



Political regimes, business cycles, seasonalities, and returns

John G. Powell^a, Jing Shi^{b,c}, Tom Smith^c, Robert E. Whaley^{d,*}

^a Department of Finance Banking and Property, Massey University, Palmerston North, New Zealand

^b School of Finance and Business, Jiangxi University of Finance and Economics, Jiangxi, China

^c School of Finance and Applied Statistics, The Australian National University, Canberra, Australia

^d Owen Graduate School of Management, Vanderbilt University, 401 21st Avenue South, Nashville, TN 37215-2422, USA

ARTICLE INFO

Article history:

Received 11 July 2008

Accepted 17 December 2008

Available online 31 December 2008

JEL classification:

G12

C15

C22

Keywords:

Persistent regimes

Regime difference tests

Spurious regression

Dummy variable

ABSTRACT

This paper provides a method for testing for regime differences when regimes are long-lasting. Standard testing procedures are generally inappropriate because regime persistence causes a spurious regression problem – a problem that has led to incorrect inference in a broad range of studies involving regimes representing political, business, and seasonal cycles. The paper outlines analytically how standard estimators can be adjusted for regime dummy variable persistence. While the adjustments are helpful asymptotically, spurious regression remains a problem in small samples and must be addressed using simulation or bootstrap procedures. We provide a simulation procedure for testing hypotheses in situations where an independent variable in a time-series regression is a persistent regime dummy variable. We also develop a procedure for testing hypotheses in situations where the dependent variable has similar properties.

© 2008 Elsevier B.V. All rights reserved.

1. Introduction

A growing body of empirical literature indicates that long-lasting regimes such as political cycles, business cycles, or seasons of the year produce significant market return and return volatility differentials. This paper revisits the question of whether political, seasonal or business regimes display significant differentials and demonstrates that once account is taken of regression variable persistence the significance of most of the differentials disappears. When long-lasting regimes are represented in a time-series regression model using indicator variables, the regression variables are extremely persistent and undermine the use of standard OLS tests of significance. After documenting the persistence properties of regimes representing presidential party affiliation, business cycles, and seasonalities, we outline how standard estimators can be adjusted for regime persistence, and demonstrate that, even after adjustment, spurious regression remains a problem in small samples. We then provide a simulation procedure that allows researchers to properly test for regime differences in small samples and show that many documented regime market return and return volatility differentials are, in fact, illusory.

The use of regime indicators (i.e., dichotomous or “dummy” variables) is commonplace in the economics literature. Usually, they are used in a regression model to test a significant difference in the intercept and/or slope where the data are observed under different environments (e.g., a time-series regression model that includes war-time and peace-time observations). What is often overlooked, however, is that this test methodology carries with it the implicit assumption that the regime dummy variables are well-behaved. Despite the general presumption that the use of dummy variables is innocuous in time-series regression, a dichotomous explanatory variable, like any other regressor, may be extremely persistent through time, thereby causing spurious regression results to arise.

This particular spurious regression problem is rooted in the work of Yule (1926) and Granger and Newbold (1974). They demonstrate that pairs of independent random walks (i.e., independent unit root processes) will be related according to standard inference in an OLS regression not because the series are, in fact, related, but because the series move together through time. The result is not specific to random walks, however. Granger et al. (2001) show that spurious regression results arise when two stationary series are positively autocorrelated (and, hence, also move together through time). Their simulation results indicate that the use of HAC standard errors helps with the problem, but spurious results still occur for sample sizes commonly used in the literature. In independent work, Ferson et al. (2003) focus on stock returns and show that a

* Corresponding author. Tel.: +1 615 343 7747; fax: +1 615 376 8879.

E-mail addresses: J.Powell@massey.ac.nz (J.G. Powell), Jing.Shi@anu.edu.au (J. Shi), Tom.Smith@anu.edu.au (T. Smith), whaley@vanderbilt.edu (R.E. Whaley).

spurious regression problem arises when stock return time-series (which have been shown to be positively autocorrelated over long periods of time) are regressed on continuous explanatory variables that are persistent (i.e., highly positively autocorrelated). They also show that the problem is exacerbated by data mining for explanatory variables since highly persistent variables are more likely to display apparent significance. Deng (2005) argues that a small net autocorrelation in returns is caused by the opposing forces of strongly autocorrelated expected returns and a negative moving average error term. He suggests that the Newey-West lag length be set as a fixed function of sample size to ameliorate the spurious regression problem (see also Richardson and Stock (1989)). Kiefer and Vogelsang (2005) provide analytical asymptotic distribution results for fixed lag length HAC adjustments when errors are serially correlated.

This paper contributes to the literature by analyzing the problem of statistical inference in time-series regressions employing persistent regime dummy variables. The problem and its solution are critically important in economics, finance, and political science, where dummy variables are often used in a time-series context. In particular, we show that hypotheses of significant return and return volatility differences related to political, seasonal and business cycle regimes can be rejected when the persistent dummy variable problem is rectified. Our solution to the persistent dummy spurious regression problem rests in simulation analysis, and we address the cases in which the independent variable and the dependent variable are dichotomous.

