On Robust Inference in Time Series Regression

Richard T. Baillie
Michigan State University
King's College, University of London

Francis X. Diebold
University of Pennsylvania
and NBER

George Kapetanios
King's College, University of London

Kun Ho Kim
Yeshiva University, New York

March 24, 2022

Abstract: Least squares regression with heteroskedasticity and autocorrelation consistent (HAC) standard errors has proved very useful in cross section environments. However, several major difficulties, which are generally overlooked, must be confronted when transferring the HAC estimation technology to time series environments. First, most economic time series have strong autocorrelation, which renders HAC regression parameter estimates highly inefficient. Second, strong autocorrelation similarly renders HAC conditional predictions highly inefficient. Finally, the structure of most popular HAC estimators is ill-suited to capture the autoregressive autocorrelation typically present in economic time series, which produces large size distortions and reduced power in hypothesis testing, in all but the largest sample sizes. We show that all three problems are largely avoided by the use of a simple dynamic regression (DynReg), which is easily implemented and also avoids possible problems concerning strong exogeneity. We demonstrate the advantages of DynReg with detailed simulations covering a range of practical issues.

Key Words: Serial correlation, heteroskedasticity and autocorrelation consistent (HAC) regression, dynamic regression

JEL Codes: C13, C22, C31

Acknowledgments: For outstanding research assistance we thank Aaron Mora Melendez. For discussions and/or comments we thank Daniel Lewis, Peter Schmidt, Tim Volgelsang, and Jeff Wooldridge. More generally we are also grateful to seminar participants at Michigan State University. All remaining errors and misunderstandings, however, are ours alone.

Contact: baillie@msu.edu, fdiebold@sas.upenn.edu, george.kapetanios@kcl.ac.uk, kunho.kim@yu.edu

1 Introduction

For nearly a century, time-series regression with heteroskedastic and/or autocorrelated disturbances has featured prominently in empirical economics research. Indeed a widely-used estimation approach in recent decades is Ordinary Least Squares (OLS) with standard errors adjusted to achieve consistency even with possible heteroskedasticity and/or serial correlation. This is commonly called HAC (Heteroskedasticity and Autocorrelation Consistent) regression, and we denote it by HAC.

In many cross-sectional situations, HAC regression is appropriate and has been successful. In cross-sectional environments, sample sizes are typically large and often huge, little or no information is available as to the form of any possible heteroskedasticity, and serial correlation is irrelevant. In such environments HAC regression is justly emphasized (e.g. Angrist and Pischke (2008)).

We will show that, in sharp contrast, HAC regression is a much less satisfactory approach in time series, for three reasons:

(1) OLS parameter estimates, and hence HAC parameter estimates, are highly sub-optimal (inefficient) in the presence of strong serial correlation, which is typical of economic time series.

(2) HAC regression discards valuable predictive information in serially-correlated disturbances and hence produces sub-optimal (inefficient) forecasts. This is true even when one ignores parameter estimation uncertainty, and it is potentially worsened – via the inefficiency of OLS – when one accounts for parameter estimation uncertainty.

(3) HAC inference is subject to significant size distortions in all but the largest samples.

The standard (and in our view feeble) refutations of points (1) and (2) are that (1) is inconsequential in practice if sample size is large enough (because OLS retains consistency even if it loses efficiency), and that (2) is inconsequential if focus does not center on prediction (obviously). But even if points (1) and (2) above are ignored or deemed inconsequential, one must still confront (3).

Let us elaborate on point (3). HAC adjustments are based on estimates of the long-run variance (that is, the spectrum at frequency zero) of the product of regressors and disturbances,

$$\Omega = \sum_{\tau=-\infty}^{\infty} \Gamma(\tau),$$

where

$$\Gamma(\tau) = cov(x_t u_t, x_{t-\tau} u_{t-\tau}), \quad \tau = 0, \pm 1, ...$$

The most common HAC estimation approach follows Newey and West (1987) in using a lag-window estimator of Ω, necessitating choice of a truncation lag. The truncation lag is implicitly restricted to low values, and by the need to use an asymptotic normal approximation, at the expense of incorrectly sized tests, as stressed by Müller (2014).

The size distortions associated with tests based on traditional Newey-West-style HAC estimators, such as Andrews (1991), are well known. More recent work proposes HAC procedures that yield more accurately sized tests. One important strand of literature stays within the Newey-West framework but uses very long truncation lags (Kiefer et al., 2000; Kiefer and Vogelsang, 2002), and another important strand moves to a different framework involving cosine transformations (Müller, 2014; Lazarus et al., 2018). Even the newer HAC procedures, however, continue to produce poorly-sized tests in typical economic time series environments (small/moderate sample size, strong serial correlation).

In this paper we take a different approach, related to, but distinct from, classic but underappreciated work of Den Haan and Levin (1997, 1998, 2000). We show that an alternative and very simple dynamic regression (DynReg) procedure simultaneously solves problems (1), (2), and (3). We proceed as follows. In section 2 we introduce and compare the HAC and dynamic regression frameworks. In section 3 we characterize the loss of both estimation efficiency and predictive accuracy from HAC regression, and we compare DynReg. In section 4 we quantify the size distortions and power losses associated with HAC-based hypothesis tests, and we again compare DynReg. We consider variations on the basic theme in section 5, and we conclude in section 6.

2 The HAC and Dynamic Regression Frameworks

Consider the data-generating process (DGP)

$$y_t = x_t' \beta + u_t, \tag{1}$$

$t = 1, 2, ...T$, where β is a k-vector of parameters, where the k-vector x_t and the scalar u_t are covariance-stationary processes with $E(x_t u_t) = 0$.

More precisely $\{u_t\}$ is a covariance stationary scalar process with Wold representation

$$u_t = \sum_{i=0}^{\infty} \psi_i \varepsilon_{t-i}, \tag{2}$$

where $\{\psi_i\}$ is a square-summable sequence of non-stochastic scalars with $\psi_0 = 1$, $\{\varepsilon_t\}$ is an ergodic process such that $E\left(\varepsilon_t \mid Z_{t-1}^{\varepsilon}\right) = 0$ a.s., $E\left(\varepsilon_t^2 \mid Z_{t-1}^{\varepsilon}\right) = \sigma^2$ a.s. with $0 < \sigma^2 < \infty$, $sup_t(|\varepsilon_t|^4) < \infty$, and Z_{t-1}^{ε} is the sigma field generated by $\{\varepsilon_s; s \leq t\}$. The conditions on (2) allow u to be both serially correlated and conditionally heteroskedastic. In what follows we will primarily emphasize serial correlation rather than heteroskedasticty, because serial correlation is the unique feature of time series relative to cross sections.[1]

In parallel, x_t is a covariance stationary k-dimensional vector process with Wold representation

$$x_t = \sum_{i=0}^{\infty} \Psi_i w_{t-i}, \tag{3}$$

and corresponding VAR representation

$$x_t = \sum_{i=1}^{\infty} \Pi_i x_{t-i} + w_t, \tag{4}$$

where $\{\Psi_i\}$ and $\{\Pi_i\}$ are square-summable sequences of nonstochastic $k \times k$ matrices with $\Psi_0 = I$, $\{w_t\}$ is an ergodic k-dimensional process such that $E\left(w_t \mid Z_{t-1}^{w}\right) = 0$ almost surely, $E\left(w_t w_t' \mid Z_{t-1}^{w}\right) = \Omega_w$ almost surely with $|\Omega_w| > 0$ and $||\Omega_w|| < \infty$, $sup_t(||w_t||^4) < \infty$, and Z_{t-1}^{w} is the sigma field generated by $\{w_s; s \leq t\}$.

2.1 HAC Regression

The OLS estimator of β satisfies

$$T^{1/2}(\hat{\beta}_{OLS} - \beta) \rightarrow N(0, M),$$

where

$$M = Q^{-1} \Omega Q^{-1}$$

[1]Interestingly, however, cross sections do sometimes have a spatial dimension and therefore a natural ordering in space if not in time, and spatial correlation has recently begun to receive attention from a HAC estimation perspective, as in Müller and Watson (2021). Spatial HAC estimation is, however, beyond the scope of this paper.

$$Q = plim\left(T^{-1}\sum_{t=1}^{T} x_t x_t'\right)$$

$$\Omega = \sum_{\tau=-\infty}^{\infty} \Gamma(\tau)$$

$$\Gamma(\tau) = cov(x_t u_t, x_{t-\tau} u_{t-\tau}), \quad \tau = 0, \pm 1, ...$$

The key object in M is Ω, the spectrum of xu at frequency zero. HAC regression delivers consistent M estimation, and hence asymptotically valid inference, via consistent Ω estimation. That is, HAC uses $\widehat{M} = Q^{-1}\widehat{\Omega}Q^{-1}$, where $\widehat{\Omega}$ is a consistent estimator of Ω.

A large literature on consistent estimation of Ω traces at least to Hansen and Hodrick (1980). The most popular approach, by far, is due to Newey and West (1987), who propose lag-window estimation with linearly-decreasing (Bartlett) lag window:

$$\widehat{\Omega} = \left(\frac{1}{T}\sum_{t=1}^{T}(x_t x_t')\hat{u}_t^2 + \sum_{\tau=1}^{h}\left(1 - \frac{\tau}{h+1}\right)(\widehat{\mathbf{\Gamma}}_\tau + \widehat{\mathbf{\Gamma}}_{-\tau})\right), \tag{5}$$

where

$$\hat{\mathbf{\Gamma}}_\tau = \frac{1}{T}\sum_{t=1}^{T} \hat{u}_t x_t x_{t-\tau}' \hat{u}_{t-\tau},$$

the $\hat{u}_t$ are OLS regression residuals, and T is sample size. Indeed, almost all leading HAC estimators are of the form (5), distinguished only by their choice of truncation lag h.

We will explore several leading truncation lag choices, including:

(1) NW: Newey-West (5) with $h = \lfloor 4(T/100)^{2/9} \rfloor + 1$. This h choice is a standard textbook recommendation (e.g., Wooldridge (2015)).

(2) NW-A: Newey-West (5) with $h = \lfloor 0.75T^{1/3} \rfloor + 1$. This h choice is also standard, arising when a formula in Andrews (1991) is specialized to the case of a first-order autoregression with coefficient 0.25.

(3) NW-LLSW: Newey-West (5) with $h = \lfloor 1.3T^{1/2} \rfloor + 1$, as proposed by Lazarus et al. (2018). Its use of $T^{1/2}$ rather than $T^{2/9}$ or $T^{1/3}$ as in NW or NW-A, respectively, produces higher truncation lags. For example, if $T = 200$, then NW selects $h = 5$ but NW-LLSW selects $h = 19$.

(4) NW-KV: Newey-West (5) with $h = T$, as proposed by Kiefer and Vogelsang (2002),

which builds on Kiefer et al. (2000). $h = T$ is of course the maximum possible truncation lag.

We also explore the Müller (2014) HAC estimator (we denote it by M), which is not in the Newey-West family. Instead it is an orthogonal series estimator, using a type-II discrete cosine transform to produce an equally-weighted average of projections on cosines. The M estimator is:

$$\widehat{\Omega} = \frac{1}{\nu} \sum_{j=1}^{\nu} \widehat{\Lambda}_j \widehat{\Lambda}_j',$$

where

$$\widehat{\Lambda}_j = \sqrt{\frac{2}{T}} \sum_{t=1}^{T} (x_t \hat{u}_t) \cos\left(\pi j \left(\frac{t - 1/2}{T}\right)\right).$$

The M truncation parameter, ν, is the total number of cosines included in the average projection. Lazarus et al. (2018) suggest setting $\nu = \lfloor 0.4T^{2/3} \rfloor$, producing the M-LLSW estimator.

2.2 Dynamic Regression (DynReg)

An alternative to HAC regression is augmenting regression (1) with lags of y and x to capture dynamics, in a fashion identical to an arbitrary equation in a vector autoregression. The p^{th}-order dynamic regression (DynReg for short) is

$$y_t = \sum_{j=1}^{p} \phi_j y_{t-j} + \sum_{i=1}^{k} \beta_i x_{i,t} + \sum_{j=1}^{p} \sum_{i=1}^{k} \gamma_{i,j} x_{i,t-j} + \varepsilon_t, \tag{6}$$

and it has $p + k + kp$ parameters. DynReg parameter estimates and standard errors are obtained by OLS.

If u_t in equation (1) is a finite-ordered $AR(p)$ process with p known, then the DynReg (6) holds exactly. That is, if the DGP is

$$y_t = x_t'\beta + u_t,$$

$$\phi(L)u_t = \varepsilon_t,$$

where $\phi(L) = 1 - \phi_1 L - \phi_2 L^2 - ... - \phi_p L^p$, then we can re-write it as

$$\phi(L)y_t = \phi(L)x_t'\beta + \varepsilon_t$$

or

$$y_t = \sum_{j=1}^{p} \phi_j y_{t-j} + \sum_{i=1}^{k} \beta_i x_{i,t} + \sum_{j=1}^{p} \sum_{i=1}^{k} \beta_i \phi_j x_{i,t-j} + \varepsilon_t. \tag{7}$$

This is commonly called a "common factor regression" in reference to the embedded coefficient constraint ($\gamma_{i,j} = \beta_i \phi_j$). The usual asymptotic inference is immediately available:

$$T^{1/2}(\widehat{\theta}_{OLS} - \theta) \to N(0, Q^{-1}), \tag{8}$$

where θ is the vector of dynamic regression parameters,

$$Q = plim \left(T^{-1} \sum_{t=1}^{T} z_t z_t' \right), \tag{9}$$

and $z_t' = (y_{t-1}, ..., y_{t-p},\ x_{1,t}, ..., x_{k,t},\ x_{1,t-1}, ..., x_{k,t-1},\ ...,\ x_{1,t-p}, ..., x_{k,t-p})$.

Alternatively, if u is an $AR(p)$ process with p *un*known, then the DynReg (6) is approximate rather than exact, but the limiting distribution (8) remains valid if the fitted autoregressive order is increased with sample size at a suitable rate (Grenander, 1981; Hannan and Deistler, 1988), as achieved by standard information criteria.

In particular, if a p_{max} is known such that $p \leq p_{max}$, then a consistent selection criterion (in the model selection sense) like BIC is a natural choice for order selection. The DynReg BIC is

$$\text{BIC} = -2lnL + \log(T)(p + k + kp). \tag{10}$$

The obvious benchmark likelihood is Gaussian, in which case model rankings by (10) match those from an "SSE version" of the BIC,

$$\text{BIC}_{\text{SSE}} = T\log(\text{SSE}) + \log(T)(p + k + kp), \tag{11}$$

where SSE is the sum of squared errors from equation (7).

Otherwise, an efficient selection criterion (in the model selection sense) like AIC is a natural choice for order selection. The DynReg AIC is

$$\text{AIC} = -2lnL + 2(p + k + kp),$$

the Gaussian SSE version of which is

$$\text{AIC}_{\text{SSE}} = T\log(\text{SSE}) + 2(p+k+kp).$$

AIC and BIC differ only in their degrees-of-freedom penalties, with AIC penalizing less harshly.

In closing this section, we note that we could impose fewer or more restrictions on the DynReg (6) than presently. As for fewer restrictions, we could allow different DynReg lag lengths for y and the x_i's. We could also allow heteroskedasticity (e.g., GARCH) in the DynReg disturbances, which we have implicitly assumed away to focus completely on serial correlation. Conversely, additional restrictions could be imposed. For example, one could impose the common-factor restriction in DynReg estimation if desired.

2.3 Discussion

Now that we have introduced both HAC and DynReg, we offer some insights into their comparative structure, which foreshadow our subsequent analysis and results in sections 3 and 4.

(1) Unlike HAC, DynReg *models* the serial correlation. One expects this to translate into superior efficiency in estimation of β and superior accuracy in forecasting y. We explore these issues in section 3.

(2) Closely related to DynReg's modeling of serial correlation, DynReg naturally leverages simple information criteria (BIC and/or AIC) with well-known optimality properties for bandwidth (p) selection. HAC regression, in contrast, relies on one or another of various "rules of thumb" for bandwidth (h or ν) selection.

(3) DynReg features an autoregressive approximation to dynamics, via its direct inclusion of autoregressive lags. Autoregressions have become a great workhorse of empirical dynamic economics precisely because low-ordered autoregressions routinely provide accurate approximations in economic contexts. In contrast, HAC regression is closely linked to low-ordered *moving average* approximations, by virtue of the HAC covariance matrix construction in terms of low-ordered autocorrelations. MA(q) dynamics, in particular, would be quickly and exactly captured by HAC regresion with $h = q$ and a rectangular lag window. One expects that HAC regression would quickly provide reliable inference there, but again, if dynamics are approximately autoregressive with

large dominant root (the leading case in macroeconomics, for example), one expects that HAC regression would perform poorly there, except in the very large samples that facilitate very long truncation lags.

(4) The validity of DynReg inference does not require strong exogeneity, just as HAC inference does not require strong exogenity. One of the first papers to advocate OLS HAC was by Hansen and Hodrick (1980) who were motivated by the situation when a variable with obervable forecast errors is regressed on explanatory variables containing lagged forecast errors. The null hypothesis of such regressions implied MA disturbances that implied violations of strong exogeneity. Hansen and Hodrick (1980) correctly realized that GLS could be inconsistent in these situations, while OLS HAC was consistent, albeit inefficient. However, this insight does not make OLS HAC preferable to DynReg. First, the situation involving lagged observable forecast errors and issues with strong exogeneity are a comparatively rare formulation in economic time series regression and are by no means a justification for the automatic implementation of OLS HAC. Second, and very importantly, the objections raised by Hansen and Hodrick (1980) are simply overcome by the use of DynReg, where only the weak exogeneity requirement $E(x_t u_t) = 0$ is necessary for consistency.

3 On the Costs of HAC Regression's Failure to Model Serial Correlation

In this section we highlight and assess the inescapable losses of estimation efficiency and forecast accuracy due to HAC regression's failure to model serial correlation, quite apart from the issue of whether HAC standard errors produce correctly-sized tests, which we explore later in section 4.

