combinations because both Wald statistics are significant at the 1% level. IVX-KMS does not account for such a high order serial correlation and suffers from extreme oversizing, which is in line with our simulation results (Table 3) for the size properties of IVX-KMS. For the other three combinations, because the order of autocorrelation is relatively lower ($q = 1$), it is still possible that these two methods give similar results, even though the Wald statistics by IVX-KMS are remarkably higher than that of IVX-AR.

4.5. *Predictability of HPI Based on Subsample (2000:Q1–2018:Q2)*

We finally apply the same procedures to examine a subsample that spans from 2000:Q1 to 2018:Q2. This subperiod is selected because, as Figure 5 shows, it witnesses the greatest fluctuation

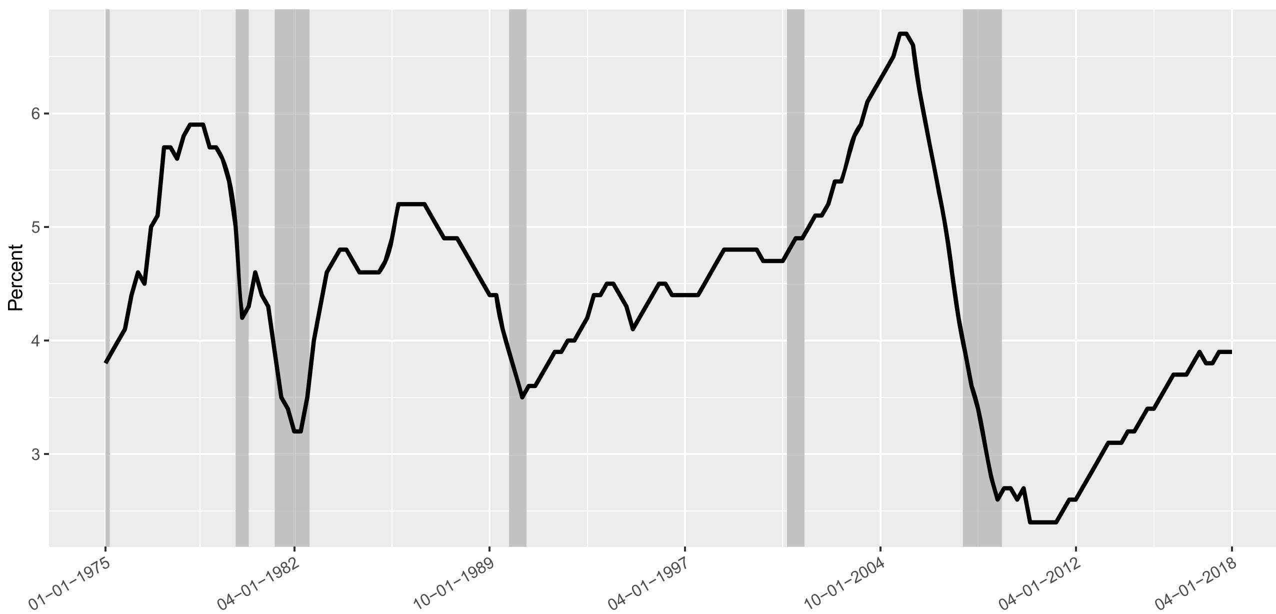


Figure 8. Percentage of residential investment in GDP. This figure plots the residential investment as a percent of GDP in the United States during 1975:Q1–2018:Q2. Recessions defined by the NBER are shown as the shaded regions.

in the housing market since 1975. The growth of HPI began to accelerate after 2000 and achieved its historical high on the eve of the 2007 subprime mortgage crisis. Shiller (2007) claimed that the boom in the housing market during this period differs from the prior ones in that "it is much more of a national, rather than regional, event" and the magnitude of the increase is "unprecedented." After a long time decline during 2008–2011, the housing price resumed to a positive growth rate, and HPI reached another historical high in 2018. Therefore, if some regressors exhibit predictive ability on the housing price, this subperiod provides an ideal scenario to test their predictive ability.