The paper is organized as follows. The first section demonstrates that dichotomous independent variables can be highly persistent in financial time-series and describes a simulation procedure used to assess the influence of dummy variable persistence and data mining on financial time-series regression results. The second section uses the simulation procedure to reassess statistical inference concerning political, seasonal and business cycle return and return volatility differences and shows many conclusions reached using standard statistical inference¹ are incorrect. The third section examines spurious regression in probit models. The dependent variable is assumed to be persistent and the independent variable positively autocorrelated. Again, simulation analysis is used to quantify the problem, and the simulation results are used to gauge the significance of the results in past work. The fourth section offers researchers some advice on how to avoid potentially incorrect inference when a dichotomous independent variable is used within a time-series regression framework. The final section provides a brief summary.

2. The persistent regime indicator dummy simulation procedure

Past research has acknowledged the potentially spurious role dichotomous variables can play in time-series analysis.² Enders (1995, p. 275), for example, warns that econometric specification problems can exist with dichotomous variable time-series regression models due to autocorrelation of the error term: “the residuals should approximate white noise. If the residuals are serially correlated, the estimated model does not mimic the actual data-generating process.” Roll (1984, p. 877, fn. 25) also points out that dummy variable regression models can be mis-specified due to regression error autocorrelation. In spite of the warnings, however, until recently,

no systematic analysis of the effects of persistent dummies on statistical inference has appeared.

Recognizing this problem, Powell et al. (2007) take account of the influence of a persistent dummy variable when examining whether monthly stock market returns are significantly higher under Democratic versus Republican presidencies. When testing for presidential regime return differentials, they use simulation analysis to adjust *t*-statistics and R^2 's for persistence-induced spurious regression bias as well as for spurious regression bias in combination with data mining. Once adjusted for regression variable persistence, the presidential regime dummy variable coefficient estimates obtained in an important study by Santa-Clara and Valkanov (2003) are inside the corresponding confidence bounds for spurious regression bias in three out of four stock index return series analyzed; they are inside the corresponding confidence bounds for all four series when both the effects of spurious regression bias and data mining are considered. The results of Powell et al. (2007) suggest that any difference in stock market performance under different political regimes is less than would be expected by chance. While they demonstrate the importance of dummy variable persistence, they do not provide a systematic analysis, nor do they indicate the widespread importance of the problem.

In independent work, Harding and Pagan (2006) explain why macroeconomic and financial researchers are likely to frequently encounter binary random variables that are highly persistent. Examples cited include economic cyclical states (e.g., recessions and expansions), bull and bear markets, financial market crises, hot and cold IPO markets, and booms and slumps in commodity or real estate markets. They emphasize that employing binary variables is a very popular strategy because of the roles of these variables in decision-making (e.g., avoiding downturns), explanatory relationships (e.g., estimating output gaps), and the understanding of cyclical features or inter-relationships (e.g., growth cycle features, or the relationship between bull markets and economic upturns). Nevertheless, the process of classifying these binary variables often creates, by construction, non-independent, highly persistent series because the cyclical states are defined to be sustained so that they are economically meaningful (e.g., the classification rules tend to ignore month by month noise and instead focus on lower frequency cyclical movements). They recommend using appropriate binary variable lags in non-parametric estimation when persistent dummy variables are used in time-series regression, and illustrate their estimation approach using term structure prediction of recession turning points. A similar recommendation, not specific to dummy variables, is made by Granger et al. (2001) who state that the problem of persistent OLS regression variables can be addressed by improving regression model specification using lags of the dependent and possibly the independent variable.

In related work, Pesaran and Timmerman (2006) emphasize the importance of serially correlated discrete data in economic and finance time-series analysis applied to problems such as the study of economic cyclical states as well as prediction of the direction of change of GDP growth, inflation, interest rates, or the stock market. They develop a trace canonical correlation test using dynamically augmented reduced rank regressions to account for serial correlation when one discrete random variable is used to predict another. Their approach is applicable to tests of market timing skills or business cycle analysis, and they illustrate their test with an application to forecasting GDP growth using survey data.

Our paper directly accounts for binary variable persistence in standard OLS regression using simulation analysis, as introduced and outlined below.³ Beforehand, however, it is, perhaps, useful to

¹ Standard statistical inference is defined as OLS estimation with Newey and West (1987) standard errors.

² It is well-known that the ordinary least squares error independence assumption can be violated in cross-sectional dummy variable regressions due to the self-selection problem in treatment effect regression analysis (see, e.g., Greene (2003), p. 788).

³ The approach accounts for persistence without altering regression model specification.

demonstrate how prevalent the use of dichotomous explanatory variables is in the time-series regression analysis literature. Table 1 lists a number of recent studies that use regime return and return volatility differences. Panel A of Table 1 summarizes the properties of the variables used. The regime dummy variables are diverse, ranging from business cycle or state-of-the-market indicators to monetary or political regime variables. Even a seasonal dummy variable is included. As the table shows, the dummy variables – marked by an asterisk (*) – are highly persistent. The first order autocorrelation of the daily presidential dummy used in Leblang and Mukherjee (2005), for example, is a whopping .999! While not reported or discussed in their study, the level of autocorrelation comes hardly as a matter of surprise. Since daily data are used, the runs of zeros or ones in the dummy variable time-series must be at least four years long with each successive administration. Similar style arguments can be made for persistence in the other dummies reported in Table 1. Regardless, the empirical results are clear. All of the dummy variable regressors reported in Panel A of Table 1 are highly persistent.