Much of the analysis will by necessity proceed by Monte Carlo. We use precisely the same data-generating process (DGP) and experimental design as Lazarus et al. (2018), with one $AR(1)$ covariance stationary right-hand-side variable and a covariance stationary $AR(1)$ error processes. More precisely,

$$y_t = x_t \beta + u_t$$

where

$$x_t = \rho x_{t-1} + \varepsilon_{x,t}$$

and

$$u_t = \rho u_{t-1} + \varepsilon_{u,t},$$

for $t = 1, ..., T$, with $\varepsilon_{x,t}$ and $\varepsilon_{u,t}$ distributed as $N(0, \sigma_x^2)$ and $N(0, \sigma_u^2)$, respectively, and uncorrelated at all leads and lags. We exploit invariance by setting $\sigma_x^2 = \sigma_u^2 = \beta = 1$.

We explore $\rho \in \{0, .3, .5, .7, .9, .95, .99\}$, which spans the relevant range for economics. All ρ values are positive, as economic time series are generally positively serially correlated, and they range from white noise to the very strong serial correlation often of relevance in macroeconomics. Including white noise ($\rho = 0$) is convenient as way to check our Monte Carlo against known results for the iid case.

We explore $T \in \{50, 200, 600, 2500\}$, which also spans the relevant range for macroeconomics, where structural change and other considerations tend to keep sample spans to roughly "the most recent fifty years"; that is, sample sizes of 50 years, or 200 quarters, or 600 months, or approximately 2500 weeks. Including $T = 2500$ also lets us check our Monte Carlo against known large-sample results.

We simulate exact realizations of x and u by drawing x_0 and u_0 from their stationary distribution at each Monte Carlo replication. We use common random numbers whenever appropriate. We perform 10,000 Monte Carlo replications.

We select the DynReg lag order using BIC, selecting the lag order that minimises

$$\text{BIC} = T \log(\text{SSE}) + \log(T)(2p + 1),$$

which is the general BIC, (11), specialized to $k = 1$.

3.1 HAC vs DynReg Estimation Efficiency

HAC regression preserves OLS parameter estimates, simply adjusting standard errors, meaning that it gives up on efficient estimation of $\hat{\beta}$, instead settling for mere consistency. In large samples the efficiency loss from using OLS is arguably inconsequential, but economic time-series sample sizes are often small. In macroeconomics, for example, one typically has just a couple hundred quarters of highly-serially-correlated data.

We present the estimation efficiency results in Table 1. The key object of interest is RE_{est}, the efficiency of DynReg relative to HAC,

$$\text{RE}_{\text{est}} = \frac{\text{MSE(HAC)}}{\text{MSE(DynReg)}}.$$

Table 1: Bias, Variance, MSE, and Efficiency of DynReg relative to OLS, Autoregressive Disturbances, BIC DynReg Lag-Order Selection

		$\rho = 0$	$\rho = 0.3$	$\rho = 0.5$	$\rho = 0.7$	$\rho = 0.9$	$\rho = 0.95$	$\rho = 0.99$
T=50								
Bias	HAC	0.0007	-0.0002	0.0005	-0.0007	-0.0029	0.0026	-0.0292
	DynReg	0.0003	-0.0009	-0.0008	-0.0012	-0.0036	-0.0004	-0.0012
Variance	HAC	0.0204	0.0248	0.0343	0.0587	0.2348	0.8037	9.6973
	DynReg	0.0216	0.025	0.0255	0.0238	0.0245	0.0255	0.0253
MSE	HAC	0.0204	0.0248	0.0343	0.0587	0.2348	0.8036	9.6972
	DynReg	0.0216	0.025	0.0255	0.0238	0.0245	0.0255	0.0253
		(0; 0.3)	(0; 0.7)	(1; 1.3)	(1; 1.4)	(1; 1.5)	(1; 1.5)	(1; 1.6)
RE_{est}		0.9462	0.993	1.342	2.4695	9.5837	31.5066	383.3278
T=200								
		$\rho = 0$	$\rho = 0.3$	$\rho = 0.5$	$\rho = 0.7$	$\rho = 0.9$	$\rho = 0.95$	$\rho = 0.99$
Bias	HAC	0.0001	0.0010	0.0010	-0.0009	-0.0020	0.0024	0.0046
	DynReg	0.0001	0.0002	0.0009	-0.0010	-0.0005	-0.0002	-0.0003
Variance	HAC	0.0051	0.0061	0.0085	0.0145	0.0514	0.1526	2.7592
	DynReg	0.0051	0.0053	0.0052	0.0052	0.0051	0.0053	0.0052
MSE	HAC	0.0051	0.0061	0.0085	0.0145	0.0514	0.1526	2.7589
	DynReg	0.0051	0.0053	0.0052	0.0052	0.0051	0.0053	0.0052
		(0; 0)	(1; 1)	(1; 1)	(1; 1)	(1; 1)	(1; 1)	(1; 1)
RE_{est}		0.9991	1.1508	1.6162	2.7929	10.0494	28.7381	535.1801
T=600								
		$\rho = 0$	$\rho = 0.3$	$\rho = 0.5$	$\rho = 0.7$	$\rho = 0.9$	$\rho = 0.95$	$\rho = 0.99$
Bias	HAC	-0.0008	-0.0002	-0.0004	0.0004	-0.0007	-0.0028	-0.0062
	DynReg	-0.0008	-0.0001	-0.0001	-0.0001	-0.0005	-0.0003	0.0004
Variance	HAC	0.0017	0.002	0.0029	0.0049	0.0168	0.0431	0.9138
	DynReg	0.0017	0.0017	0.0017	0.0017	0.0017	0.0017	0.0017
MSE	HAC	0.0017	0.002	0.0029	0.0049	0.0168	0.0431	0.9138
	DynReg	0.0017	0.0017	0.0017	0.0017	0.0017	0.0017	0.0017
		(0; 0)	(1; 1)	(1; 1)	(1; 1)	(1; 1)	(1; 1)	(1; 1)
RE_{est}		0.9999	1.2002	1.6605	2.8961	10.1162	26.0652	541.6975
T=2500								
		$\rho = 0$	$\rho = 0.3$	$\rho = 0.5$	$\rho = 0.7$	$\rho = 0.9$	$\rho = 0.95$	$\rho = 0.99$
Bias	HAC	-0.0001	-0.0001	-0.0001	0.0001	0.0004	-0.0004	-0.0005
	DynReg	-0.0001	-0.0001	-0.0002	0.0001	-0.0003	0.0001	-0.0001
Variance	HAC	0.0004	0.0005	0.0007	0.0012	0.0039	0.0087	0.1959
	DynReg	0.0004	0.0004	0.0004	0.0004	0.0004	0.0004	0.0004
MSE	HAC	0.0004	0.0005	0.0007	0.0012	0.0039	0.0087	0.1958
	DynReg	0.0004	0.0004	0.0004	0.0004	0.0004	0.0004	0.0004
		(0; 0)	(1; 1)	(1; 1)	(1; 1)	(1; 1)	(1; 1)	(1; 1)
RE_{est}		1.0000	1.2090	1.6925	2.9168	9.6878	21.9029	493.2473

Notes: The data-generating process is $y_t = x_t + u_t$, $x_t = \rho x_{t-1} + \epsilon_{x,t}$, $u_t = \rho u_{t-1} + \epsilon_{u,t}$, $t = 1, ..., T$. All shocks are $N(0,1)$, orthogonal at all leads and lags. RE_{est} is the estimation efficiency of DynReg relative to OLS: RE_{est}=MSE(OLS)/MSE(DynReg). We perform 10000 Monte Carlo replications. The median and mean BIC-selected DynReg lags appear in parentheses. See text for details.

We also show MSE, bias, and variance for both HAC and DynReg.[2]

Let us begin directly with the bottom-line RE_{est} results. For any fixed sample size T, RE_{est} is increasing in serial correlation strength ρ. Consider, for example, a leading case like $T = 200$ corresponding, to fifty years of quarterly data. For $\rho = 0$, $\text{RE}_{\text{est}} = 1$, as it should since there is no serial correlation. RE_{est} grows quickly as ρ increases, however, reaching 2.8 when $\rho = 0.7$ and 28.7 when $\rho = 0.95$.

In contrast, for any fixed serial correlation strength ρ, RE_{est} stabilizes quickly in sample size T and remains approximately constant. Consider, for example, a realistic case like $\rho = 0.9$. RE_{est} remains at approximately $\text{RE}_{\text{est}} = 10$ for all sample sizes $T \in \{50, 200, 600, 2500\}$. Hence RE_{est} is clearly driven by serial correlation strength and not by sample size.

Now consider separately the HAC and DynReg MSEs that underlie RE_{est} is composed. For any fixed sample size T, MSE(HAC) is strongly increasing in serial correlation strength ρ (because the HAC estimator ignores serial correlation), whereas MSE(DynReg) is invariant to serial correlation strength (because the DynReg estimator controls for serial correlation). That is why the RE_{est} ratio is also strongly increasing in ρ, as documented earlier.

In contrast, for any fixed serial correlation strength ρ, *both* MSE(HAC) and MSE(DynReg) decrease with sample size T (as they must, since both the HAC and DynReg estimators are consistent), but they decrease proportionately, so that the RE_{est} ratio is invariant to T, as documented earlier.

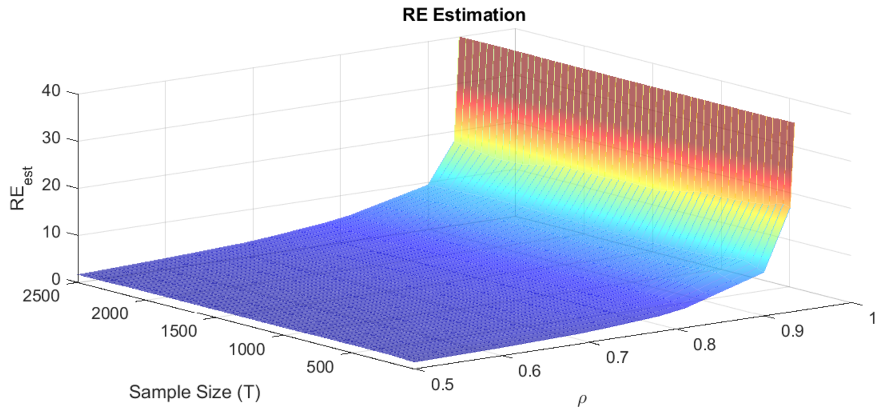
Finally, let us examine the bias and variance components that underlie the MSEs. First consider bias. Both the HAC and DynReg estimators are theoretically unbiased for any serial correlation strength and sample size, and the Monte Carlo confirms the theory: the estimated biases are always negligible and invariant to ρ.[3]

Now consider variance. HAC variance increases sharply with serial correlation strength (because HAC ignores serial correlation), whereas DynReg variance does not (because DynReg controls for serial correlation), and both variances decrease with sample size (by consistency), but they do so proportionately. Assembling all the variance results reveals that they drive the MSE results. That is, comparative HAC vs DynReg MSE patterns, and hence ultimately comparative RE_{est} patterns, are driven by entirely by variance.

[2]Note that we can simply speak of "HAC" for now, rather than worrying about particular HAC estimators like NW, NW-A, and so on, because *all* HAC estimators simply use the OLS estimator of β. Particular HAC estimators will have particular effects on the *standard errors* of $\hat{\beta}$, but not on $\hat{\beta}$ itself, which always remains just $\hat{\beta}_{OLS}$.

[3]Moreover the estimated biases decrease with T, as expected, by consistency.

Figure 1: Estimation Efficiency of DynReg Relative to HAC Regression, Autoregressive Disturbances, BIC DynReg Lag-Order Selection



Notes: The data-generating process is $y_t = x_t + u_t$, $x_t = \rho x_{t-1} + \epsilon_{x,t}$, $u_t = \rho u_{t-1} + \epsilon_{u,t}$, $t = 1, ..., T$. All shocks are $N(0, 1)$, orthogonal at all leads and lags. DynReg lag order selected by BIC. We perform 10000 Monte Carlo replications, using common random numbers whenever appropriate, with x_0 and u_0 drawn from the stationary distribution. Values for $\rho = 0.99$ are not shown in the plot due to their relative extreme magnitude as shown in Table 1 and 2. See text for details.

3.2 HAC vs DynReg Forecast Accuracy

Explicit modeling of autocorrelation can be used for improved prediction. OLS-HAC estimators neglect this and therefore produce suboptimal forecasts. First consider the case of known parameters, which we can solve analytically. As before the DGP is

$$y_t = x_t + u_t$$

$$x_t = \rho x_{t-1} + \epsilon_{x,t}$$

$$u_t = \rho u_{t-1} + \epsilon_{u,t},$$

with all shocks $N(0, 1)$ and orthogonal at all leads and lags. In an obvious notation, the optimal forecast accounting for serial correlation in u is

$$\begin{aligned} y^{opt}_{t+1,t} &= x_{t+1,t} + u_{t+1,t} \\ &= \rho x_t + \rho u_t. \end{aligned} \tag{12}$$

The corresponding forecast error is $e^{opt}_{t+1} = \varepsilon_{x,t+1} + \varepsilon_{u,t+1}$, with variance σ^2_{opt}=2.

Table 2: Forecast Accuracy of DynReg relative to OLS,
Autoregressive Disturbances, BIC DynReg Lag-Order Selection

Relative Prediction Efficiency (RE_{pred})							
T	$\rho = 0$	$\rho = 0.3$	$\rho = 0.5$	$\rho = 0.7$	$\rho = 0.9$	$\rho = 0.95$	$\rho = 0.99$
50	0.989	1.042	1.160	1.452	3.033	5.865	391.908
200	0.997	1.051	1.168	1.476	3.121	5.698	47.361
600	1.000	1.047	1.152	1.505	3.214	5.605	25.569
2500	1.000	1.049	1.163	1.469	3.101	5.656	25.648

Notes: The data-generating process is $y_t = x_t + u_t$, $x_t = \rho x_{t-1} + \epsilon_{x,t}$, $u_t = \rho u_{t-1} + \epsilon_{u,t}$, $t = 1, ..., T$. All shocks are $N(0,1)$, orthogonal at all leads and lags. RE_{pred} is the predictive efficiency of DynReg relative to OLS: RE_{pred}=MSPE(OLS)/MSPE(DynReg), where $MSPE$ is 1-step-ahead mean squared prediction error. We perform 10000 Monte Carlo replications. See text for details.

The suboptimal forecast, failing to account for serial correlation in u, is just the first term in (12),

$$y_{t+1,t}^{subopt} = \rho x_t.$$

The corresponding forecast error is $e_{t+1}^{subopt} = \varepsilon_{x,t+1} + u_{t+1}$, with variance $\sigma^2_{subopt} = 1 + \frac{1}{1-\rho^2}$.

Both forecasts are unbiased, so the relative prediction efficiency (relative MSPE) is just the relative variance, which is

$$RE_{pred} = \frac{\sigma^2_{\text{subopt}}}{\sigma^2_{\text{opt}}} = \frac{1}{2} + \frac{1}{2(1-\rho^2)}. \quad (13)$$

RE_{pred} is bounded below by 1, which occurs when ρ=0, and $RE_{pred} \rightarrow \infty$ monotonically as $\rho \rightarrow 1$. Crucially, note that *RE_{pred} is effectively the predictive efficiency of DynReg relative to HAC, because DynReg produces the optimal forecast (which exploits serial correlation) and HAC produces the suboptimal forecast (which ignores serial correlation).*

We now consider the case of estimated parameters, which is more complicated. In Table 2 we show RE_{pred} estimated by Monte Carlo, accounting for parameter estimation uncertainty. For all but the most extreme cases (e.g., $T = 50$ with $\rho = 0.99$) the Monte Carlo results are almost identical to the analytic result (13) that ignores parameter estimation uncertainty.[4] Hence RE_{pred} depends strongly on ρ but not on T. More precisely, for any T we of course

[4]This is because the effects of parameter estimation uncertainty on MSPE vanish quickly (like $1/T$ rather than $1/\sqrt{T}$), as is well known. Hence the earlier-documented poor estimation efficiency of HAC relative to DynReg, although a large problem for some purposes, is not an important problem for forecasting.

obtain $\mathrm{RE}_{\mathrm{pred}} = 1$ in the white noise case ($\rho = 0$), but then $\mathrm{RE}_{\mathrm{pred}}$ grows quickly in ρ, and for any ρ, $\mathrm{RE}_{\mathrm{pred}}$ stabilizes extremely quickly in T and is basically constant.

4 On the Performance of HAC-Based Hypothesis Tests

Now we consider the finite-sample properties of hypothesis tests associated with the various procedures. We first consider test sizes, after which we consider rejection frequencies.

4.1 Size

We show estimated size distortions for a 5% test of $H_0\!: \beta{=}1$ numerically in Table 3 and graphically in the response surfaces of Figure 2.[5] There are three key results:

(1) Tests based on OLS are incorrectly sized for all (ρ, T) combinations except when $\rho = 0$, and the size distortions become huge as ρ grows.

(2) The various NW and M HAC corrections reduce but do not eliminate the size distortion. In particular, important distortion generally remains in the economically crucial region of $\rho \in [0.6, 0.99]$, depending on the sample size and the precise NW version used. NW and NW-A are worst, M-LLSW and NW-LLSW are better, and NW-KV is best.

(3) *Tests based on DynReg, in contrast, are correctly sized for all (ρ, T) combinations, even with extremely strong autocorrelation.*

4.2 Rejection Frequencies and Power

Only tests that are correctly sized for all (T, ρ) combinations are of real interest, because only correctly-sized tests produce unambiguously interpretable rejections. DynReg satisfies this requirement, and HAC regression does not. One could simply stop here, but it may be of interest to compare rejection frequencies in a few laboratory environments where the DGP is known.