The second panel in Table 1 reports the results of the univariate analysis for observations during this subperiod.[5] In Panel 2, OLS and IVX-KMS have consensus on the significant predictive ability demonstrated by CPI, DEF, GDP, INC, INV, and UNE. However, compared with Panel 1, only the coefficients associated with CPI, GDP, INV, and UNE remain significant at the 1% level. IVX-AR agrees with IVX-KMS only on INV and UNE. Besides these two regressors, the slope coefficient estimate for MOG by IVX-AR becomes significant at the 5% level during the subperiod, even though that of OLS and IVX-KMS are insignificant at any conventional level.

Moreover, we check the joint significance of the same combinations and report the results in the second panel of Table 8. Comparing Panel 2 with Panel 1, we find two main differences. First, the two methods show contradictory results in combination 2: during the subperiod (Panel 2), the joint Wald statistic by IVX-AR is no longer significant at any conventional levels, while that by IVX-KMS is still significant at the 1% level, even though none of the five regressors is individually significant. Considering that the optimal choice of q for $\check{u}_t$ increases to 3 and the corresponding Wald statistic in column (13) is significant at the 1% level during the subperiod, we argue that the joint significance of the five regressors in combination 2 is driven by the strong serial correlation in $\check{u}_t$. When the correlation is appropriately modeled, as IVX-AR does, the significant joint predictive ability demonstrated by these regressors disappears. Second, in combination 5, even though both methods agree on the joint predictive ability of the ten regressors at the 1% level, during the subperiod, INT and MOG are no longer individually significant as they are in Panel 1, while DEF and IND become significant at the 1% and 10% levels, respectively.

Taken together, regardless of the univariate or multivariate analysis, only INV uniformly displays significant predictive ability on the growth rate of the housing price during the full sample period and the subperiod. The residential investment represents all economic activities related to housing structures and usually includes three parts: construction of new single-family and multifamily structures, residential remodeling and production of manufactured homes, and brokers fees. Therefore, INV serves as an important channel of housing's contribution to GDP. From 1980 to 2005, the residential investment increased from 333 billion dollars to 873 billion dollars, but then sharply dropped to 382 billion dollars in 2010 (U.S. Bureau of Economic Analysis). Figure 8 shows the residential investment as a percent of GDP during 1975:Q1–2018:Q2. As Leamer (2007) argued, the residential investment as a percent of GDP has had an obvious peak before all five recessions since 1975, and the end of a recession is usually accompanied by a sharp return. However, Shiller (2007) claimed that "the relation between housing investment and the business cycle may be changing" because the recession in 2001 does not show a substantial drop in residential investment compared with prior cases. The latest recession in 2007 again confirms Leamer's (2007) findings by showing another substantial drop in housing investment, and this time it takes a relatively longer time for housing investment in GDP to bounce back. Given this evidence, the predictive ability of INV should

[5]We implement the same procedures to examine the serial correlation in the residuals for the subsample and confirm that the serial correlation still exists. The serial correlation test results for the subsample are available upon request.

be expected. For the other nine variables, even though some of them are individually significant in either the univariate or the multivariate analysis, we do not observe consistent evidence about their predictive ability. For these regressors, the significance of their slope coefficients might be purely driven by the serial correlation in the residuals of the predictive regression rather than their predictive ability. Therefore, predictive ability detected by IVX-KMS, which fails to account for the serial correlation in the residuals, deserves more careful investigations.

5. Conclusion

In this study, we test the predictability of the growth rate of HPI by proposing an IVX-AR Wald statistic that accounts for the serial correlation and heteroscedasticity in the error terms of the linear predictive regression model. We develop two algorithms to facilitate the estimation and establish two Wald statistics to, respectively, test the predictability of regressors and the existence of serial correlation in the error terms. We conduct exhaustive Monte Carlo simulations to investigate the finite-sample performance of the testing procedures and find that, if error terms in the predictive regression are indeed serially correlated, the proposed IVX-AR method exhibits excellent size control regardless of the degree of serial correlation in the error terms and the persistence of the predictive variables, while the IVX-KMS method proposed by Kostakis, Magdalinos, and Stamatogiannis (2015) suffers severe size distortions. Using the quarterly growth rate of housing price in the United States and ten common macroeconomic variables, we find that both IVX-AR and IVX-KMS agree that the percentage of residential fixed investment in GDP is a robust predictor of the growth rate of housing price. In contrast, no strong or consistent evidence of predictability is observed by IVX-AR when considering the other nine macroeconomic variables as predictors. Results from several serial correlation tests imply that the predictive ability falsely detected by IVX-KMS is likely to be driven by the highly correlated error terms in the predictive regressions, and thus becomes insignificant when IVX-AR is implemented. Because of its generalizability, IVX-AR could be a useful and reliable tool for empirical researchers to test for the predictability of returns of other assets such as stocks and commodities.