With persistence in the regressors, the potential for spurious regression results looms large. The remaining symptom is autocorrelation in the dependent variable. Panel A of Table 1 shows that all of the dependent variables have positive autocorrelation to varying degrees. The estimated first-order autocorrelation coefficient of the Dow Jones Industrial Average daily return volatility obtained in Leblang and Mukherjee (2005) is an extreme .99, for instance. The seasonally-adjusted money supply variable, *M1SL*, in Jensen et al. (1996) is 0.97.

The results reported in Panel A of Table 1 contain obvious symptoms of spurious regression – both the dependent and independent variables are autocorrelated. This means that standard OLS tests of significance are inappropriate. To assess the effect of the autocorrelation on statistical inference, we generalize the simulation analysis in Powell et al. (2007). We simulate critical cut-off values of the coefficient estimates, the Newey and West (1987) *t*-statistics, and adjusted *R*-squared (hereafter, \bar{R}^2) for testing whether the coefficients and significance levels estimated in time-series dummy variable regressions are less than would be expected by chance. In this section, we conduct simulations to generate a matrix of critical *t*-statistics for different pairings of autocorrelation in the dependent variable and transition probability of the dichotomous explanatory variable. Underlying each simulation run is a standard dummy variable regression model,

$$r_{t+1} = \alpha + \beta\pi_t + u_{t+1}. \quad (1)$$

The dependent variable is denoted r_{t+1} and represents, say, stock market return over the interval the end of month *t* to month *t* + 1. The dummy independent variable, π_t , is equal to 1 if a particular regime, say, a recession, is present at the beginning of the month and is 0 otherwise. The null hypothesis that there is no effect implies $\beta = 0$, whereas, if the coefficient on the dummy is significantly positive (negative), the null hypothesis is rejected in favour of the alternative, that is, monthly stock market returns are higher (lower) during recessions. Each simulation run is conducted under the assumption that the dependent and independent variables are uncorrelated, however, the autocorrelation properties of the variables are set in order to match those present in actual data. In the remainder of this section, we describe the simulation procedure and apply it using a range of levels of autocorrelation in the dependent variable in combination with a range of transition probabilities of the independent dummy variable. In the next section, we revisit the question of whether political, seasonal or business regimes display significant return or return volatility differentials and match exactly the observed autocorrelation, dummy variable transition probabilities, and sample sizes in prior tests in the literature that use dummy variables in a time-series regression context.

2.1. The simulation procedure

Table 1 demonstrates that the dependent variables used in many dummy variable regression tests are positively autocorrelated. A natural way to model such behavior is as an AR(1) process.⁴ We follow Powell et al. (2007) and generate observations for the dependent variable r_{t+1} using

$$r_{t+1} = \alpha_r + \rho_r r_t + e_{t+1} \quad \text{for } t = 1, 2, 3, \dots, n-1, \quad (2)$$

where *n* is the number of time intervals in the sample period, α_r is the intercept and ρ_r is the first-order autocorrelation coefficient. The unconditional mean and variance of the dependent variable are

$$\mu_r = \frac{\alpha_r}{1 - \rho_r}$$

and

$$\sigma_r^2 = \frac{\sigma_e^2}{1 - \rho_r^2}.$$

In the general simulations that follow, we generalize the simulation procedure by assuming that the unconditional mean and variance are set equal to zero and one, respectively. The dependent variable simulation is started at the unconditional mean μ_r and the error term for process (2) is generated from a normal distribution with a mean of zero and a variance of $1 - \rho_r^2$. The length of each series is set equal to *n*, and below we choose *n* to be 250, 500, and 1000 – typical numbers of time-series observations used in stock return studies.

The dummy variable in the regression model (1) underlying the simulation is generated using a transition matrix that represents the conditional probability of remaining in or exiting a particular dummy variable state or regime. In our general simulations, the independent dummy variable series π_t ($t = 2, 3, \dots, n-1$) is generated as a first-order Markov Chain such that the transition probabilities for potential dummy variable regime changes are

$$\Pr(\pi_t = 0 | \pi_{t-1} = 0) = q,$$

$$\Pr(\pi_t = 1 | \pi_{t-1} = 0) = 1 - q,$$

$$\Pr(\pi_t = 1 | \pi_{t-1} = 1) = q,$$

$$\Pr(\pi_t = 0 | \pi_{t-1} = 1) = 1 - q.$$

The simulated dummy variable starting value $\pi_{t=0}$ is determined by the regime dummy variable unconditional probability. The dummy variable series is generated independently of the dependent variable series, and has a first-order autocorrelation coefficient, $\rho_\pi \equiv 2q - 1$ (Hamilton, 1994, p. 684). The process is more likely than not to persist in its current state when $2q > 1$ or $q > .5$. The first-order autocorrelation is zero where $q = .5$, and is negative where $q < .5$.