In Figure 3 we show rejection frequency curves for a fixed sample size $T = 200$ and $\rho \in \{0.0, 0.3, 0.5, 0.7, 0.9, 0.95\}$. We include in the figure only the tests based on NW-KV,

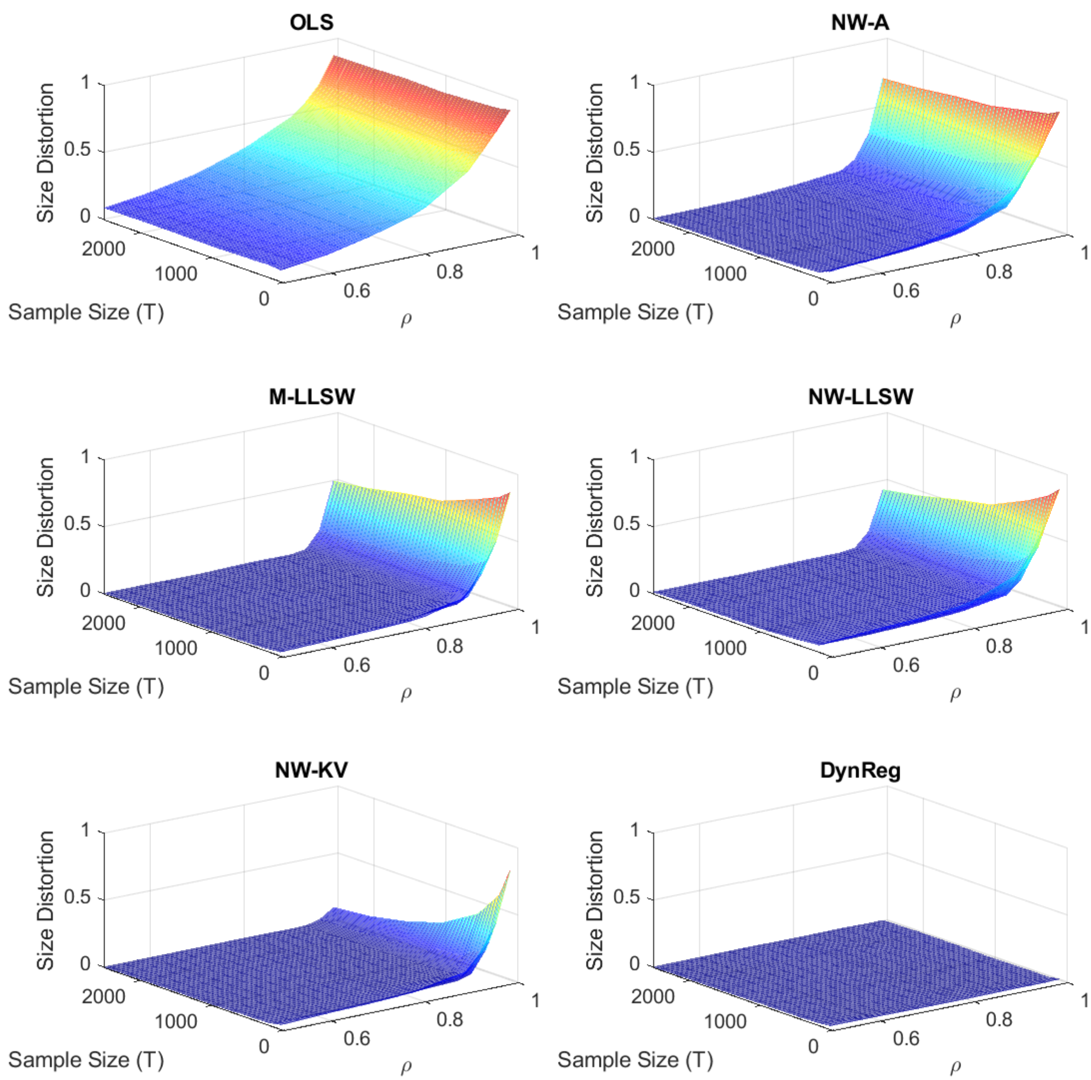
[5] The response surfaces use a larger set of ρ and T values than used elsewhere. In particular, we compute empirical size distortions for the 5% test of $H_0 : \beta = 1$ for all combinations of $\rho \in \{0, 0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9, 0.95, 0.99\}$ and $T \in \{50, 200, 500, 1000, 1500, 2000, 2500\}$. Then, based on the grid of empirical size distortions for all (ρ, T) combinations, we fit response surfaces.

Table 3: Empirical size of nominal 5% t-test of H_0: β=1,
Autoregressive Disturbances, BIC DynReg Lag-Order Selection

	Truncation	$\rho = 0$	$\rho = 0.3$	$\rho = 0.5$	$\rho = 0.7$	$\rho = 0.9$	$\rho = 0.95$	$\rho = 0.99$
T=50								
OLS	–	0.048	0.072	0.127	0.242	0.564	0.774	0.970
NW	$h = \lfloor 4[T/100]^{2/9} \rfloor + 1$	0.079	0.089	0.119	0.169	0.421	0.674	0.953
NW-A	$h = \lfloor 0.75T^{1/3} \rfloor + 1$	0.073	0.083	0.116	0.180	0.452	0.702	0.958
NW-LLSW	$h = \lfloor 1.3T^{1/2} \rfloor + 1$	0.118	0.123	0.149	0.179	0.359	0.603	0.938
NW-KV	$h = T$	0.053	0.061	0.072	0.092	0.189	0.388	0.879
M-LLSW	$\nu = \lfloor 0.41T^{2/3} \rfloor$	0.055	0.064	0.073	0.080	0.197	0.455	0.914
DynReg	–	0.055	0.072	0.073	0.060	0.063	0.066	0.067
T=200								
OLS	–	0.049	0.073	0.135	0.247	0.544	0.727	0.955
NW	$h = \lfloor 4[T/100]^{2/9} \rfloor + 1$	0.060	0.063	0.081	0.114	0.280	0.498	0.916
NW-A	$h = \lfloor 0.75T^{1/3} \rfloor + 1$	0.060	0.063	0.081	0.114	0.280	0.498	0.916
NW-LLSW	$h = \lfloor 1.3T^{1/2} \rfloor + 1$	0.082	0.085	0.095	0.107	0.176	0.319	0.858
NW-KV	$h = T$	0.049	0.052	0.058	0.061	0.089	0.136	0.671
M-LLSW	$\nu = \lfloor 0.41T^{2/3} \rfloor$	0.051	0.054	0.056	0.067	0.113	0.238	0.847
DynReg	–	0.050	0.055	0.049	0.053	0.050	0.055	0.051
T=600								
OLS	–	0.052	0.071	0.137	0.248	0.536	0.688	0.944
NW	$h = \lfloor 4[T/100]^{2/9} \rfloor + 1$	0.056	0.055	0.074	0.092	0.219	0.380	0.878
NW-A	$h = \lfloor 0.75T^{1/3} \rfloor + 1$	0.056	0.055	0.072	0.087	0.199	0.353	0.868
NW-LLSW	$h = \lfloor 1.3T^{1/2} \rfloor + 1$	0.069	0.066	0.074	0.080	0.116	0.171	0.746
NW-KV	$h = T$	0.047	0.045	0.050	0.053	0.061	0.069	0.429
M-LLSW	$\nu = \lfloor 0.41T^{2/3} \rfloor$	0.055	0.050	0.056	0.058	0.083	0.136	0.756
DynReg	–	0.052	0.048	0.056	0.051	0.051	0.049	0.051
T=2500								
OLS	–	0.050	0.075	0.128	0.249	0.535	0.673	0.924
NW	$h = \lfloor 4[T/100]^{2/9} \rfloor + 1$	0.052	0.053	0.061	0.074	0.150	0.263	0.771
NW-A	$h = \lfloor 0.75T^{1/3} \rfloor + 1$	0.053	0.054	0.060	0.070	0.129	0.229	0.748
NW-LLSW	$h = \lfloor 1.3T^{1/2} \rfloor + 1$	0.057	0.059	0.061	0.064	0.069	0.095	0.475
NW-KV	$h = T$	0.050	0.046	0.051	0.049	0.052	0.052	0.144
M-LLSW	$\nu = \lfloor 0.41T^{2/3} \rfloor$	0.050	0.051	0.053	0.056	0.056	0.082	0.535
DynReg	–	0.050	0.049	0.047	0.049	0.048	0.049	0.048

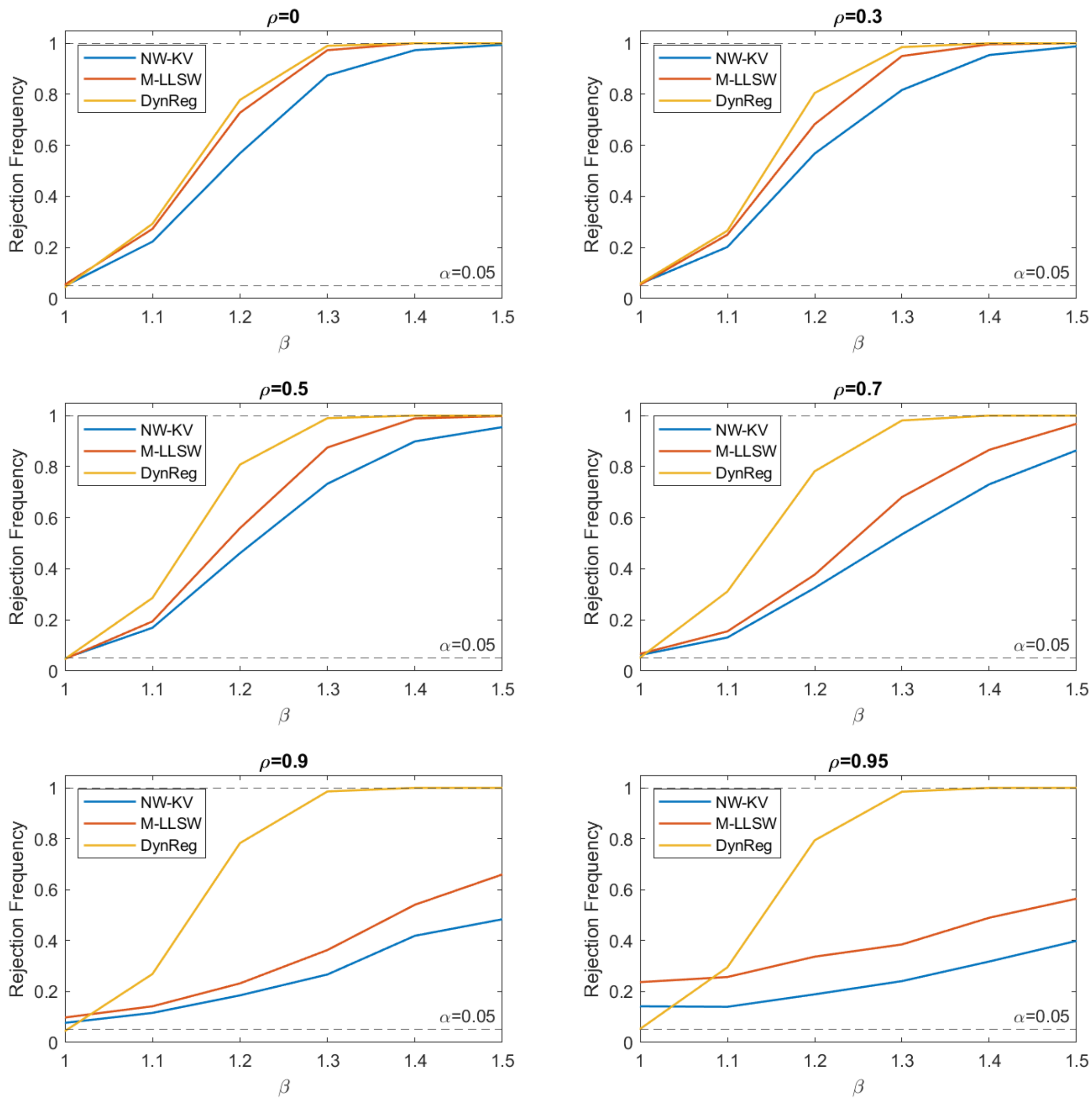
Notes: The data-generating process is $y_t = x_t + u_t$, $x_t = \rho x_{t-1} + \epsilon_{x,t}$, $u_t = \rho u_{t-1} + \epsilon_{u,t}$, $t = 1, ..., T$. All shocks are $N(0, 1)$, orthogonal at all leads and lags. We perform 10 000 Monte Carlo replications. See text for details.

Figure 2: Empirical Size Distortion of Nominal 5% t-Test of H_0: β=1, Autoregressive Disturbances, BIC DynReg Lag-Order Selection



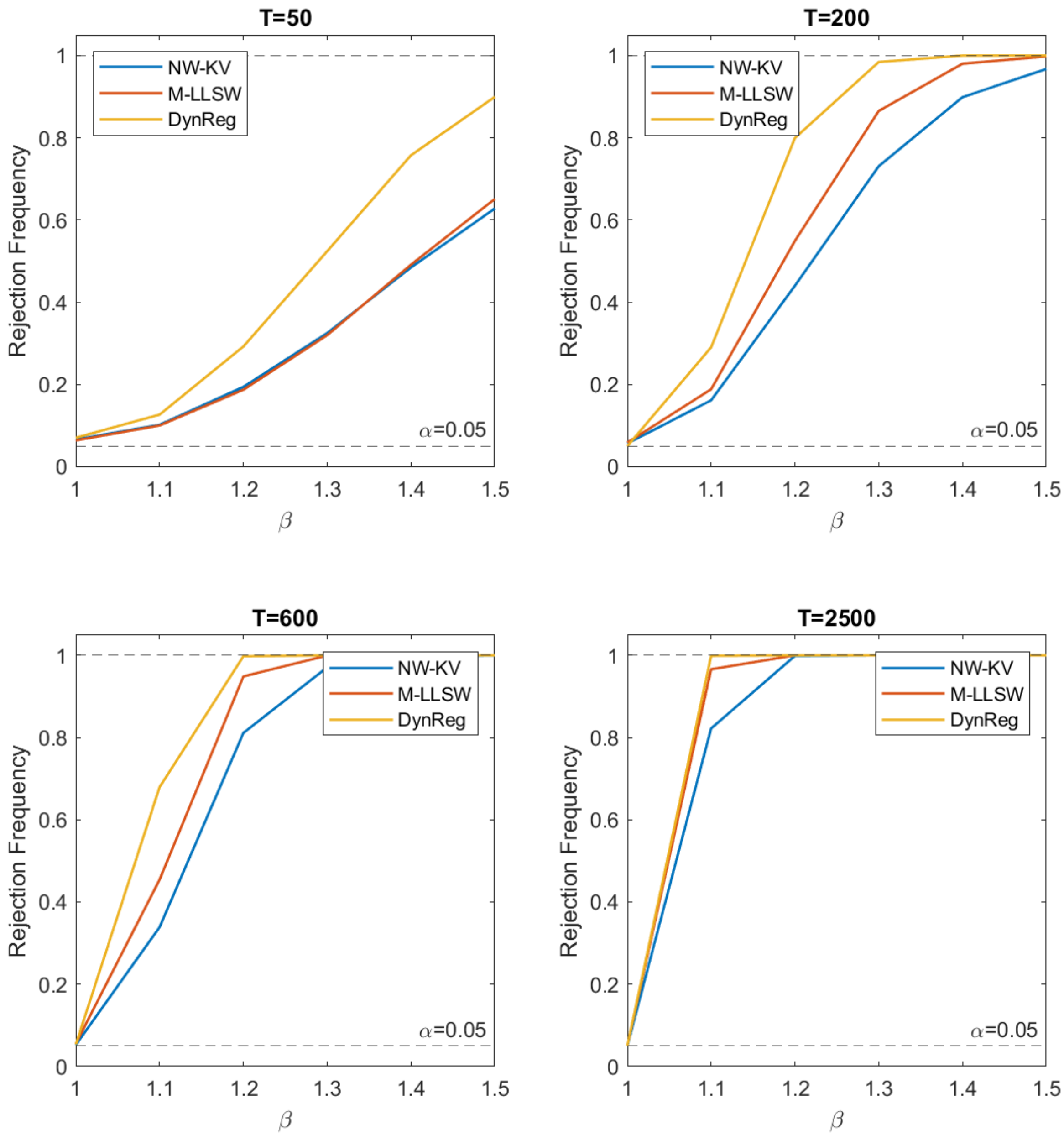
Notes: The data-generating process is $y_t = x_t + u_t$, $x_t = \rho x_{t-1} + \epsilon_{x,t}$, $u_t = \rho u_{t-1} + \epsilon_{u,t}$, $t = 1, ..., T$. All shocks are $N(0,1)$, orthogonal at all leads and lags. DynReg lag order selected by BIC. We perform 10 000 Monte Carlo replications, using common random numbers whenever appropriate, with x_0 and u_0 drawn from the stationary distribution. See text for details.

Figure 3: Empirical Rejection frequencies of Nominal 5% t-Test of H_0: β=1, Autoregressive Disturbances, $T = 200$, BIC DynReg Lag-Order Selection



Notes: The data-generating process is $y_t = \beta x_t + u_t$, $x_t = \rho x_{t-1} + \epsilon_{x,t}$, $u_t = \rho u_{t-1} + \epsilon_{u,t}$, $t = 1, ..., T$. All shocks are $N(0, 1)$, orthogonal at all leads and lags. DynReg lag order selected by BIC. We perform 10 000 Monte Carlo replications, using common random numbers whenever appropriate, with x_0 and u_0 drawn from the stationary distribution. See text for details.

Figure 4: Empirical Rejection frequencies of Nominal 5% t-Test of H_0: β=1, Autoregressive Disturbances, $\rho = 0.5$, BIC DynReg Lag-Order Selection



Notes: The data-generating process is $y_t = \beta x_t + u_t$, $x_t = \rho x_{t-1} + \epsilon_{x,t}$, $u_t = \rho u_{t-1} + \epsilon_{u,t}$, $t = 1, ..., T$. All shocks are $N(0, 1)$, orthogonal at all leads and lags. DynReg lag order selected by BIC. We perform 10 000 Monte Carlo replications, using common random numbers whenever appropriate, with x_0 and u_0 drawn from the stationary distribution. See text for details.

M-LLSW and DynReg because, as shown in Table 3, other HAC-based tests are over-sized for most values of ρ. NW-KV, M-LLSW, and DynReg, in contrast, are approximately correctly sized when $T = 200$ for $\rho \leq 0.7$ so we can actually compare power for $\rho \in \{0.0, 0.3, 0.5, 0.7\}$. For $\rho = 0.9$ and $\rho = 0.95$ the HAC tests are over-sized, so rejections are not purely indicative of "power," but the rejection frequency curves may nevertheless be informative, so we show curves for $\rho = 0.9$ and $\rho = 0.95$ as well.

There are three basic results regarding finite-sample rejection frequencies:

(1) For null and moderate values of ρ (0, .3, .5 .7) we can compare the power of NW-KV, M-LLSW and DynReg. The HAC tests are dominated by DynReg, and the DynReg dominance increases with ρ.