Supplementary Materials

The supplementary materials for this article contain an Appendix and the replication files. The Appendix provides supplemental lemmas and proofs of the main results in the main text. The replication files provide codes and datasets used in the simulations and empirical study.

Acknowledgments

We thank Professor Heping Zhang (editor), one associate editor, and two anonymous reviewers for helpful suggestions.

Funding

Yang's research is partly supported by Humanity and Social Science Youth Foundation of Ministry of Education of China (19C10421028), the National Nature Science Foundation of China (71401066), and the Major Program of the National Social Science Foundation of China (17ZDA073). Long's research is partly supported by the Kurzius Family fund and the Carol Lavin Bernick Faculty Grant Program at Tulane University. Peng's research is partly supported by the Simons Foundation. Cai's research is partly supported by the NSFC grant (71631004) (Key Project).

References

Amihud, Y., and Hurvich, C. (2004), "Predictive Regressions: A Reduced-Bias Estimation Method," *Journal of Financial and Quantitative Analysis*, 39, 813–841. [1599]

Ball, R., and Kothari, S. P. (1989), "Nonstationary Expected Returns: Implications for Tests of Market Efficiency and Serial Correlation in Returns," *Journal of Financial Economics*, 25, 51–74. [1599]

Cai, Z., and Wang, Y. (2014), "Testing Predictive Regression Models With Nonstationary Regressors," *Journal of Econometrics*, 178, 4–14. [1599]

Case, K., and Shiller, R. (1989), "The Efficiency of the Market for Single-Family Homes," *American Economic Review*, 79, 125–137. [1599,1611]

Campbell, J. Y., and Yogo, M. (2006), "Efficient Tests of Stock Return Predictability," *Journal of Financial Economics*, 81, 27–60. [1599]

Cavanagh, C., Elliott, G., and Stock, J. (1995), "Inference in Models With Nearly Integrated Regressors," *Econometric Theory*, 11, 1131–1147. [1613]

Cochrane, J. (2011), "Presidential Address: Discount Rates," *The Journal of Finance*, 66, 1047–1108. [1615]

Del Negro, M., and Otrok, C. (2007), "Luftballons: Monetary Policy and the House Price Boom Cross U.S. States," *Journal of Monetary Economics*, 54, 1962–1985. [1598,1611,1615]

Demetrescu, M., and Rodrigues, P. (2016), "Residual-Augmented IVX Predictive Regression," Working Papers w201605, Banco de Portugal, Economics and Research Department. [1599]

Elliott, G., Rothenberg, T., and Stock, J. (1996), "Efficient Tests for an Autoregressive Unit Root," *Econometrica*, 64, 813–836. [1613]

Elliott, G., and Stock, J. (1994), "Inference in Time Series Regression When the Order of Integration of a Regressor Is Unknown," *Econometric Theory*, 10, 672–700. [1599]

Fu, D. (2007), "National, Regional and Metro-Specific Factors of the U.S. Housing Market," Federal Reserve Bank of Dallas, Working Papers 0707. [1615]

Getmansky, M., Lo, A. W., and Makarov, I. (2004), "An Econometric Model of Serial Correlation and Illiquidity in Hedge Fund Returns," *Journal of Financial Economics*, 74, 529–609. [1599]

Hall, P., and Yao, Q. (2003), "Data Tilting for Time Series," *Journal of the Royal Statistical Society*, Series B, 65, 425–442. [1600]

Hill, J., Li, D., and Peng, L. (2016), "Uniform Interval Estimation for an AR(1) Process With AR(p) Errors," *Statistica Sinica*, 26, 119–136. [1600]

Hjalmarsson, E. (2011), "New Methods for Inference in Long-Horizon Regressions," *Journal of Financial and Quantitative Analysis*, 46, 815–839. [1599]