In the sample-specific tests in the next section (see also Powell et al. (2007)), the conditional probabilities are calibrated from the actual time-series data used in the time-series tests listed in Panel A of Table 1. Panel B of Table 1 summarizes the information from which the transition probabilities are computed. The independent dummy variable series π_t ($t = 2, 3, \dots, n-1$) in the specific simulations is once again generated from a first-order Markov Chain, but now with transition probabilities

⁴ Naturally, different simulation results will be obtained depending on the nature of the assumed process. We choose an AR(1) since it is frequently used to model the infrequent trading effect in stock market returns. See, for example, Lo and MacKinlay (1990a) and Stoll and Whaley (1990).

Table 1

Summary statistics for variables used in various studies and transition probabilities of dummy variables.

Panel A: Summary statistics					
Study/sample period	Variable	N	Mean	Standard deviation	Auto-correlation
<i>A. Political regimes</i>					
Leblang and Mukherjee (2005)	<i>DJIA volatility</i>	28,987	0.92	0.57	0.990
June 1896–December 2001	<i>D</i>	28,987	0.46	0.50	0.999
<i>B. Seasonalities</i>					
Bouman and Jacobsen (2002)	<i>US Return</i>	344	0.96	4.42	0.012
January 1970–August 1998	<i>HLW</i>	344	0.50	0.50	0.663
<i>C. Business cycle and momentum profits</i>					
Chordia and Shivakumar (2002)	<i>UMD6M</i>	811	3.64	13.03	0.845
July 1926–December 1994	<i>BC</i>	822	0.79	0.41	0.904
Cooper et al. (2004)	<i>UMD6M</i>	804	3.55	13.04	0.844
January 1929–December 1995	<i>SOM</i>	804	0.85	0.36	0.923
Jensen et al. (1996)	<i>M1SL</i>	408	385.29	244.88	0.974
February 1954–December 1992	<i>AMBSL</i>	467	117.62	89.01	0.974
	<i>EXRESERVE</i>	467	0.50	0.30	0.938
	<i>FFPREM</i>	467	0.58	0.84	0.872
	<i>DIR</i>	467	0.54	0.50	0.900
Resnick and Shoesmith (2002)	<i>SPREAD</i>	480	1.37	1.22	0.948
January 1960–December 1999	<i>BEAR</i>	480	0.24	0.43	0.899
Panel B: Frequency and transition probability of dummy variables					
Study/subsample	Frequency	Transition probability	Frequency	Transition probability	Frequency
<i>A. Political regimes</i>					
Leblang and Mukherjee (2005)					
Republican	15,876	99.97%	5	0.03%	15,881
Democrat	5	0.04%	13,101	99.96%	13,106
Total	15,881	54.79%	13,106	45.21%	28,987
<i>B. Seasonalities</i>					
Bouman and Jacobsen (2002)	Months in November–April		Months in May–October		Total
Months in November–April	144	83.24%	29	16.76%	173
Months in May–October	28	16.37%	143	83.63%	171
Total	172	50.00%	172	50.00%	344
<i>C. Business cycle and momentum profits</i>					
Chordia and Shivakumar (2002)	Expansionary period		Contractionary period		Total
Expansionary period	636	98.00%	13	2.00%	649
Contractionary period	13	7.51%	160	92.49%	173
Total	649	78.95%	173	21.05%	822
Cooper et al. (2004)	UP state		DOWN state		Total
UP state	666	98.81%	8	1.19%	674
DOWN state	8	6.45%	116	93.55%	124
Total	674	84.46%	124	15.54%	798
Jensen et al. (1996)	Expansive period		Restrictive period		Total
Expansive period	205	94.91%	11	5.09%	216
Restrictive period	11	4.40%	239	95.60%	250
Total	216	46.35%	250	53.65%	466
Resnick and Shoesmith (2002)	Bear market		Bull market		Total
Bear market	108	92.31%	9	7.69%	117
Bull market	8	2.20%	355	97.80%	363
Total	116	24.17%	364	75.83%	480

Panel A reports summary statistics. The notation is as follows:

DJIA volatility is the 20-day moving standard deviation of the daily return on the Dow Jones Industrial Average index.*UMD6M* is cumulative six-month return to the monthly momentum strategy (estimated for this paper using the Fama–French monthly momentum factor).*M1SL* is seasonally adjusted money supply.*D_t* is 1 if a democratic president is in power and 0 otherwise.*HLW_t* takes the value 1 if month *t* falls during the period November through April and 0 otherwise.*BC_t* is 1 if month *t* falls into expansionary periods as determined by the NBER and 0 otherwise.*SOM_t* is 1 if the lagged three-year market return is non-negative and 0 otherwise.*AMBSL* is seasonally adjusted adjusted-monetary-base.*EXRESNS* is excess reserves.*FFPREM* is the federal funds premium and is calculated as the difference between the federal funds rate and the three-month T-bill rate.*DIR_t* is 1 if the previous change in the discount rate was an increase and 0 otherwise.*SPREAD* is the difference between average quarterly ten-year government bond rate and average quarterly three-month T-Bill rate.*BEAR_t* is 1 if the six-month trend in monthly returns is negative and 0 otherwise.