(2) *DynReg dominance increases with ρ not because DynReg power curves are shifting up, but because the HAC power curves are shifting down and flattening.* That is, the DynReg rejection curves are approximately invariant to ρ, attaining, for example, unit power at approximately the same alternative value ($\beta = 1.3$) regardless of the value of ρ. HAC rejection frequencies, in contrast, increase progressively less quickly as dynamics become progressively more persistent. Effectively the DynReg information criteria nail the proper DynReg lag order, regardless of ρ, whereas the HAC procedures struggle to produce adequate corrections as the dynamics become progressively more persistent.

(3) Despite the fact that NW-KV and M-LLSW are over-sized in the high serial correlation environments of $\rho = 0.90$ and $\rho = 0.95$, they nevertheless reject much less often that DynReg under the alternative.

In Figure 4 we show power curves for $\rho = 0.5$ and sample sizes $T = 50, 200, 600, 2500$. Of course all curves shift up as T increases, but DynReg always dominates.

5 Additional Analysis

In this section we provide additional analysis that supports and enhances our main earlier conclusions. First, we explore a DGP with moving-average, as opposed to autoregressive, dynamics. Second, we explore a DGP characterized by weak, but not strong, exogeneity.

5.1 Moving Average Disturbances and/or AIC DynReg Order Selection

In this subsection we consider results obtained using moving average (MA) disturbances, rather than autoregresive ones, as well as lag selection using AIC, given by $T\log(\text{SSE}) + 2(2p+1)$, rather than BIC. For the MA case we use a single-lag with parameter values $\rho \in \{0, .3, .5, .7, .9, .95, .99\}$. All these results are presented in Appendix Tables A1 to C2 and Figures A1 to C3.

We start by looking at estimation results. Allowing for AIC with its less parsimonious lag selection properties (Table B1) leads to slightly less favourable results for DynReg. However, its superiority compared to HAC estimation remains overwhelming. Finally, the use of an MA structure for u_t (Tables A1 and C1) produces far fewer problems for HAC estimation. This is potentially due to the fact that MA structures induce less pronounced autocorrelation patterns, but also because the structure of HAC inference promotes fine behaviour in MA environments. While DynReg and HAC have comparable performance in this case, the fact that MA error terms induce much better behaved estimation patterns, suggests that such errors are not a cause of concern, unlike the clear problematic issues associated with AR errors and HAC estimation.

Moving on to the size performance and considering AIC lag selection (Table B2), we note that again we obtain slightly less favourable results for DynReg with marginally increased rejection probabilities for small sample sizes. Nevertheless, its superior performance compared to HAC remains overwhelming. The use of an MA structure for u_t (Tables A2 and C2) produces, as for estimation, fewer problems for HAC inference, suggesting again that MA error structures should not be an important focus for our analysis.

Finally, commenting on the rejection probabilities under the alternative hypothesis, results for AIC lag selection and MA structures for u_t (Figures A1 to C3) retain the relative performance patterns noted earlier. DynReg remains better performing, apart from the case of the very pronounced MA error structure where HAC inference has a very slight power advantage probably driven by the need for DynReg, to have a lag structure with many lags to capture this form of autocorrelation.

5.2 Weak Exogeneity

Our Monte Carlo has thus far featured a DGP with strongly exogenous regressors, despite the fact that DynReg only assumes weak exogeneity. Before presenting some simulations for this case, we set the scene for this analysis. To do this, let $z_t = (x_t', u_t)'$ and assume that

$$z_t = \sum_{i=0}^{\infty} \Xi_i \varepsilon_{t-i}, \tag{14}$$

where $\varepsilon_t = (\varepsilon_{x,t}', \varepsilon_{u,t})'$. Of course, it is clear that in this case OLS estimation of (1) will not be consistent unless Ξ_i are of the form $\begin{pmatrix} \Xi_{xi} & 0 \\ 0 & \Xi_{ui} \end{pmatrix}$. On the other hand, for sufficiently long lag orders, p, the DynReg model retains its form given in (6), and repeated below for ease of reference.

$$y_t = \sum_{j=1}^{p} \phi_j y_{t-j} + \sum_{i=1}^{k} \beta_i x_{i,t} + \sum_{j=1}^{p} \sum_{i=1}^{k} \gamma_{i,j} x_{i,t-j} + \varepsilon_{u,t}, \tag{15}$$

A proof of this result follows easily given, e.g., Theorem 1.2.1. of Hannan and Deistler (1988). Clearly, DynReg remains consistent, unless $E(\varepsilon_{x,t}\varepsilon_{u,t}) \neq 0$. Now we analyze a DGP with only weakly exogenous regressors that illustrates this point. In particular, we set

$$\begin{pmatrix} x_t \\ u_t \end{pmatrix} = \begin{pmatrix} 0.5 & 0.2 \\ 0.3 & 0.4 \end{pmatrix} \begin{pmatrix} x_{t-1} \\ u_{t-1} \end{pmatrix} + \begin{pmatrix} \epsilon_{x,t} \\ \epsilon_{u,t} \end{pmatrix}. \tag{16}$$

where $(\varepsilon_{x,t}, \varepsilon_{u,t})' \sim_{iid} N(0, I)$. Table 4 reports results for this simulation design. It is clear that DynReg retains its attractive properties, including the absence of any bias, whereas OLS estimation is biased as expected given the above analysis.

6 Concluding Remarks

This paper has considered issues surrounding the application of OLS regression with HAC standard errors, in time-series environments. While the HAC methodology is very sensible in many cross section regression situations, it is not generally an effective procedure in time series regressions. Such regressions usually possess persistent autocorrelation. This causes HAC regressions to be very inefficient and sub-optimal in terms of parameter estimation and inference. HAC produces inefficient conditional predictions and leads to significant size distortions and reduced power in hypothesis testing of regression parameters. These problems are largely avoided by the use of dynamic regressions (DynReg), which are easily implemented and also avoid issues arising from the violation of strong exogeneity assumptions. The

Table 4: Estimation Statistics and Rejection Frequencies for a DGP with Weak Exogeneity

T=50

Estimation Statistics			Rejection Frequencies t-test H_0: $\beta = 1$		
			Test	$\beta = 1$	$\beta = 0.8$
Bias	HAC	0.2904	OLS	0.5549	0.1936
	DynReg	0.0180	NW	0.5140	0.1747
Variance	HAC	0.0317	NW-A	0.5207	0.1792
	DynReg	0.0295	NW-LLSW	0.5353	0.2038
MSE	HAC	0.1160	NW-KV	0.3390	0.1098
	DynReg	0.0298	M-LLSW	0.3407	0.1082
		(1; 1.3)	DynReg	0.0986	0.2792
RE_{est}		3.8958			

T=200

Estimation Statistics			Rejection Frequencies t-test H_0: $\beta = 1$		
			Test	$\beta = 1$	$\beta = 0.8$
Bias	HAC	0.3017	OLS	0.9669	0.3984
	DynReg	-0.0001	NW	0.9370	0.2967
Variance	HAC	0.0079	NW-A	0.9370	0.2967
	DynReg	0.0051	NW-LLSW	0.9309	0.2933
MSE	HAC	0.0990	NW-KV	0.7503	0.1764
	DynReg	0.0051	M-LLSW	0.8824	0.2174
		(1; 1.0)	DynReg	0.0477	0.7927
RE_{est}		19.5859			

T=600

Estimation Statistics			Rejection Frequencies t-test H_0: $\beta = 1$		
			Test	$\beta = 1$	$\beta = 0.8$
Bias	HAC	0.3066	OLS	1.0000	0.7219
	DynReg	0.0001	NW	1.0000	0.5928
Variance	HAC	0.0027	NW-A	1.0000	0.5864
	DynReg	0.0017	NW-LLSW	0.9999	0.5694
MSE	HAC	0.0967	NW-KV	0.9780	0.3764
	DynReg	0.0017	M-LLSW	0.9999	0.5141
		(1; 1.0)	DynReg	0.0474	0.9978
RE_{est}		58.6061			

T=2500

Estimation Statistics			Rejection Frequencies t-test H_0: $\beta = 1$		
			Test	$\beta = 1$	$\beta = 0.8$
Bias	HAC	0.3074	OLS	1.0000	0.9967
	DynReg	0.0003	NW	1.0000	0.9891
Variance	HAC	0.0007	NW-A	1.0000	0.9884
	DynReg	0.0004	NW-LLSW	1.0000	0.9858
MSE	HAC	0.0951	NW-KV	1.0000	0.8707
	DynReg	0.0004	M-LLSW	1.0000	0.9836
		(1; 1.0)	DynReg	0.0511	1.0000
RE_{est}		232.1908			

Notes: The data-generating process is $y_t = \beta x_t + u_t$, $x_t = \phi_{1,1}x_{t-1} + \phi_{1,2}u_{t-1} + \epsilon_{x,t}$, $u_t = \phi_{2,1}x_{t-1} + \phi_{2,2}u_{t-1} + \epsilon_{u,t}$, $t = 1, ..., T$ and $(\phi_{1,1}, \phi_{1,2}, \phi_{2,1}, \phi_{2,2}) = (0.5, 0.2, 0.3, 0.4)$. All shocks are $N(0, 1)$, orthogonal at all leads and lags. The median and mean BIC-selected DynReg lags appear in parentheses. We perform 10000 Monte Carlo replications, using common random numbers whenever appropriate, with x_0 and u_0 drawn from the stationary distribution. See text for details.

significant advantages of using DynReg are demonstrated with detailed simulations, which cover a range of practical and implementation issues.

Appendices

A Additional Monte Carlo: MA Errors, BIC Selection

Table A1: Bias, Variance, MSE, and Efficiency of DynReg relative to OLS
Moving Average Disturbances, BIC DynReg Lag-Order Selection

T=50								
		$\rho = 0$	$\rho = 0.3$	$\rho = 0.5$	$\rho = 0.7$	$\rho = 0.9$	$\rho = 0.95$	$\rho = 0.99$
Bias	HAC	0.0007	-0.0002	0.0002	0.0005	-0.0016	-0.0008	-0.0004
	DynReg	0.0003	-0.0009	-0.0009	-0.0011	-0.0044	0.0001	-0.0018
Variance	HAC	0.0204	0.024	0.0277	0.0269	0.0137	0.007	0.001
	DynReg	0.0216	0.025	0.0274	0.0292	0.035	0.0382	0.0377
MSE	HAC	0.0204	0.024	0.0277	0.0269	0.0137	0.007	0.001
	DynReg	0.0216	0.025	0.0274	0.0292	0.035	0.0382	0.0377
		(0; 0.3)	(0; 0.8)	(1; 1.4)	(2; 2.2)	(2; 3.2)	(2; 3.4)	(2; 3.5)
RE_{est}		0.9462	0.9606	1.0103	0.921	0.3922	0.1834	0.027
RE_{pred}		0.9899	1.0445	1.1001	1.2495	1.4047	1.4208	1.4723
T=200								
		$\rho = 0$	$\rho = 0.3$	$\rho = 0.5$	$\rho = 0.7$	$\rho = 0.9$	$\rho = 0.95$	$\rho = 0.99$
Bias	HAC	0.0001	0.0001	0.0009	-0.0007	-0.0002	-0.0001	-0.0001
	DynReg	0.0001	0.0003	0.0006	-0.0011	-0.0001	-0.0004	-0.0009
Variance	HAC	0.0051	0.0058	0.0067	0.0064	0.0032	0.0017	0.0002
	DynReg	0.0051	0.0053	0.0055	0.0057	0.006	0.0065	0.0066
MSE	HAC	0.0051	0.0058	0.0067	0.0064	0.0032	0.0017	0.0002
	DynReg	0.0051	0.0053	0.0055	0.0057	0.006	0.0065	0.0066
		(0; 0.0)	(1; 0.9)	(1; 1.5)	(2; 2.5)	(4; 3.6)	(4; 3.8)	(4; 3.9)
RE_{est}		0.9991	1.0991	1.22	1.1318	0.5299	0.2525	0.0307
RE_{pred}		0.9993	1.0403	1.1297	1.2451	1.4090	1.4555	1.4831
T=600								
		$\rho = 0$	$\rho = 0.3$	$\rho = 0.5$	$\rho = 0.7$	$\rho = 0.9$	$\rho = 0.95$	$\rho = 0.99$
Bias	HAC	-0.0008	-0.0001	-0.0006	0.0001	-0.0004	-0.0002	0.0000
	DynReg	-0.0008	0.0000	0.0000	-0.0001	-0.0003	-0.0002	0.0005
Variance	HAC	0.0017	0.0019	0.0023	0.0021	0.0011	0.0006	0.0001
	DynReg	0.0017	0.0017	0.0018	0.0018	0.0018	0.0019	0.0020
MSE	HAC	0.0017	0.0019	0.0023	0.0021	0.0011	0.0006	0.0001
	DynReg	0.0017	0.0017	0.0018	0.0018	0.0018	0.0019	0.0020
		(0; 0.0)	(1; 1.1)	(2; 2.1)	(3; 3.5)	(6; 5.7)	(6; 6.3)	(6; 6.5)
RE_{est}		0.9999	1.1476	1.2967	1.1807	0.6042	0.3105	0.0364
RE_{pred}		0.9990	1.0539	1.1404	1.2408	1.3901	1.4434	1.4886
T=2500								
		$\rho = 0$	$\rho = 0.3$	$\rho = 0.5$	$\rho = 0.7$	$\rho = 0.9$	$\rho = 0.95$	$\rho = 0.99$
Bias	HAC	-0.0001	-0.0001	-0.0001	0.0000	0.0000	-0.0001	0.0000
	DynReg	-0.0001	-0.0001	-0.0002	0.0001	-0.0002	0.0002	0.0001
Variance	HAC	0.0004	0.0005	0.0005	0.0005	0.0003	0.0001	0.0000
	DynReg	0.0004	0.0004	0.0004	0.0004	0.0004	0.0004	0.0004
MSE	HAC	0.0004	0.0005	0.0005	0.0005	0.0003	0.0001	0.0000
	DynReg	0.0004	0.0004	0.0004	0.0004	0.0004	0.0004	0.0004
		(0; 0.0)	(2; 1.7)	(3; 2.9)	(5; 5.0)	(10; 9.8)	(11; 11.6)	(12; 12.6)
RE_{est}		1.0000	1.1598	1.3240	1.2481	0.6137	0.3373	0.0525
RE_{pred}		0.9998	1.0391	1.1382	1.2496	1.4049	1.4281	1.4642

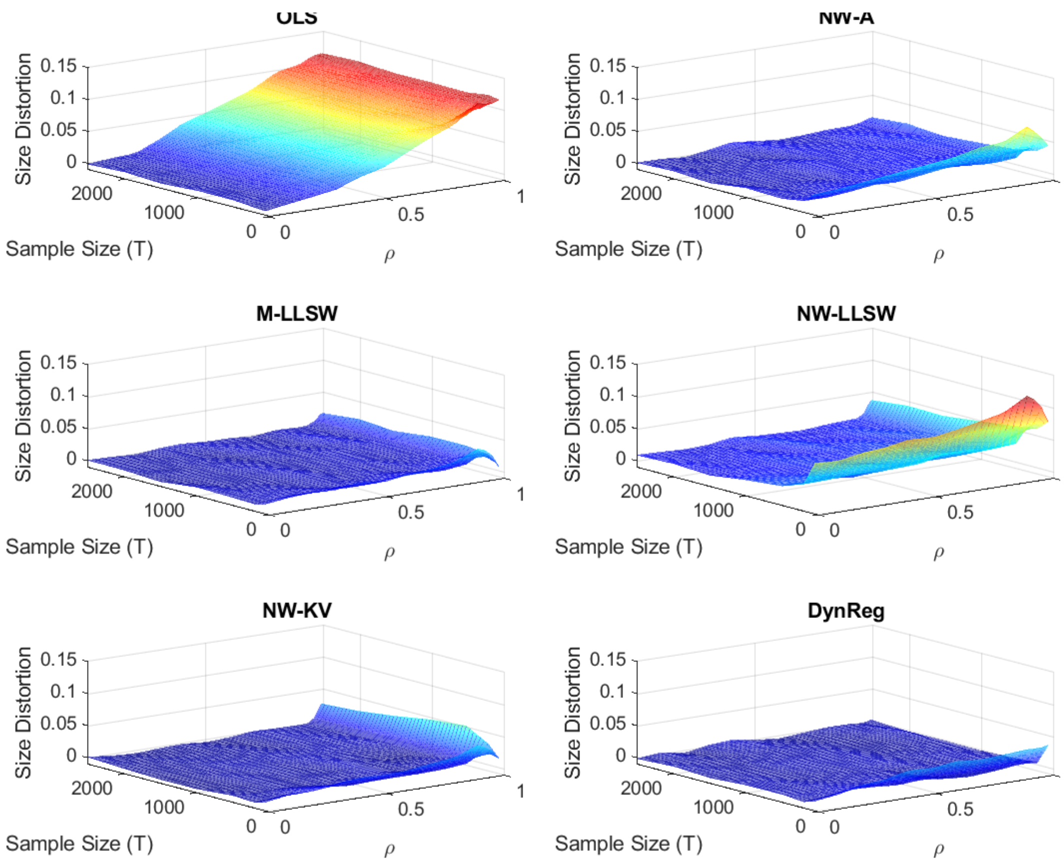
Notes: The data-generating process is $y_t = x_t + u_t$, $x_t = \rho x_{t-1} + \epsilon_{x,t}$, $u_t = \epsilon_{u,t} + \rho \epsilon_{u,t-1}$, $t = 1, ..., T$. All shocks are $N(0, 1)$, orthogonal at all leads and lags. RE_{est} is the estimation efficiency of DynReg relative to OLS: RE_{est}=MSE(OLS)/MSE(DynReg). RE_{pred} is the predictive efficiency of DynReg relative to OLS: RE_{pred}=MSPE(OLS)/MSPE(DynReg), where $MSPE$ is 1-step-ahead mean squared prediction error. We perform 10000 Monte Carlo replications. The median and mean BIC-selected DynReg lags appear in parentheses. See text for details.