Jansson, M., and Moreira, M. (2006), "Optimal Inference in Regression Models With Nearly Integrated Regressors," *Econometrica*, 74, 681–714. [1599]

Joe, H. (1997), *Dependence Modeling With Copulas*, Boca Raton, FL: Chapman & Hall/CRC. [1602]

Kallberg, J., Liu, C., and Pasquariello, P. (2014), "On the Price Comovement of U.S. Residential Real Estate Markets," *Real Estate Economics*, 42, 71–108. [1598,1611,1615]

Kostakis, A., Magdalinos, T., and Stamatogiannis, M. (2015), "Robust Econometric Inference for Stock Return Predictability," *The Review of Financial Studies*, 28, 1506–1553. [1599,1600,1601,1603,1605,1606,1607,1609,1610,1611,1612,1613,1616,1618]

Kwiatkowski, D., Phillips, P. C. B., Schmidt, P., and Shin, Y. (1992), "Testing the Null Hypothesis of Stationarity Against the Alternative of a Unit Root," *Journal of Econometrics*, 54, 159–178. [1613]

Leamer, E. (2007), "Housing and Business Cycle," *Paper Presented at the Federal Reserve Bank of Kansas City Symposium "Housing, Housing Finance, and Monetary Policy"*, Jackson Hole, Wyoming, August 31, 2007. [1617]

Lee, J. (2016), "Predictive Quantile Regression With Persistent Covariates: IVX-QR Approach," *Journal of Econometrics*, 192, 105–118. [1599]

Lewellen, L. (2004), "Predicting Return With Financial Ratios," *Journal of Financial Economics*, 74, 209–235. [1599]

Liu, J., Chen, R., and Yao, Q. (2010), "Nonparametric Transfer Function Models," *Journal of Econometrics*, 157, 151–164. [1600]

Liu, X., Yang, B., Cai, Z., and Peng, L. (2019), "A Unified Test for Predictability of Asset Returns Regardless of Properties of Predicting Variables," *Journal of Econometrics*, 208, 141–159. [1599]

Magdalinos, T., and Phillips, P. C. B. (2009), "Limit Theory for Cointegrated Systems With Moderately Integrated and Moderately Explosive Regressors," *Econometric Theory*, 25, 482–526. [1599]

Phillips, P. C. B., and Lee, J. (2013), "Predictive Regression Under Various Degrees of Persistence and Robust Long-Horizon Regression," *Journal of Econometrics*, 177, 250–264. [1599]

——— (2016), "Robust Econometric Inference With Mixed Integrated and Mildly Explosive Regressors," *Journal of Econometrics*, 192, 433–450. [1599,1600,1602]

Phillips, P. C. B., and Magdalinos, T. (2009), "Econometric Inference in the Vicinity of Unity," CoFie Working Paper No. 7, Singapore Management University. [1602]

Phillips, P. C. B., and Perron, P. (1988), "Testing for a Unit Root in Time Series Regression," *Biometrika*, 75, 335–346. [1613]

Said, S., and Dickey, D. (1984), "Testing for Unit Roots in Autoregressive-Moving Average Models of Unknown Order," *Biometrika*, 71, 599–607. [1613]

Shiller, R. (2007), "Understanding Recent Trends in House Prices and Home Ownership," NBER Working Paper No. 13553. [1598,1611,1615, 1617]

Stambaugh, R. (1999), "Predictive Regressions," *Journal of Financial Economics*, 54, 375–421. [1599]

Tsay, R. (2010), *Analysis of Financial Time Series* (3rd ed.), Hoboken, NJ: Wiley. [1600]

Welch, I., and Goyal, A. (2008), "A Comprehensive Look at the Empirical Performance of Equity Premium Prediction," *Review of Financial Studies*, 21, 1455–1508. [1598,1615]

Xiao, Z., Linton, O., Carroll, R., and Mammen, E. (2003), "More Efficient Local Polynomial Estimation in Nonparametric Regression With Autocorrelated Errors," *Journal of the American Statistical Association*, 98, 980–992. [1600]

Zhu, F., Cai, Z., and Peng, L. (2014), "Predictive Regressions for Macroeconomic Data," *Annals of Applied Statistics*, 8, 577–594. [1599]