All variables marked by an asterisk (*) are dummy variables, and all returns are expressed in %. Data for Leblang and Mukherjee (2005) are obtained from Global Financial Data. Data for Chordia and Shivakumar (2002) and Cooper et al. (2004) are obtained from Kenneth R. French's data library. Data for Bouman and Jacobsen (2002) are obtained from MSCI, and data for Jensen et al. (1996) and Resnick and Shoesmith (2002) are obtained from the Federal Reserve Bank St. Louis. Note that, while the sample period for Chordia and Shivakumar (2002) is between July 1926 and December 1994, the Fama and French momentum factor starts in January 1927. Also note that, because the six-month cumulative log return is calculated, the number of observations for *UMD6M* is different from that of *BC*.

Panel B presents frequencies and transition probabilities for the dummy variables. Expansionary and contractionary periods are classified based on the NBER published business cycle dates (see Table 2 of Chordia and Shivakumar (2002)), UP is when the lagged three-year CRSP value-weighted index return is non-negative, DOWN is when the lagged three-year CRSP value-weighted index return is negative, expansive (restrictive) periods are those occurring during series of discount rate decreases (increases), and a bear market is when the six-month trend in monthly returns is negative.

$$\begin{aligned}\Pr(\pi_t = 0 | \pi_{t-1} = 0) &= q, \\ \Pr(\pi_t = 1 | \pi_{t-1} = 0) &= 1 - q, \\ \Pr(\pi_t = 1 | \pi_{t-1} = 1) &= p, \\ \Pr(\pi_t = 0 | \pi_{t-1} = 1) &= 1 - p.\end{aligned}$$

The dummy variable series has a first-order autocorrelation coefficient equal to $q + p - 1$ (see Hamilton, 1994, p. 684).

2.2. Critical value table for spurious regression bias

To develop 95% confidence bands for the coefficient estimates and Newey and West (1987) t -statistics as well as the \bar{R}^2 's, 10,000 simulation runs are performed. The estimates are recorded for each simulated run and the five vectors of estimates (i.e., one each for the values of α , β , t_α , t_β , and \bar{R}^2) are ranked from lowest to highest. The 95% confidence interval (i.e., the lower and upper critical levels) of the coefficient estimates and the t -statistics are defined as the observations corresponding to 2.5th and 97.5th percentiles. The 95% confidence interval of the \bar{R}^2 falls at the 95th percentile. The \bar{R}^2 and other regression estimates obtained using the actual data are compared to the critical \bar{R}^2 level, critical coefficient estimates, and critical t -statistics to assess statistical significance.

Panel A of Table 2 contains the critical Newey-West t -statistics⁵ for different combinations of autocorrelation in the dependent variable (i.e., the rows in the table) and transition probability of the dummy variable explanatory variable (i.e., the columns). The first-order autocorrelation of the explanatory variable is $2q - 1$. Lower and upper values of the critical t -statistics for a given sample size are reported. As noted earlier, the sample sizes of 250, 500, and 1000 observations in Table 2 are chosen to correspond to the various sample sizes in the studies reported in Table 1. The mean and standard deviation of the dependent variable series are set equal to 0 and 1, respectively.

Table 2 reveals that even with adjustments of the Newey-West estimator to minimize the potential for spurious regression, spurious regression is a significant problem in sample sizes that are common in the literature. The first set of results in Panel A of Table 2 are for a sample size of 250 observations. Where the dependent variable autocorrelation is modest, say, 0.1, and the dummy variable has a transition probability of $q = .99$ (i.e., autocorrelation of 0.98), the critical 95% confidence interval t -statistic values are $(-2.32, 2.39)$ compared with the standard large sample 95% confidence interval t -statistic values of $(-1.96, 1.96)$. As the sample size increases, the critical 95% confidence interval t -statistic values narrow to $(-2.09, 2.11)$ for sample size 500 and $(-2.04, 2.06)$ for sample size 1000, still above conventional levels, showing that the potential for spurious regression exists for the typical sample sizes used in monthly return studies. The critical 95% confidence interval t -statistic values obtained are quite general across a number of different choices of Newey West estimator lag length (including a plug in AR lag suggested by Andrews (1991) and a lag equal to a fixed function of the sample size, bT , where $b = 0.05$) and kernels (including the Bartlett and quadratic spectral kernels).⁶

Interestingly, autocorrelation of the dependent variable has a "one-off" effect on the critical t -statistics even when there is no autocorrelation of the dichotomous explanatory variable. Consider the case where dependent variable autocorrelation is 0.1, q is 0.5 (i.e., zero autocorrelation in the dichotomous explanatory variable), and the sample size is 250. The 95% confidence interval is $(-2.03,$

2.04), still wider than the conventional values $(-1.96, 1.96)$. In addition, the size of the interval remains approximately the same as the dependent variable autocorrelation increases. This finding can be explained by noting that the 95% confidence t -statistic interval does not converge asymptotically to $(-1.96, 1.96)$ when autocorrelation is present, as shown in the analytical HAC asymptotic sub-section 2.4 below (see also Kiefer and Vogelsang (2005)). Further widening of the critical t -statistic intervals occurs only when dependent and independent variable autocorrelation increase together, and is especially apparent when autocorrelation of both variables is high. As Panel A of Table 2 shows, the autocorrelation adjustment more than doubles the critical t -statistics relative to conventional levels when autocorrelation of the dependent variable exceeds 0.9 and the transition probability exceeds 0.99, that is $(-4.49, 4.61)$.