Table A2: Empirical size of nominal 5% t-test of H_0: β=1
Moving Average Disturbances, BIC DynReg Lag-Order Selection

T=50								
	Truncation	$\rho = 0$	$\rho = 0.3$	$\rho = 0.5$	$\rho = 0.7$	$\rho = 0.9$	$\rho = 0.95$	$\rho = 0.99$
OLS	–	0.048	0.067	0.102	0.124	0.148	0.169	0.167
NW	$h = \lfloor 4[T/100]^{2/9} \rfloor + 1$	0.079	0.086	0.106	0.113	0.120	0.122	0.096
NW-A	$h = \lfloor 0.75T^{1/3} \rfloor + 1$	0.073	0.083	0.103	0.108	0.115	0.117	0.095
NW-LLSW	$h = \lfloor 1.3T^{1/2} \rfloor + 1$	0.118	0.121	0.140	0.145	0.160	0.167	0.130
NW-KV	$h = T$	0.053	0.061	0.070	0.075	0.086	0.089	0.064
M-LLSW	$\nu = \lfloor 0.41T^{2/3} \rfloor$	0.055	0.062	0.070	0.074	0.083	0.089	0.058
DynReg	–	0.055	0.071	0.083	0.078	0.086	0.087	0.083
T=200								
	Truncation	$\rho = 0$	$\rho = 0.3$	$\rho = 0.5$	$\rho = 0.7$	$\rho = 0.9$	$\rho = 0.95$	$\rho = 0.99$
OLS	–	0.049	0.070	0.098	0.129	0.148	0.162	0.166
NW	$h = \lfloor 4[T/100]^{2/9} \rfloor + 1$	0.060	0.064	0.070	0.074	0.077	0.086	0.080
NW-A	$h = \lfloor 0.75T^{1/3} \rfloor + 1$	0.060	0.064	0.070	0.074	0.077	0.086	0.080
NW-LLSW	$h = \lfloor 1.3T^{1/2} \rfloor + 1$	0.082	0.084	0.088	0.092	0.098	0.116	0.111
NW-KV	$h = T$	0.049	0.051	0.056	0.056	0.057	0.074	0.086
M-LLSW	$\nu = \lfloor 0.41T^{2/3} \rfloor$	0.051	0.054	0.055	0.060	0.062	0.079	0.074
DynReg	–	0.050	0.055	0.053	0.054	0.055	0.062	0.055
T=600								
	Truncation	$\rho = 0$	$\rho = 0.3$	$\rho = 0.5$	$\rho = 0.7$	$\rho = 0.9$	$\rho = 0.95$	$\rho = 0.99$
OLS	–	0.052	0.068	0.103	0.125	0.157	0.162	0.160
NW	$h = \lfloor 4[T/100]^{2/9} \rfloor + 1$	0.056	0.054	0.066	0.062	0.070	0.070	0.070
NW-A	$h = \lfloor 0.75T^{1/3} \rfloor + 1$	0.056	0.054	0.065	0.062	0.069	0.069	0.070
NW-LLSW	$h = \lfloor 1.3T^{1/2} \rfloor + 1$	0.069	0.065	0.073	0.070	0.080	0.085	0.101
NW-KV	$h = T$	0.047	0.047	0.050	0.051	0.054	0.058	0.103
M-LLSW	$\nu = \lfloor 0.41T^{2/3} \rfloor$	0.055	0.049	0.056	0.052	0.062	0.065	0.074
DynReg	–	0.052	0.048	0.056	0.052	0.051	0.051	0.052
T=2500								
	Truncation	$\rho = 0$	$\rho = 0.3$	$\rho = 0.5$	$\rho = 0.7$	$\rho = 0.9$	$\rho = 0.95$	$\rho = 0.99$
OLS	–	0.050	0.070	0.097	0.130	0.150	0.161	0.164
NW	$h = \lfloor 4[T/100]^{2/9} \rfloor + 1$	0.052	0.053	0.058	0.061	0.057	0.057	0.064
NW-A	$h = \lfloor 0.75T^{1/3} \rfloor + 1$	0.053	0.053	0.057	0.060	0.056	0.057	0.065
NW-LLSW	$h = \lfloor 1.3T^{1/2} \rfloor + 1$	0.057	0.058	0.060	0.062	0.059	0.064	0.086
NW-KV	$h = T$	0.050	0.048	0.049	0.053	0.047	0.055	0.074
M-LLSW	$\nu = \lfloor 0.41T^{2/3} \rfloor$	0.050	0.051	0.053	0.055	0.050	0.052	0.068
DynReg	–	0.050	0.050	0.048	0.049	0.049	0.049	0.049

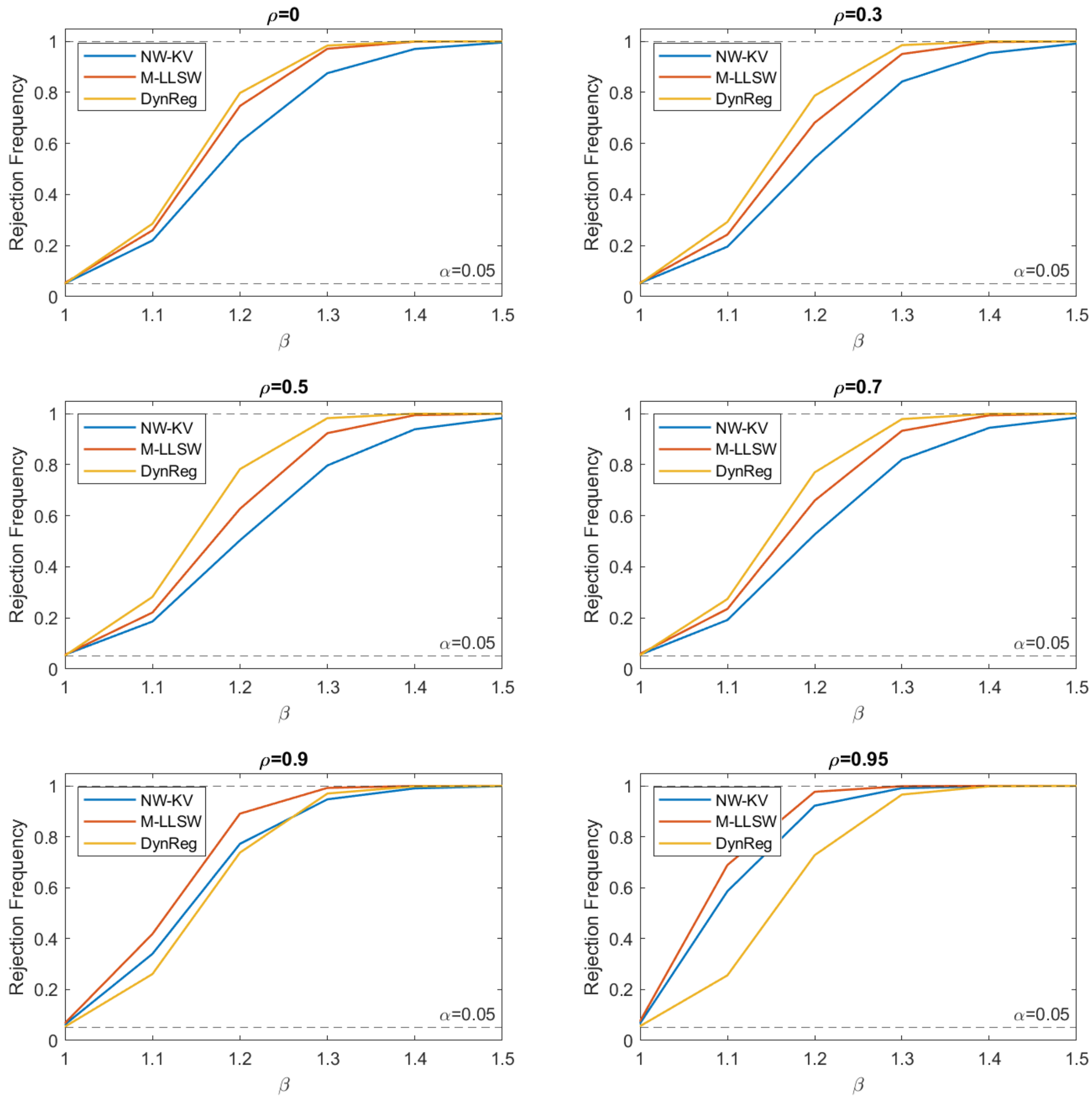
Notes: The data-generating process is $y_t = x_t + u_t$, $x_t = \rho x_{t-1} + \epsilon_{x,t}$, $u_t = \epsilon_{u,t} + \rho\epsilon_{u,t-1}$, $t = 1, ..., T$. All shocks are $N(0, 1)$, orthogonal at all leads and lags. We perform 10000 Monte Carlo replications. See text for details.

Figure A1: Empirical Size Distortion of Nominal 5% t-Test of H_0: β=1, Moving Average Disturbances, BIC DynReg Lag-Order Selection



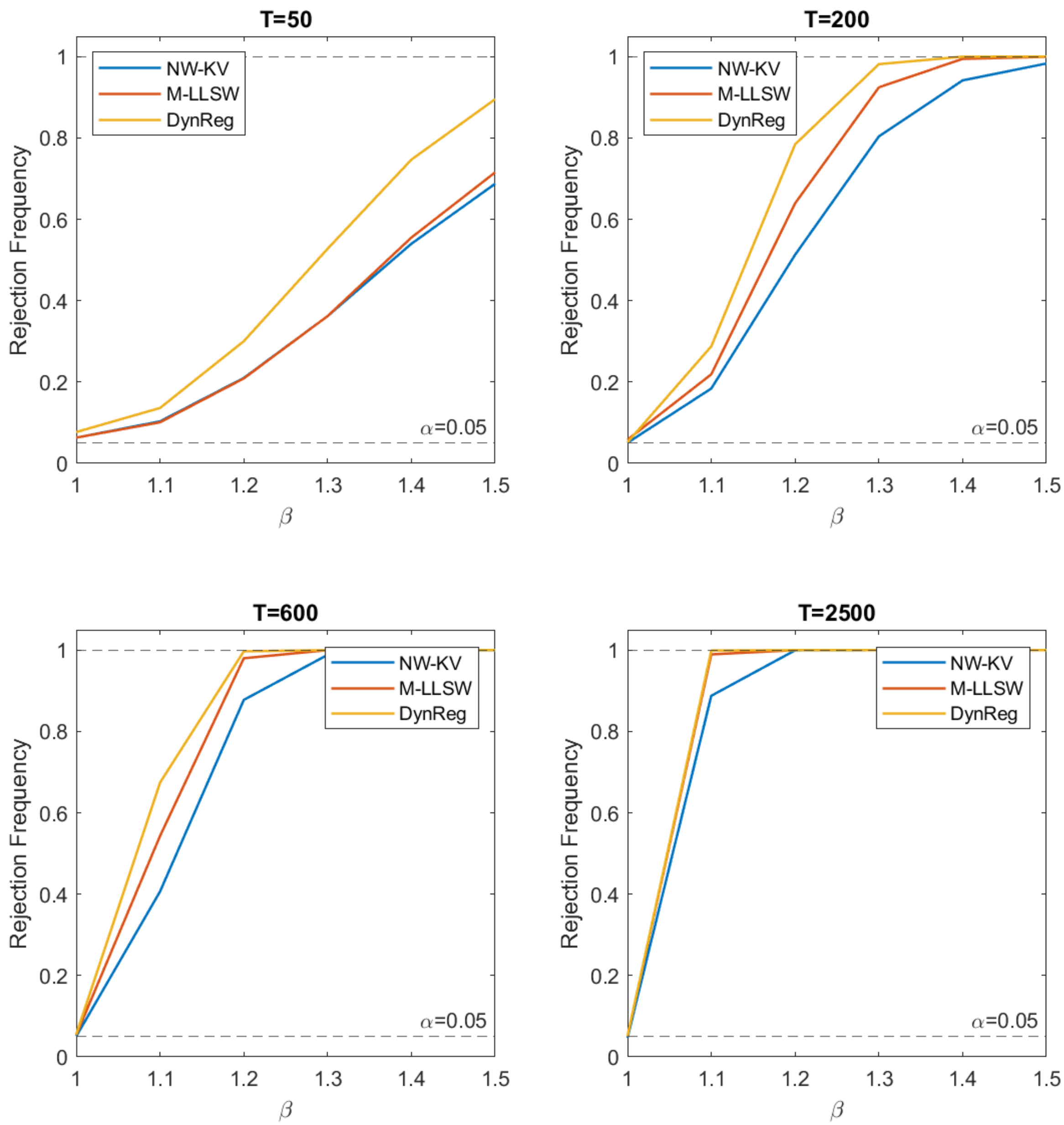
Notes: The data-generating process is $y_t = x_t + u_t$, $x_t = \rho x_{t-1} + \epsilon_{x,t}$, $u_t = \epsilon_{u,t} + \rho\epsilon_{u,t-1}$, $t = 1, ..., T$. All shocks are $N(0, 1)$, orthogonal at all leads and lags. DynReg lag order selected by BIC. We perform 10000 Monte Carlo replications, using common random numbers whenever appropriate, with x_0 and u_0 drawn from the stationary distribution. See text for details.

Figure A2: Empirical Rejection frequencies of Nominal 5% t-Test of H_0: β=1, Moving Average Disturbances, $T = 200$, BIC DynReg Lag-Order Selection



Notes: The data-generating process is $y_t = \beta x_t + u_t$, $x_t = \rho x_{t-1} + \epsilon_{x,t}$, $u_t = \epsilon_{u,t} + \rho\epsilon_{u,t-1}$, $t = 1, ..., T$. All shocks are $N(0, 1)$, orthogonal at all leads and lags. DynReg lag order selected by BIC. We perform 10000 Monte Carlo replications, using common random numbers whenever appropriate, with x_0 and u_0 drawn from the stationary distribution. See text for details.

Figure A3: Empirical Rejection frequencies of Nominal 5% t-Test of H_0: β=1, Moving Average Disturbances, $\rho = 0.5$, BIC DynReg Lag-Order Selection



Notes: The data-generating process is $y_t = \beta x_t + u_t$, $x_t = \rho x_{t-1} + \epsilon_{x,t}$, $u_t = \epsilon_{u,t} + \rho\epsilon_{u,t-1}$, $t = 1, ..., T$. All shocks are $N(0, 1)$, orthogonal at all leads and lags. DynReg lag order selected by BIC. We perform 10000 Monte Carlo replications, using common random numbers whenever appropriate, with x_0 and u_0 drawn from the stationary distribution. See text for details.

Table A3: Bias, Variance, MSE, and Efficiency of DynReg relative to OLS
Moving Average Disturbances, BIC DynReg Lag-Order Selection, Negative ρ

T=50								
		$\rho=0$	$\rho=-0.3$	$\rho=-0.5$	$\rho=-0.7$	$\rho=-0.9$	$\rho=-0.95$	$\rho=-0.99$
Bias	HAC	-0.0008	0.0007	0.0003	-0.0010	-0.0009	0.0001	-0.0001
	DynReg	-0.0002	0.0000	-0.0012	-0.0009	-0.0040	0.0004	0.0016
Variance	HAC	0.0110	0.0080	0.0071	0.0070	0.0078	0.0072	0.0032
	DynReg	0.0144	0.0183	0.0246	0.0284	0.0349	0.0379	0.0393
MSE	HAC	0.0110	0.0080	0.0071	0.0070	0.0078	0.0072	0.0032
	DynReg	0.0144	0.0183	0.0246	0.0284	0.0349	0.0379	0.0393
		(0; 0.3)	(0; 0.8)	(1; 1.4)	(2; 2.2)	(2; 3.2)	(2; 3.4)	(2; 3.5)
RE_{est}		0.7630	0.4398	0.2905	0.2473	0.2244	0.1902	0.0826
T=200								
		$\rho=0$	$\rho=-0.3$	$\rho=-0.5$	$\rho=-0.7$	$\rho=-0.9$	$\rho=-0.95$	$\rho=-0.99$
Bias	HAC	-0.0003	-0.0001	0.0004	-0.0007	-0.0004	0.0002	0.0004
	DynReg	-0.0002	0.0000	0.0007	-0.0013	-0.0003	0.0002	-0.0007
Variance	HAC	0.0027	0.0018	0.0015	0.0014	0.0016	0.0017	0.0011
	DynReg	0.0029	0.0049	0.0055	0.0057	0.0061	0.0064	0.0066
MSE	HAC	0.0027	0.0018	0.0015	0.0014	0.0016	0.0017	0.0011
	DynReg	0.0029	0.0049	0.0055	0.0057	0.0061	0.0064	0.0066
		(0; 0.0)	(1; 0.9)	(1; 1.5)	(2; 2.5)	(4; 3.6)	(4; 3.8)	(4; 3.9)
RE_{est}		0.9314	0.3566	0.2752	0.2501	0.2602	0.2607	0.1719
T=600								
		$\rho=0$	$\rho=-0.3$	$\rho=-0.5$	$\rho=-0.7$	$\rho=-0.9$	$\rho=-0.95$	$\rho=-0.99$
Bias	HAC	-0.0004	-0.0002	-0.0002	-0.0005	-0.0003	0.0000	-0.0001
	DynReg	-0.0005	0.0000	-0.0001	-0.0001	-0.0005	-0.0001	0.0004
Variance	HAC	0.0009	0.0006	0.0005	0.0004	0.0005	0.0005	0.0005
	DynReg	0.0009	0.0017	0.0018	0.0017	0.0018	0.0019	0.0020
MSE	HAC	0.0009	0.0006	0.0005	0.0004	0.0005	0.0005	0.0005
	DynReg	0.0009	0.0017	0.0018	0.0017	0.0018	0.0019	0.0020
		(0; 0.0)	(1; 1.1)	(2; 2.1)	(3; 3.5)	(6; 5.7)	(6; 6.3)	(6; 6.5)
RE_{est}		0.9752	0.3393	0.2725	0.2526	0.2683	0.2832	0.2396
T=2500								
		$\rho=0$	$\rho=-0.3$	$\rho=-0.5$	$\rho=-0.7$	$\rho=-0.9$	$\rho=-0.95$	$\rho=-0.99$
Bias	HAC	0.0000	0.0000	-0.0001	0.0001	-0.0002	0.0001	0.0001
	DynReg	0.0000	0.0000	-0.0002	0.0001	-0.0002	0.0001	-0.0001
Variance	HAC	0.0002	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001
	DynReg	0.0002	0.0004	0.0004	0.0004	0.0004	0.0004	0.0004
MSE	HAC	0.0002	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001
	DynReg	0.0002	0.0004	0.0004	0.0004	0.0004	0.0004	0.0004
		(0; 0.0)	(2; 1.7)	(3; 2.9)	(5; 5.0)	(10; 9.8)	(11; 11.6)	(12; 12.6)
RE_{est}		0.9916	0.3406	0.2841	0.2516	0.2722	0.2849	0.3179

Notes: The data-generating process is $y_t = x_t + u_t$, $x_t = 0.7x_{t-1} + \epsilon_{x,t}$, $u_t = \epsilon_{u,t} + \rho\epsilon_{u,t-1}$, $t = 1, ..., T$. All shocks are $N(0, 1)$, orthogonal at all leads and lags. RE_{est} is the estimation efficiency of DynReg relative to OLS: RE_{est}=MSE(OLS)/MSE(DynReg). We perform 10000 Monte Carlo replications. The median and mean BIC-selected DynReg lags appear in parentheses. See text for details.