Also of interest is the Table 2 finding that simulated confidence intervals actually tighten (all else being equal) relative to conventional confidence intervals when the explanatory variable autocorrelation is negative, thus further enriching the simulation analysis results and their implications. With positive dependent variable autocorrelation but negative independent variable autocorrelation, conventional standard error estimates are upward (not downward) biased, thus causing standard confidence intervals to over-reject rather than under-reject as happens when autocorrelation is exclusively positive.

An important point to note is that the critical t -statistics reported in Panel A of Table 2 can be directly used in any time-series regression analysis study whenever a dichotomous explanatory variable is likely to be persistent.⁷ This point is illustrated in the following section when the parameters of the simulation procedure are adapted to match the parameters used in six studies referenced in Table 1, since the Table 2 critical t -statistic values are reasonably close to the values obtained in the customized simulations. For now, it is sufficient to recognize that the simulation procedure t -statistics are easy to implement. They simply replace conventional t -statistics when it is important to take account of the influence of autocorrelation in testing whether regression coefficient estimates are less than would be expected by chance.⁸

2.3. Critical values for spurious regression bias and more than one variable being examined

Panel B of Table 2 reports a second set of critical t -statistics that use Bonferonni correction intervals to take account of the number of series potentially examined by researchers in the search for statistical significance. While this issue has been addressed in asset pricing theory tests by Foster et al. (1997),⁹ Sullivan et al. (2001) discuss why data mining considerations are especially relevant in seasonal dummy variable studies. Their point can be illustrated in relation to Bouman and Jacobsen's (2002) "Halloween Indicator" study. The starting point for the Halloween Indicator is based on a popular press expression "Sell in May and go away", but there are many potential ending points for the indicator, with there being no strong reason to support a Halloween ending date. There are also

⁵ The lag length choice follows Granger et al. (2001) who use a Newey-West lag length equal to integer $(4(T \cdot 100)^{1/4})$ in their simulations to deal with the effects of serial correlation.

⁶ In unreported results, Table 2 is re-estimated using the Bartlett method with lag length set by AR fit as well as the quadratic spectral method with lag length set by AR fit, and the results are substantially the same as those reported in Table 2. See also the discussion of Table 4 in subsection 2.4 below.

⁷ The state transition probabilities in many applications are asymmetric, as illustrated in Panel B of Table 1, and this asymmetry can be important, especially when states are highly persistent. The simulation procedure dichotomous variable transition probability process is therefore customized for each study when assessing statistical inference, but the simulation procedure t -statistics provided by Table 2 are usually quite close to those provided by the customized simulations, so Table 2 can be used as a first step in assessing statistical inference.

⁸ The values of the simulated confidence intervals and critical t -statistics do not depend upon unobserved latent variables, so our procedure is simple to implement and provides a conservative determination as to whether a coefficient estimate is less than would be expected by chance. See Ferson et al. (2003).

⁹ See, also, Lo and MacKinlay (1990b), Sullivan et al. (1999, 2001), Ferson et al. (2003), and Santa-Clara and Valkanov (2003).

Table 2
Cut-off dummy variable t-statistics for dependent and dummy variable autocorrelation combinations.

Sample size (n)	Auto-correlation of dependent variable	Transition probability of dummy variable (q)								
		0.01	0.05	0.10	0.25	0.50	0.75	0.90	0.95	0.99
<i>Panel A: Cut-offs for spurious regression bias</i>										
250	0.1	-1.95/1.97	-1.99/1.99	-2.00/1.96	-2.01/1.99	-2.03/2.04	-2.04/2.02	-2.07/2.05	-2.11/2.09	-2.32/2.39
250	0.2	-1.94/2.01	-1.97/1.92	-1.97/1.96	-2.00/2.01	-2.02/2.03	-2.07/2.08	-2.10/2.14	-2.14/2.15	-2.36/2.41
250	0.3	-1.89/1.90	-1.93/1.90	-1.94/1.91	-1.95/1.94	-2.00/2.05	-2.09/2.10	-2.13/2.15	-2.23/2.18	-2.52/2.44
250	0.4	-1.85/1.86	-1.84/1.88	-1.93/1.88	-1.96/1.91	-2.03/1.99	-2.10/2.11	-2.18/2.16	-2.25/2.23	-2.60/2.55
250	0.5	-1.82/1.84	-1.84/1.86	-1.84/1.90	-1.93/1.92	-2.02/2.02	-2.10/2.17	-2.22/2.16	-2.33/2.32	-2.74/2.69
250	0.6	-1.77/1.78	-1.81/1.80	-1.80/1.80	-1.95/1.91	-1.98/2.06	-2.15/2.13	-2.29/2.39	-2.46/2.37	-2.87/2.87
250	0.7	-1.72/1.68	-1.76/1.74	-1.78/1.78	-1.86/1.93	-2.04/1.99	-2.22/2.24	-2.27/2.42	-2.57/2.57	-3.22/3.12
250	0.8	-1.61/1.64	-1.72/1.69	-1.75/1.71	-1.86/1.88	-2.02/2.00	-2.22/2.27	-2.58/2.58	-2.84/2.84	-3.48/3.62
250	0.9	-1.56/1.55	-1.58/1.56	-1.62/1.63	-1.87/1.81	-2.07/2.00	-2.26/2.22	-2.75/2.76	-3.20/3.23	-4.49/4.61
500	0.1	-1.94/1.96	-1.96/1.98	-1.96/1.92	-1.96/2.00	-1.96/1.99	-2.02/2.00	-2.02/1.98	-2.03/2.08	-2.09/2.11
500	0.2	-1.93/1.89	-1.93/1.95	-1.93/1.90	-1.95/1.96	-2.01/1.97	-2.02/2.01	-2.09/2.02	-2.06/2.11	-2.16/2.17
500	0.3	-1.90/1.89	-1.93/1.91	-1.91/1.93	-1.96/2.00	-2.00/1.96	-2.02/2.08	-2.11/2.10	-2.12/2.18	-2.22/2.18
500	0.4	-1.82/1.85	-1.89/1.92	-1.92/1.87	-1.97/1.93	-1.96/2.00	-2.05/2.08	-2.15/2.13	-2.17/2.16	-2.25/2.22
500	0.5	-1.82/1.83	-1.83/1.82	-1.88/1.82	-1.93/1.88	-1.98/2.06	-2.06/2.08	-2.22/2.19	-2.23/2.24	-2.34/2.37
500	0.6	-1.76/1.73	-1.80/1.76	-1.80/1.80	-1.86/1.90	-2.01/2.00	-2.15/2.08	-2.27/2.26	-2.37/2.37	-2.40/2.49
500	0.7	-1.67/1.64	-1.75/1.72	-1.76/1.75	-1.85/1.90	-2.00/2.01	-2.18/2.15	-2.28/2.29	-2.48/2.47	-2.74/2.71
500	0.8	-1.60/1.62	-1.65/1.69	-1.75/1.74	-1.86/1.86	-1.97/1.99	-2.18/2.22	-2.46/2.51	-2.63/2.64	-3.09/3.03
500	0.9	-1.53/1.51	-1.59/1.55	-1.66/1.67	-1.79/1.79	-1.99/1.98	-2.20/2.25	-2.70/2.65	-3.05/3.10	-3.89/3.99
1000	0.1	-1.92/1.98	-1.93/1.93	-1.96/1.97	-1.99/2.02	-2.00/1.96	-1.96/1.98	-2.03/2.04	-2.01/2.02	-2.04/2.06
1000	0.2	-1.88/1.96	-1.92/1.96	-1.95/1.93	-1.94/2.01	-2.00/2.01	-2.00/2.04	-2.04/1.99	-2.03/2.04	-2.05/2.05
1000	0.3	-1.88/1.90	-1.95/1.90	-1.93/1.93	-1.96/1.92	-1.98/2.01	-1.98/2.05	-2.06/2.06	-2.07/2.08	-2.09/2.09
1000	0.4	-1.87/1.91	-1.88/1.86	-1.87/1.89	-1.91/1.92	-2.02/1.91	-2.03/2.01	-2.11/2.09	-2.12/2.11	-2.17/2.12
1000	0.5	-1.82/1.83	-1.85/1.83	-1.83/1.91	-1.90/1.93	-1.98/1.98	-2.05/2.08	-2.13/2.20	-2.16/2.21	-2.25/2.23
1000	0.6	-1.81/1.78	-1.84/1.79	-1.88/1.79	-1.96/1.90	-1.99/1.94	-2.08/2.09	-2.14/2.14	-2.19/2.17	-2.28/2.29
1000	0.7	-1.69/1.72	-1.77/1.76	-1.81/1.83	-1.95/1.90	-2.00/1.98	-2.09/2.12	-2.19/2.22	-2.35/2.31	-2.37/2.41