Table A4: Empirical size of nominal 5% t-test of H_0: β=1
Moving Average Disturbances, BIC DynReg Lag-Order Selection, Negative ρ

T=50								
	Truncation Lag	$\rho=0$	$\rho=-0.3$	$\rho=-0.5$	$\rho=-0.7$	$\rho=-0.9$	$\rho=-0.95$	$\rho=-0.99$
OLS	–	0.046	0.014	0.007	0.003	0.005	0.010	0.033
NW	$h=\lfloor 4[T/100]^{2/9}\rfloor+1$	0.080	0.056	0.043	0.030	0.034	0.058	0.199
NW-A	$h=\lfloor 0.75T^{1/3}\rfloor+1$	0.075	0.044	0.034	0.021	0.025	0.042	0.146
NW-LLSW	$h=\lfloor 1.3T^{1/2}\rfloor+1$	0.118	0.101	0.087	0.070	0.077	0.125	0.394
NW-KV	$h=T$	0.062	0.049	0.040	0.029	0.030	0.055	0.285
M-LLSW	$\nu=\lfloor 0.41T^{2/3}\rfloor$	0.061	0.057	0.054	0.042	0.046	0.088	0.440
DynReg	–	0.059	0.051	0.065	0.068	0.075	0.081	0.083

T=200								
	Truncation Lag	$\rho=0$	$\rho=-0.3$	$\rho=-0.5$	$\rho=-0.7$	$\rho=-0.9$	$\rho=-0.95$	$\rho=-0.99$
OLS	–	0.052	0.011	0.004	0.002	0.002	0.004	0.028
NW	$h=\lfloor 4[T/100]^{2/9}\rfloor+1$	0.063	0.040	0.034	0.023	0.024	0.031	0.167
NW-A	$h=\lfloor 0.75T^{1/3}\rfloor+1$	0.063	0.040	0.034	0.023	0.024	0.031	0.167
NW-LLSW	$h=\lfloor 1.3T^{1/2}\rfloor+1$	0.087	0.074	0.067	0.059	0.059	0.077	0.339
NW-KV	$h=T$	0.053	0.048	0.046	0.040	0.040	0.043	0.252
M-LLSW	$\nu=\lfloor 0.41T^{2/3}\rfloor$	0.060	0.050	0.049	0.045	0.043	0.053	0.325
DynReg	–	0.056	0.049	0.053	0.055	0.055	0.058	0.055

T=600								
	Truncation Lag	$\rho=0$	$\rho=-0.3$	$\rho=-0.5$	$\rho=-0.7$	$\rho=-0.9$	$\rho=-0.95$	$\rho=-0.99$
OLS	–	0.053	0.013	0.004	0.001	0.001	0.001	0.018
NW	$h=\lfloor 4[T/100]^{2/9}\rfloor+1$	0.056	0.042	0.034	0.023	0.023	0.026	0.114
NW-A	$h=\lfloor 0.75T^{1/3}\rfloor+1$	0.057	0.044	0.037	0.026	0.025	0.030	0.127
NW-LLSW	$h=\lfloor 1.3T^{1/2}\rfloor+1$	0.070	0.063	0.065	0.050	0.055	0.061	0.226
NW-KV	$h=T$	0.050	0.048	0.049	0.041	0.043	0.042	0.151
M-LLSW	$\nu=\lfloor 0.41T^{2/3}\rfloor$	0.053	0.049	0.053	0.043	0.044	0.047	0.193
DynReg	–	0.054	0.048	0.056	0.050	0.050	0.050	0.054

T=2500								
	Truncation Lag	$\rho=0$	$\rho=-0.3$	$\rho=-0.5$	$\rho=-0.7$	$\rho=-0.9$	$\rho=-0.95$	$\rho=-0.99$
OLS	–	0.051	0.011	0.004	0.001	0.000	0.001	0.007
NW	$h=\lfloor 4[T/100]^{2/9}\rfloor+1$	0.053	0.042	0.037	0.030	0.027	0.030	0.069
NW-A	$h=\lfloor 0.75T^{1/3}\rfloor+1$	0.053	0.043	0.039	0.032	0.030	0.033	0.076
NW-LLSW	$h=\lfloor 1.3T^{1/2}\rfloor+1$	0.058	0.057	0.054	0.051	0.053	0.056	0.117
NW-KV	$h=T$	0.049	0.046	0.046	0.047	0.048	0.046	0.073
M-LLSW	$\nu=\lfloor 0.41T^{2/3}\rfloor$	0.052	0.050	0.049	0.046	0.047	0.047	0.085
DynReg	–	0.051	0.049	0.047	0.050	0.047	0.050	0.049

Notes: The data-generating process is $y_t = x_t + u_t$, $x_t = 0.7x_{t-1} + \epsilon_{x,t}$, $u_t = \epsilon_{u,t} + \rho\epsilon_{u,t-1}$, $t = 1, ..., T$. All shocks are $N(0,1)$, orthogonal at all leads and lags. We perform 10000 Monte Carlo replications. See text for details.

B Additional Monte Carlo: AR Errors, AIC Selection

Table B1: Bias, Variance, MSE, and Efficiency of DynReg relative to OLS
Autoregressive Disturbances, AIC DynReg Lag-Order Selection

T=50								
		$\rho = 0$	$\rho = 0.3$	$\rho = 0.5$	$\rho = 0.7$	$\rho = 0.9$	$\rho = 0.95$	$\rho = 0.99$
Bias	HAC	0.0007	-0.0002	0.0005	-0.0007	-0.0029	0.0026	-0.0292
	DynReg	0.0024	-0.0003	-0.0025	-0.0039	-0.0061	0.0006	0.0003
Variance	HAC	0.0204	0.0248	0.0343	0.0587	0.2348	0.8037	9.6973
	DynReg	0.0434	0.0453	0.0464	0.0453	0.0461	0.0492	0.0475
MSE	HAC	0.0204	0.0248	0.0343	0.0587	0.2348	0.8036	9.6972
	DynReg	0.0434	0.0453	0.0464	0.0453	0.0461	0.0492	0.0475
		(9; 6.6)	(9; 6.9)	(9; 7.1)	(9; 7.3)	(9; 7.4)	(9; 7.4)	(9; 7.4)
RE_{est}		0.4707	0.5481	0.7395	1.2968	5.0925	16.3235	204.1635
RE_{pred}		0.9893	1.0417	1.1605	1.4520	3.0330	5.8646	391.9083
T=200								
		$\rho = 0$	$\rho = 0.3$	$\rho = 0.5$	$\rho = 0.7$	$\rho = 0.9$	$\rho = 0.95$	$\rho = 0.99$
Bias	HAC	0.0001	0.0000	0.0010	-0.0009	-0.0020	0.0024	0.0046
	DynReg	-0.0002	-0.0001	0.0007	-0.0011	-0.0009	0.0005	-0.0003
Variance	HAC	0.0051	0.0061	0.0085	0.0145	0.0514	0.1526	2.7592
	DynReg	0.0071	0.0072	0.0071	0.0073	0.0071	0.0073	0.0071
MSE	HAC	0.0051	0.0061	0.0085	0.0145	0.0514	0.1526	2.7589
	DynReg	0.0071	0.0072	0.0071	0.0073	0.0071	0.0073	0.0071
		(8; 13.2)	(10; 14.2)	(10; 14.2)	(10; 14.1)	(11; 14.4)	(12; 14.5)	(13; 14.7)
RE_{est}		0.7222	0.8407	1.1840	1.9882	7.2732	20.8362	388.0916
RE_{pred}		0.9969	1.0510	1.1678	1.4758	3.1211	5.6978	47.3609
T=600								
		$\rho = 0$	$\rho = 0.3$	$\rho = 0.5$	$\rho = 0.7$	$\rho = 0.9$	$\rho = 0.95$	$\rho = 0.99$
Bias	HAC	-0.0008	-0.0002	-0.0004	0.0004	-0.0007	-0.0028	-0.0062
	DynReg	-0.0009	-0.0001	0.0000	-0.0001	-0.0005	-0.0003	0.0003
Variance	HAC	0.0017	0.0020	0.0029	0.0049	0.0168	0.0431	0.9138
	DynReg	0.0017	0.0017	0.0018	0.0017	0.0017	0.0017	0.0017
MSE	HAC	0.0017	0.0020	0.0029	0.0049	0.0168	0.0431	0.9138
	DynReg	0.0017	0.0017	0.0018	0.0017	0.0017	0.0017	0.0017
		(1; 4.0)	(2; 5.2)	(2; 5.1)	(2; 5.1)	(2; 5.0)	(2; 5.2)	(2; 5.0)
RE_{est}		0.9729	1.1644	1.6106	2.8058	9.8082	25.1653	523.3413
RE_{pred}		0.9996	1.0470	1.1524	1.5047	3.2144	5.6054	25.5608
T=2500								
		$\rho = 0$	$\rho = 0.3$	$\rho = 0.5$	$\rho = 0.7$	$\rho = 0.9$	$\rho = 0.95$	$\rho = 0.99$
Bias	HAC	-0.0001	-0.0001	0.0001	0.0001	0.0004	-0.0004	-0.0005
	DynReg	-0.0001	0.0001	-0.0002	0.0001	-0.0003	0.0001	-0.0001
Variance	HAC	0.0004	0.0005	0.0007	0.0012	0.0039	0.0087	0.1959
	DynReg	0.0004	0.0004	0.0004	0.0004	0.0004	0.0004	0.0004
MSE	HAC	0.0004	0.0005	0.0007	0.0012	0.0039	0.0087	0.1958
	DynReg	0.0004	0.0004	0.0004	0.0004	0.0004	0.0004	0.0004
		(1; 2.9)	(2; 4.0)	(2; 3.9)	(2; 4.0)	(1; 3.8)	(1; 3.8)	(1; 3.8)
RE_{est}		0.9933	1.2016	1.6800	2.9048	9.6350	21.7653	490.6450
RE_{pred}		0.9991	1.0493	1.1625	1.4686	3.1010	5.6562	25.6477

Notes: The data-generating process is $y_t = x_t + u_t$, $x_t = \rho x_{t-1} + \epsilon_{x,t}$, $u_t = \rho u_{t-1} + \epsilon_{u,t}$, $t = 1, ..., T$. All shocks are $N(0,1)$, orthogonal at all leads and lags. RE_{est} is the estimation efficiency of DynReg relative to OLS: RE_{est}=MSE(OLS)/MSE(DynReg). RE_{pred} is the predictive efficiency of DynReg relative to OLS: RE_{pred}=MSPE(OLS)/MSPE(DynReg), where $MSPE$ is 1-step-ahead mean squared prediction error. We perform 10000 Monte Carlo replications. The median and mean AIC-selected DynReg lags appear in parentheses. See text for details.

Table B2: Empirical size of nominal 5% t-test of H_0: β=1
Autoregressive Disturbances, AIC DynReg Lag-Order Selection

T=50

	Truncation	$\rho = 0$	$\rho = 0.3$	$\rho = 0.5$	$\rho = 0.7$	$\rho = 0.9$	$\rho = 0.95$	$\rho = 0.99$
OLS	–	0.048	0.072	0.127	0.242	0.564	0.774	0.970
NW	$h = \lfloor 4[T/100]^{2/9} \rfloor + 1$	0.079	0.089	0.119	0.169	0.421	0.674	0.953
NW-A	$h = \lfloor 0.75T^{1/3} \rfloor + 1$	0.073	0.083	0.116	0.180	0.452	0.702	0.958
NW-LLSW	$h = \lfloor 1.3T^{1/2} \rfloor + 1$	0.118	0.123	0.149	0.179	0.359	0.603	0.938
NW-KV	$h = T$	0.053	0.061	0.072	0.092	0.189	0.388	0.879
M-LLSW	$\nu = \lfloor 0.41T^{2/3} \rfloor$	0.055	0.064	0.073	0.080	0.197	0.455	0.914
DynReg	–	0.074	0.078	0.078	0.073	0.076	0.084	0.079

T=200

	Truncation	$\rho = 0$	$\rho = 0.3$	$\rho = 0.5$	$\rho = 0.7$	$\rho = 0.9$	$\rho = 0.95$	$\rho = 0.99$
OLS	–	0.049	0.073	0.135	0.247	0.544	0.727	0.955
NW	$h = \lfloor 4[T/100]^{2/9} \rfloor + 1$	0.060	0.063	0.081	0.114	0.280	0.498	0.916
NW-A	$h = \lfloor 0.75T^{1/3} \rfloor + 1$	0.060	0.063	0.081	0.114	0.280	0.498	0.916
NW-LLSW	$h = \lfloor 1.3T^{1/2} \rfloor + 1$	0.082	0.085	0.095	0.107	0.176	0.319	0.858
NW-KV	$h = T$	0.049	0.052	0.058	0.061	0.089	0.136	0.671
M-LLSW	$\nu = \lfloor 0.41T^{2/3} \rfloor$	0.051	0.054	0.056	0.067	0.113	0.238	0.847
DynReg	–	0.062	0.061	0.062	0.067	0.060	0.064	0.061

T=600

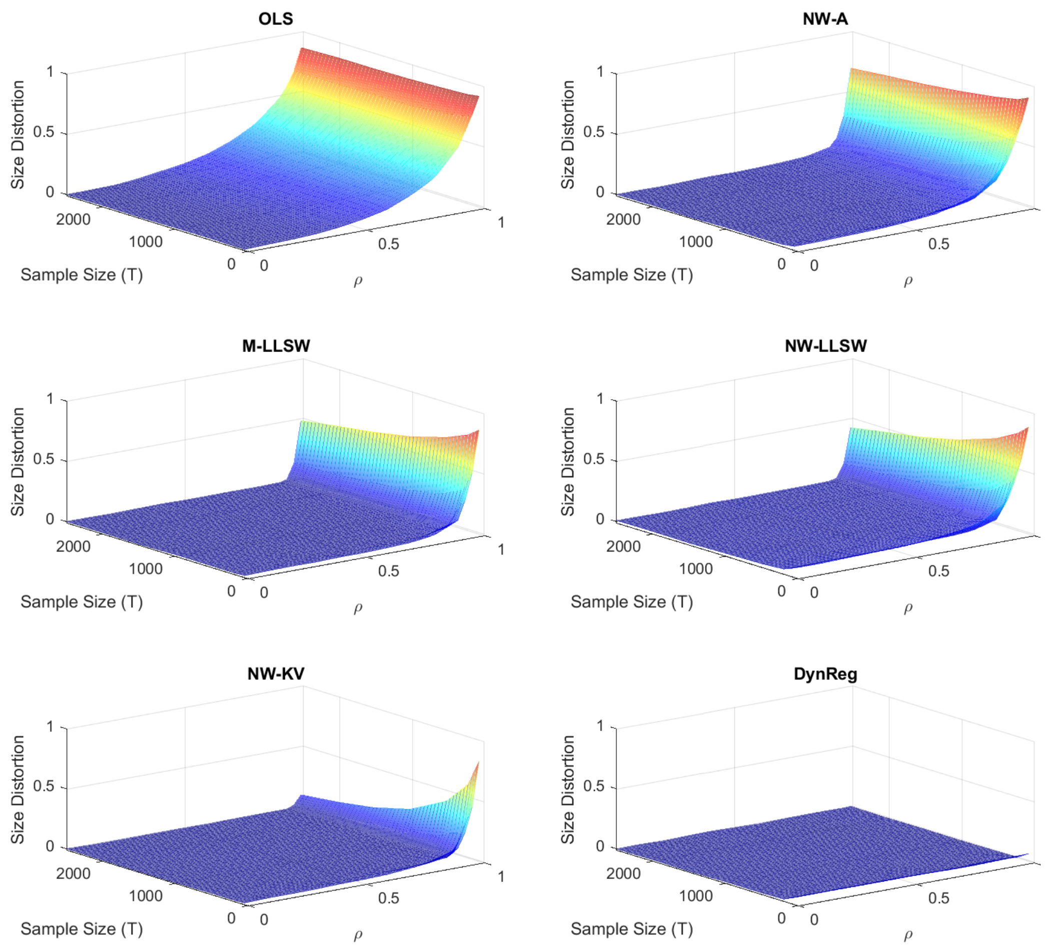
	Truncation	$\rho = 0$	$\rho = 0.3$	$\rho = 0.5$	$\rho = 0.7$	$\rho = 0.9$	$\rho = 0.95$	$\rho = 0.99$
OLS	–	0.052	0.071	0.137	0.248	0.536	0.688	0.944
NW	$h = \lfloor 4[T/100]^{2/9} \rfloor + 1$	0.056	0.055	0.074	0.092	0.219	0.380	0.878
NW-A	$h = \lfloor 0.75T^{1/3} \rfloor + 1$	0.056	0.055	0.072	0.087	0.199	0.353	0.868
NW-LLSW	$h = \lfloor 1.3T^{1/2} \rfloor + 1$	0.069	0.066	0.074	0.080	0.116	0.171	0.746
NW-KV	$h = T$	0.047	0.045	0.050	0.053	0.061	0.069	0.429
M-LLSW	$\nu = \lfloor 0.41T^{2/3} \rfloor$	0.055	0.050	0.056	0.058	0.083	0.136	0.756
DynReg	–	0.055	0.049	0.058	0.052	0.052	0.050	0.054

T=2500

	Truncation	$\rho = 0$	$\rho = 0.3$	$\rho = 0.5$	$\rho = 0.7$	$\rho = 0.9$	$\rho = 0.95$	$\rho = 0.99$
OLS	–	0.050	0.075	0.128	0.249	0.535	0.673	0.924
NW	$h = \lfloor 4[T/100]^{2/9} \rfloor + 1$	0.052	0.053	0.061	0.074	0.150	0.263	0.771
NW-A	$h = \lfloor 0.75T^{1/3} \rfloor + 1$	0.053	0.054	0.060	0.070	0.129	0.229	0.748
NW-LLSW	$h = \lfloor 1.3T^{1/2} \rfloor + 1$	0.057	0.059	0.061	0.064	0.069	0.095	0.475
NW-KV	$h = T$	0.050	0.046	0.051	0.049	0.052	0.052	0.144
M-LLSW	$\nu = \lfloor 0.41T^{2/3} \rfloor$	0.050	0.051	0.053	0.056	0.056	0.082	0.535
DynReg	–	0.050	0.051	0.047	0.050	0.049	0.048	0.049

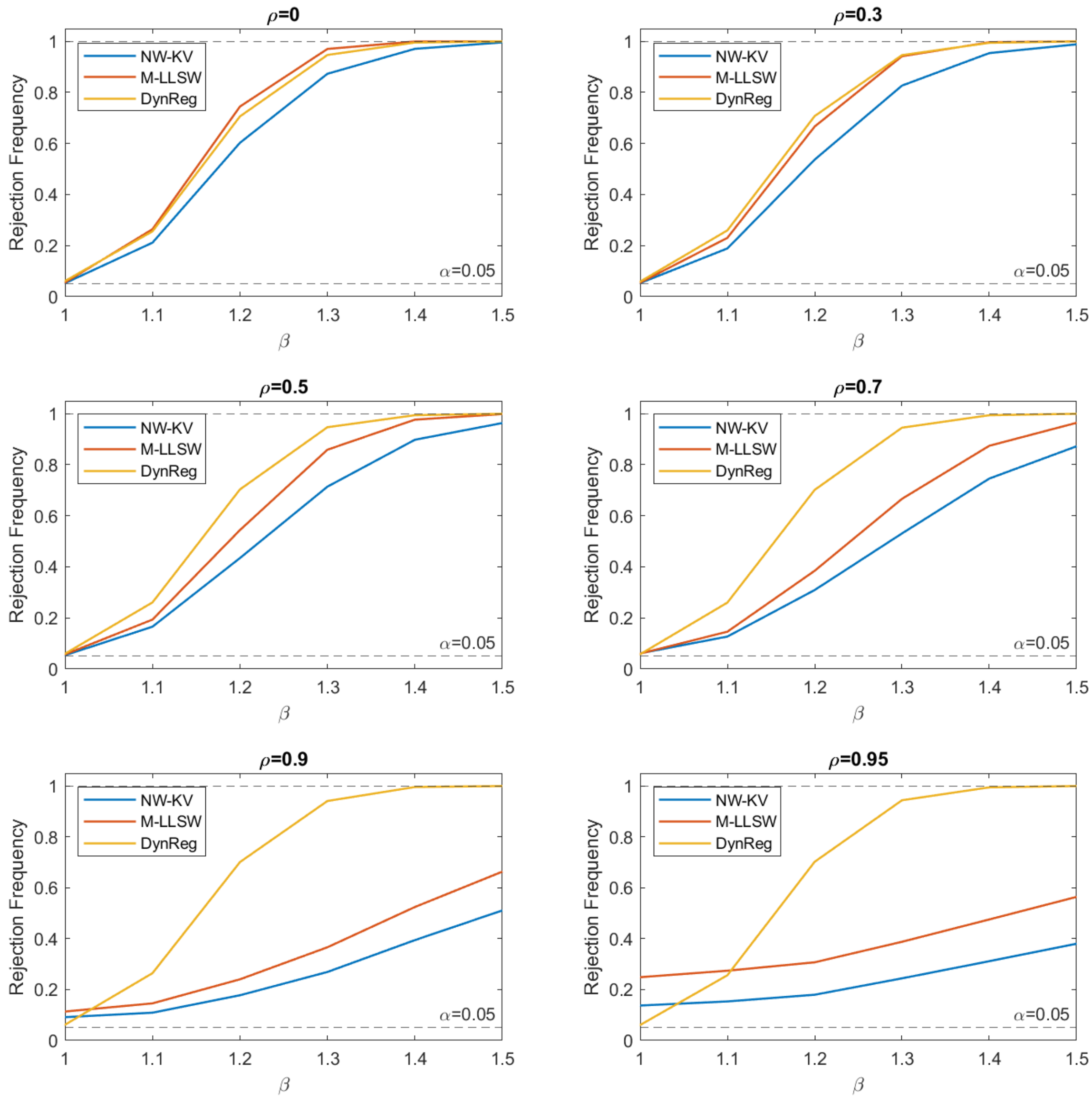
Notes: The data-generating process is $y_t = x_t + u_t$, $x_t = \rho x_{t-1} + \epsilon_{x,t}$, $u_t = \rho u_{t-1} + \epsilon_{u,t}$, $t = 1, ..., T$. All shocks are $N(0, 1)$, orthogonal at all leads and lags. We perform 10000 Monte Carlo replications. See text for details.

Figure B1: Empirical Size Distortion of Nominal 5% t-Test of H_0: β=1, Autoregressive Disturbances, AIC DynReg Lag-Order Selection



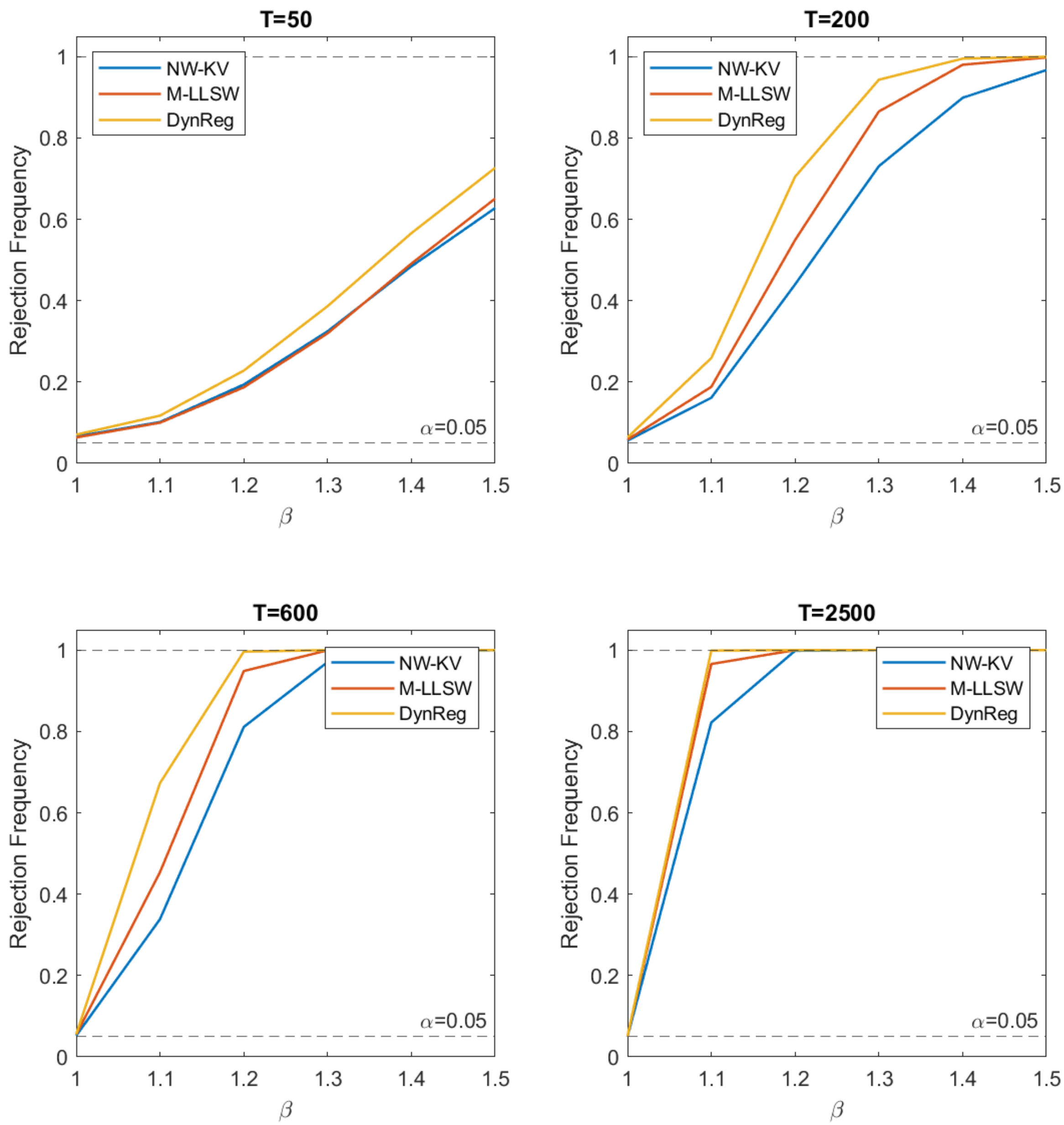
Notes: The data-generating process is $y_t = x_t + u_t$, $x_t = \rho x_{t-1} + \epsilon_{x,t}$, $u_t = \rho u_{t-1} + \epsilon_{u,t}$, $t = 1, ..., T$. All shocks are $N(0, 1)$, orthogonal at all leads and lags. DynReg lag order selected by AIC. We perform 10000 Monte Carlo replications, using common random numbers whenever appropriate, with x_0 and u_0 drawn from the stationary distribution. See text for details.

Figure B2: Empirical Rejection frequencies of Nominal 5% t-Test of H_0: β=1, Autoregressive Disturbances, $T = 200$, AIC DynReg Lag-Order Selection



Notes: The data-generating process is $y_t = \beta x_t + u_t$, $x_t = \rho x_{t-1} + \epsilon_{x,t}$, $u_t = \rho u_{t-1} + \epsilon_{u,t}$, $t = 1, ..., T$. All shocks are $N(0, 1)$, orthogonal at all leads and lags. DynReg lag order selected by AIC. We perform 10000 Monte Carlo replications, using common random numbers whenever appropriate, with x_0 and u_0 drawn from the stationary distribution. See text for details.

Figure B3: Empirical Rejection frequencies of Nominal 5% t-Test of H_0: β=1, Autoregressive Disturbances, $\rho = 0.5$, AIC DynReg Lag-Order Selection



Notes: The data-generating process is $y_t = \beta x_t + u_t$, $x_t = \rho x_{t-1} + \epsilon_{x,t}$, $u_t = \rho u_{t-1} + \epsilon_{u,t}$, $t = 1, ..., T$. All shocks are $N(0, 1)$, orthogonal at all leads and lags. DynReg lag order selected by BIC. We perform 10000 Monte Carlo replications, using common random numbers whenever appropriate, with x_0 and u_0 drawn from the stationary distribution. See text for details.

C Additional Monte Carlo: MA Errors, AIC Selection

Table C1: Bias, Variance, MSE, and Efficiency of DynReg relative to OLS
Moving Average Disturbances, AIC DynReg Lag-Order Selection

		$\rho = 0$	$\rho = 0.3$	$\rho = 0.5$	$\rho = 0.7$	$\rho = 0.9$	$\rho = 0.95$	$\rho = 0.99$
T=50								
Bias	HAC	0.0007	-0.0002	0.0002	0.0005	-0.0016	-0.0008	-0.0004
	DynReg	0.0024	-0.0010	-0.0020	-0.0028	-0.0062	-0.0001	0.0009
Variance	HAC	0.0204	0.0240	0.0277	0.0269	0.0137	0.0070	0.0010
	DynReg	0.0434	0.0456	0.0472	0.0467	0.0502	0.0558	0.0551
MSE	HAC	0.0204	0.0240	0.0277	0.0269	0.0137	0.0070	0.0010
	DynReg	0.0434	0.0456	0.0472	0.0467	0.0502	0.0558	0.0551
		(9; 6.6)	(9; 7.0)	(9; 7.4)	(9; 7.8)	(10; 8.5)	(10; 8.7)	(10; 8.7)
RE_{est}		0.4707	0.5252	0.5864	0.5760	0.2733	0.1254	0.0185
RE_{pred}		0.9899	1.0445	1.1001	1.2495	1.4047	1.4208	1.4723
T=200								
		$\rho = 0$	$\rho = 0.3$	$\rho = 0.5$	$\rho = 0.7$	$\rho = 0.9$	$\rho = 0.95$	$\rho = 0.99$
Bias	HAC	0.0001	0.0001	0.0009	-0.0007	-0.0002	-0.0001	0.0000
	DynReg	-0.0002	-0.0001	0.0005	-0.0006	-0.0010	0.0002	-0.0005
Variance	HAC	0.0051	0.0058	0.0067	0.0064	0.0032	0.0017	0.0002
	DynReg	0.0071	0.0072	0.0073	0.0076	0.0081	0.0085	0.0085
MSE	HAC	0.0051	0.0058	0.0067	0.0064	0.0032	0.0017	0.0002
	DynReg	0.0071	0.0072	0.0073	0.0076	0.0081	0.0085	0.0085
		(8; 13.2)	(11; 14.6)	(15; 15.7)	(20; 17.6)	(25; 21.7)	(27; 23.4)	(28; 24.3)
RE_{est}		0.7222	0.8077	0.9131	0.8420	0.3924	0.1947	0.0236
RE_{pred}		0.9993	1.0403	1.1297	1.2451	1.4090	1.4555	1.4831
T=600								
		$\rho = 0$	$\rho = 0.3$	$\rho = 0.5$	$\rho = 0.7$	$\rho = 0.9$	$\rho = 0.95$	$\rho = 0.99$
Bias	HAC	-0.0008	-0.0001	-0.0006	0.0001	-0.0004	-0.0002	0.0000
	DynReg	-0.0009	0.0000	0.0000	-0.0001	-0.0004	-0.0003	0.0002
Variance	HAC	0.0017	0.0019	0.0023	0.0021	0.0011	0.0006	0.0001
	DynReg	0.0017	0.0017	0.0018	0.0018	0.0018	0.0019	0.0020
MSE	HAC	0.0017	0.0019	0.0023	0.0021	0.0011	0.0006	0.0001
	DynReg	0.0017	0.0017	0.0018	0.0018	0.0018	0.0019	0.0020
		(1; 4.0)	(3; 6.0)	(4; 7.4)	(7; 10.1)	(16; 17.6)	(21; 21.3)	(24; 23.3)
RE_{est}		0.9729	1.1143	1.2583	1.1562	0.5966	0.3076	0.0356
RE_{pred}		0.9990	1.0539	1.1404	1.2408	1.3901	1.4434	1.4886
T=2500								
		$\rho = 0$	$\rho = 0.3$	$\rho = 0.5$	$\rho = 0.7$	$\rho = 0.9$	$\rho = 0.95$	$\rho = 0.99$
Bias	HAC	-0.0001	-0.0001	-0.0001	0.0000	0.0000	-0.0001	0.0000
	DynReg	-0.0001	0.0000	-0.0002	0.0001	-0.0003	0.0002	-0.0001
Variance	HAC	0.0004	0.0005	0.0005	0.0005	0.0003	0.0001	0.0000
	DynReg	0.0004	0.0004	0.0004	0.0004	0.0004	0.0004	0.0004
MSE	HAC	0.0004	0.0005	0.0005	0.0005	0.0003	0.0001	0.0000
	DynReg	0.0004	0.0004	0.0004	0.0004	0.0004	0.0004	0.0004
		(1; 2.9)	(3; 5.3)	(5; 7.1)	(9; 10.7)	(21; 21.4)	(27; 26.8)	(29; 28.6)
RE_{est}		0.9933	1.1532	1.3178	1.2469	0.6212	0.3429	0.0529
RE_{pred}		0.9998	1.0391	1.1382	1.2496	1.4049	1.4281	1.4642

Notes: The data-generating process is $y_t = x_t + u_t$, $x_t = \rho x_{t-1} + \epsilon_{x,t}$, $u_t = \epsilon_{u,t} + \rho\epsilon_{u,t-1}+$, $t = 1, ..., T$. All shocks are $N(0, 1)$, orthogonal at all leads and lags. RE_{est} is the estimation efficiency of DynReg relative to OLS: RE_{est}=MSE(OLS)/MSE(DynReg). RE_{pred} is the predictive efficiency of DynReg relative to OLS: RE_{pred}=MSPE(OLS)/MSPE(DynReg), where $MSPE$ is 1-step-ahead mean squared prediction error. We perform 10000 Monte Carlo replications. The median and mean AIC-selected DynReg lags appear in parentheses. See text for details.

Table C2: Empirical size of nominal 5% t-test of H_0: β=1
Moving Average Disturbances, AIC DynReg Lag-Order Selection

	T=50							
	Truncation	$\rho=0$	$\rho=0.3$	$\rho=0.5$	$\rho=0.7$	$\rho=0.9$	$\rho=0.95$	$\rho=0.99$
OLS	–	0.048	0.067	0.102	0.124	0.148	0.169	0.167
NW	$h=\lfloor 4[T/100]^{2/9}\rfloor+1$	0.079	0.086	0.106	0.113	0.120	0.122	0.096
NW-A	$h=\lfloor 0.75T^{1/3}\rfloor+1$	0.073	0.083	0.103	0.108	0.115	0.117	0.095
NW-LLSW	$h=\lfloor 1.3T^{1/2}\rfloor+1$	0.118	0.121	0.140	0.145	0.160	0.167	0.130
NW-KV	$h=T$	0.053	0.061	0.070	0.075	0.086	0.089	0.064
M-LLSW	$\nu=\lfloor 0.41T^{2/3}\rfloor$	0.055	0.062	0.070	0.074	0.083	0.089	0.058
DynReg	–	0.074	0.078	0.078	0.071	0.068	0.077	0.073

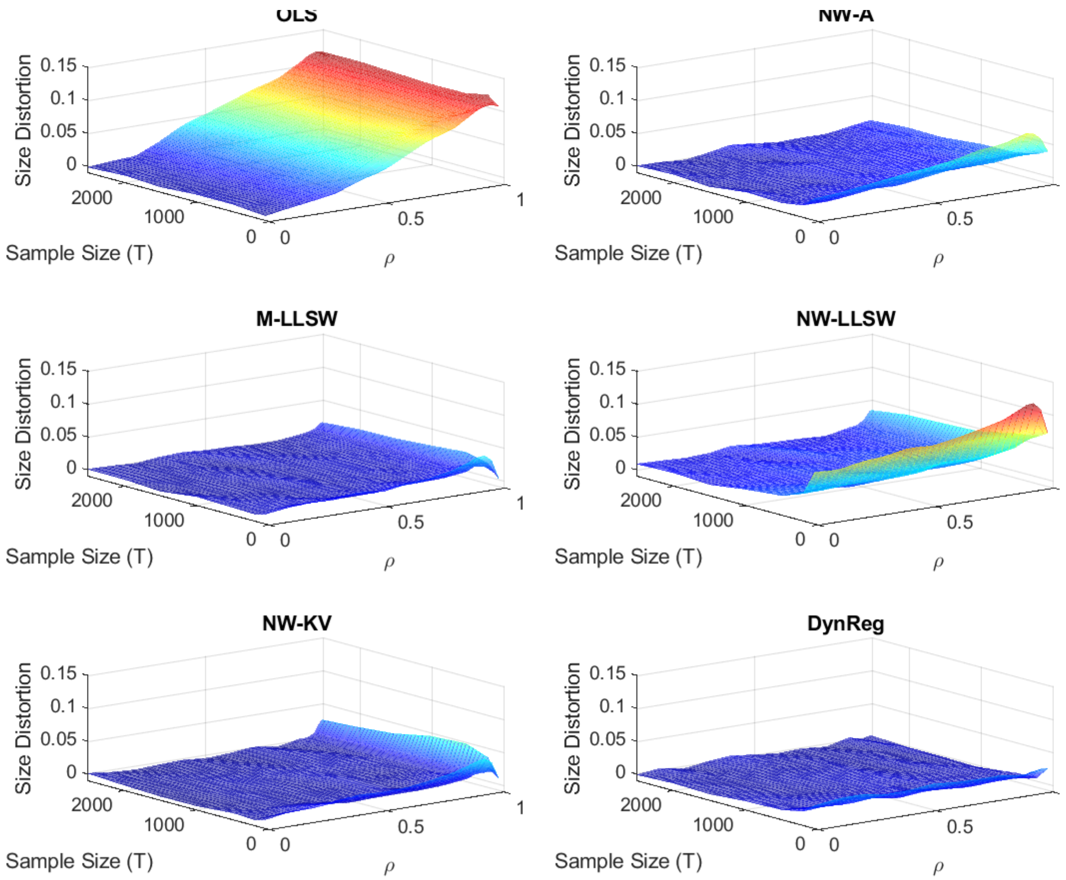
	T=200							
	Truncation	$\rho=0$	$\rho=0.3$	$\rho=0.5$	$\rho=0.7$	$\rho=0.9$	$\rho=0.95$	$\rho=0.99$
OLS	–	0.049	0.070	0.098	0.129	0.148	0.162	0.166
NW	$h=\lfloor 4[T/100]^{2/9}\rfloor+1$	0.060	0.064	0.070	0.074	0.077	0.086	0.080
NW-A	$h=\lfloor 0.75T^{1/3}\rfloor+1$	0.060	0.064	0.070	0.074	0.077	0.086	0.080
NW-LLSW	$h=\lfloor 1.3T^{1/2}\rfloor+1$	0.082	0.084	0.088	0.092	0.098	0.116	0.111
NW-KV	$h=T$	0.049	0.051	0.056	0.056	0.057	0.074	0.086
M-LLSW	$\nu=\lfloor 0.41T^{2/3}\rfloor$	0.051	0.054	0.055	0.060	0.062	0.079	0.074
DynReg	–	0.062	0.061	0.064	0.064	0.058	0.061	0.059

	T=600							
	Truncation	$\rho=0$	$\rho=0.3$	$\rho=0.5$	$\rho=0.7$	$\rho=0.9$	$\rho=0.95$	$\rho=0.99$
OLS	–	0.052	0.068	0.103	0.125	0.157	0.162	0.160
NW	$h=\lfloor 4[T/100]^{2/9}\rfloor+1$	0.056	0.054	0.066	0.062	0.070	0.070	0.070
NW-A	$h=\lfloor 0.75T^{1/3}\rfloor+1$	0.056	0.054	0.065	0.062	0.069	0.069	0.070
NW-LLSW	$h=\lfloor 1.3T^{1/2}\rfloor+1$	0.069	0.065	0.073	0.070	0.080	0.085	0.101
NW-KV	$h=T$	0.047	0.047	0.050	0.051	0.054	0.058	0.103
M-LLSW	$\nu=\lfloor 0.41T^{2/3}\rfloor$	0.055	0.049	0.056	0.052	0.062	0.065	0.074
DynReg	–	0.055	0.048	0.058	0.052	0.051	0.053	0.056

	T=2500							
	Truncation	$\rho=0$	$\rho=0.3$	$\rho=0.5$	$\rho=0.7$	$\rho=0.9$	$\rho=0.95$	$\rho=0.99$
OLS	–	0.050	0.070	0.097	0.130	0.150	0.161	0.164
NW	$h=\lfloor 4[T/100]^{2/9}\rfloor+1$	0.052	0.053	0.058	0.061	0.057	0.057	0.064
NW-A	$h=\lfloor 0.75T^{1/3}\rfloor+1$	0.053	0.053	0.057	0.060	0.056	0.057	0.065
NW-LLSW	$h=\lfloor 1.3T^{1/2}\rfloor+1$	0.057	0.058	0.060	0.062	0.059	0.064	0.086
NW-KV	$h=T$	0.050	0.048	0.049	0.053	0.047	0.055	0.074
M-LLSW	$\nu=\lfloor 0.41T^{2/3}\rfloor$	0.050	0.051	0.053	0.055	0.050	0.052	0.068
DynReg	–	0.050	0.051	0.048	0.050	0.048	0.049	0.049

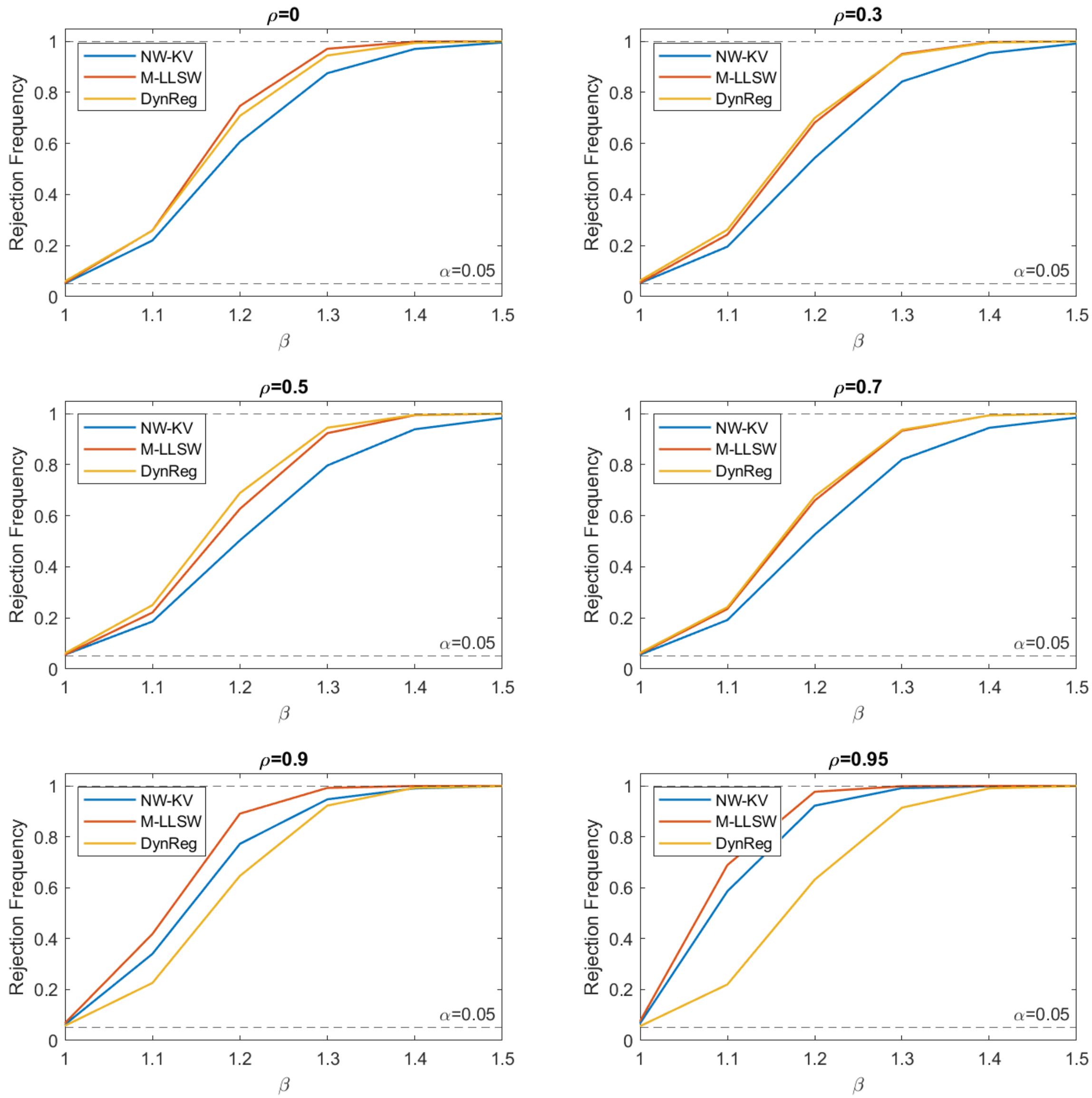
Notes: The data-generating process is $y_t = x_t + u_t$, $x_t = \rho x_{t-1} + \epsilon_{x,t}$, $u_t = \epsilon_{u,t} + \rho\epsilon_{u,t-1}$, $t = 1, ..., T$. All shocks are $N(0, 1)$, orthogonal at all leads and lags. We perform 10000 Monte Carlo replications. See text for details.

Figure C1: Empirical Size Distortion of Nominal 5% t-Test of H_0: β=1, Moving Average Disturbances, AIC DynReg Lag-Order Selection



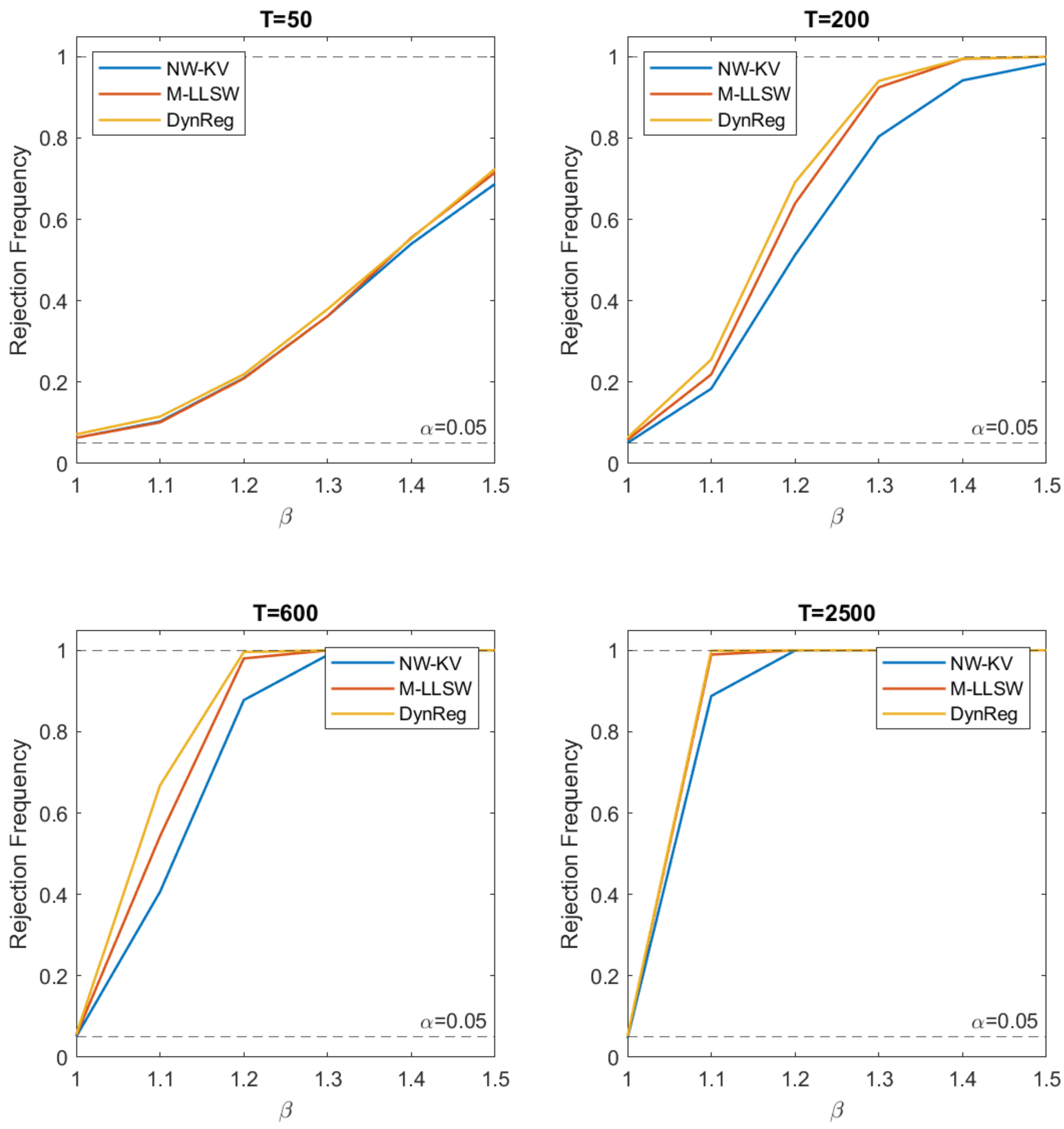
Notes: The data-generating process is $y_t = x_t + u_t$, $x_t = \rho x_{t-1} + \epsilon_{x,t}$, $u_t = \epsilon_{u,t} + \rho\epsilon_{u,t-1}$, $t = 1, ..., T$. All shocks are $N(0, 1)$, orthogonal at all leads and lags. DynReg lag order selected by AIC. We perform 10000 Monte Carlo replications, using common random numbers whenever appropriate, with x_0 and u_0 drawn from the stationary distribution. See text for details.

Figure C2: Empirical Rejection frequencies of Nominal 5% t-Test of H_0: β=1, Moving Average Disturbances, $T = 200$, AIC DynReg Lag-Order Selection



Notes: The data-generating process is $y_t = \beta x_t + u_t$, $x_t = \rho x_{t-1} + \epsilon_{x,t}$, $u_t = \epsilon_{u,t} + \rho\epsilon_{u,t-1}$, $t = 1, ..., T$. All shocks are $N(0, 1)$, orthogonal at all leads and lags. DynReg lag order selected by AIC. We perform 10000 Monte Carlo replications, using common random numbers whenever appropriate, with x_0 and u_0 drawn from the stationary distribution. See text for details.

Figure C3: Empirical Rejection frequencies of Nominal 5% t-Test of H_0: β=1, Moving Average Disturbances, $\rho = 0.5$, AIC DynReg Lag-Order Selection



Notes: The data-generating process is $y_t = \beta x_t + u_t$, $x_t = \rho x_{t-1} + \epsilon_{x,t}$, $u_t = \epsilon_{u,t} + \rho\epsilon_{u,t-1}$, $t = 1, ..., T$. All shocks are $N(0,1)$, orthogonal at all leads and lags. DynReg lag order selected by AIC. We perform 10000 Monte Carlo replications, using common random numbers whenever appropriate, with x_0 and u_0 drawn from the stationary distribution. See text for details.

References

Andrews, D.W.K. (1991), "Heteroskedasticity and Autocorrelation Consistent Covariance Matrix Estimation," *Econometrica*, 817–858.

Angrist, J.D. and J.-S. Pischke (2008), *Mostly Harmless Econometrics: An Empiricist's Companion*, Princeton University Press.

Den Haan, W.J. and A. Levin (1997), "A Practitioner's Guide to Robust Covariance Matrix Estimation," In G.S. Maddala and C.R. Rao (eds.), *Handbook of Statistics*, 15, 299-342.

Den Haan, W.J. and A. Levin (1998), "Vector Autoregressive Covariance Matrix Estimation," Working Paper, Board of Governors of the Federal Reserve System, `http://www.wouterdenhaan.com/papers/asym35.pdf`.

Den Haan, W.J. and A. Levin (2000), "Robust Covariance Matrix Estimation with Data-Dependent VAR Prewhitening Order," `http://www.wouterdenhaan.com/papers/T0255.pdf`.

Grenander, U. (1981), *Abstract Inference*, Wiley, New York.

Hannan, E.J. and M. Deistler (1988), *The Statistical Theory Of Linear Systems*, Wiley.

Hansen, L.P. and R.J. Hodrick (1980), "Forward Exchange Rates as Optimal Predictors of Future Spot Rates: An Econometric Analysis," *Journal of Political Economy*, 88, 829–853.

Kiefer, N.M. and T. Vogelsang (2002), "Heteroskedasticity-Autocorrelation Robust Standard Errors using the Bartlett Kernel without Truncation," *Econometrica*, 70, 2093–2095.

Kiefer, N.M., T. J. Vogelsang, and H. Bunzel (2000), "Simple Robust Testing of Regression Hypotheses," *Econometrica*, 68, 695–714.

Lazarus, E., D.J. Lewis, J.H. Stock, and M.W. Watson (2018), "HAR Inference: Recommendations For Practice," *Journal of Business & Economic Statistics*, 36, 541–559.

Müller, U.K. (2014), "HAC Corrections for Strongly Autocorrelated Time Series," *Journal of Business & Economic Statistics*, 32, 311–322.

Müller, U.K. and M.W. Watson (2021), "Spatial Correlation Robust Inference," ArXiv:2102.09353 (`https://arxiv.org/abs/2102.09353`).

Newey, W.K. and K.D. West (1987), "A Simple, Positive Semi-Definite, Heteroskedasticity and Autocorrelation Consistent Covariance Matrix," *Econometrica*, 55, 703–708.

Wooldridge, J.M. (2015), *Introductory Econometrics: A Modern Approach*, Cengage Learning.