1000	0.8	-1.65/1.65	-1.74/1.71	-1.80/1.75	-1.91/1.86	-2.00/1.93	-2.10/2.08	-2.31/2.31	-2.50/2.48	-2.74/2.69
1000	0.9	-1.56/1.55	-1.64/1.64	-1.71/1.67	-1.84/1.88	-1.99/1.96	-2.18/2.13	-2.52/2.49	-2.75/2.74	-3.31/3.37
<i>Panel B: Cut-offs for spurious regression bias and data mining</i>										
250	0.1	-2.61/2.54	-2.61/2.62	-2.60/2.58	-2.67/2.68	-2.73/2.71	-2.78/2.71	-2.72/2.82	-2.82/2.72	-3.63/3.59
250	0.2	-2.60/2.61	-2.61/2.55	-2.68/2.57	-2.61/2.62	-2.76/2.73	-2.79/2.72	-2.76/2.80	-2.87/2.85	-3.65/3.69
250	0.3	-2.53/2.61	-2.58/2.60	-2.55/2.59	-2.57/2.63	-2.70/2.72	-2.81/2.83	-2.83/2.85	-2.87/2.86	-3.69/3.99
250	0.4	-2.43/2.49	-2.47/2.46	-2.58/2.50	-2.61/2.52	-2.74/2.68	-2.79/2.88	-2.85/2.82	-3.00/2.93	-4.20/4.16
250	0.5	-2.41/2.49	-2.47/2.43	-2.47/2.54	-2.60/2.52	-2.80/2.67	-2.79/2.91	-2.89/2.88	-3.23/3.06	-4.38/4.13
250	0.6	-2.30/2.28	-2.40/2.40	-2.48/2.33	-2.58/2.47	-2.63/2.73	-2.89/2.79	-2.94/3.22	-3.30/3.25	-4.59/4.37
250	0.7	-2.26/2.22	-2.35/2.36	-2.37/2.34	-2.48/2.53	-2.67/2.58	-2.93/2.88	-3.31/3.15	-3.53/3.41	-5.06/4.65
250	0.8	-2.10/2.10	-2.21/2.17	-2.33/2.27	-2.45/2.45	-2.64/2.57	-2.96/3.01	-3.48/3.41	-3.70/3.76	-5.78/6.26
250	0.9	-1.99/2.06	-2.06/2.04	-2.14/2.15	-2.37/2.36	-2.70/2.64	-2.91/2.87	-3.62/3.65	-4.24/4.33	-7.05/6.90
500	0.1	-2.59/2.51	-2.64/2.64	-2.59/1.51	-2.62/2.60	-2.61/2.58	-2.62/2.72	-2.70/2.73	-2.74/2.79	-2.76/2.91
500	0.2	-2.52/2.52	-2.51/2.54	-2.53/2.53	-2.54/2.56	-2.66/2.71	-2.67/2.72	-2.78/2.72	-2.72/2.70	-2.88/2.84
500	0.3	-2.49/2.45	-2.51/2.53	-2.51/2.50	-2.65/2.65	-2.64/2.67	-2.73/2.69	-2.81/2.77	-2.78/2.89	-2.99/2.89
500	0.4	-2.38/2.50	-2.41/2.50	-2.47/2.41	-2.61/2.51	-2.64/2.56	-2.75/2.76	-2.87/2.79	-2.92/2.83	-2.97/3.06
500	0.5	-2.35/2.36	-2.37/2.45	-2.49/2.40	-2.56/2.55	-2.68/2.62	-2.79/2.82	-2.90/2.87	-3.03/3.08	-3.23/3.20
500	0.6	-2.25/2.30	-2.30/2.37	-2.42/2.41	-2.52/2.49	-2.62/2.53	-2.79/2.71	-2.95/2.96	-3.03/3.13	-3.47/3.44
500	0.7	-2.19/2.15	-2.27/2.19	-2.30/2.27	-2.39/2.50	-2.58/2.56	-2.84/2.82	-3.03/3.08	-3.29/3.19	-3.85/3.74
500	0.8	-2.07/2.15	-2.21/2.21	-2.27/2.23	-2.40/2.44	-2.57/2.52	-2.84/2.92	-3.26/3.29	-3.47/3.50	-4.37/4.21
500	0.9	-2.04/2.02	-2.07/2.09	-2.15/2.18	-2.39/2.40	-2.67/2.57	-2.88/3.03	-3.47/3.45	-4.20/4.15	-5.35/5.56
1000	0.1	-2.59/2.60	-2.56/2.59	-2.53/2.66	-2.59/2.59	-2.65/2.62	-2.58/2.71	-2.66/2.56	-2.61/2.63	-2.70/2.63
1000	0.2	-2.49/2.55	-2.54/2.55	-2.55/2.59	-2.62/2.61	-2.55/2.56	-2.60/2.66	-2.69/2.72	-2.68/2.67	-2.71/2.75
1000	0.3	-2.48/2.45	-2.53/2.51	-2.59/2.47	-2.50/2.58	-2.62/2.68	-2.70/2.63	-2.75/2.63	-2.71/2.70	-2.72/2.87
1000	0.4	-2.46/2.49	-2.45/2.46	-2.58/2.52	-2.59/2.57	-2.57/2.57	-2.71/2.73	-2.78/2.67	-2.78/2.79	-2.82/2.88
1000	0.5	-2.37/2.43	-2.46/2.40	-2.50/2.50	-2.45/2.45	-2.53/2.68	-2.74/2.79	-2.79/2.90	-2.85/2.77	-2.87/2.98
1000	0.6	-2.40/2.28	-2.36/2.36	-2.48/2.46	-2.59/2.49	-2.61/2.65	-2.73/2.71	-2.95/2.88	-2.91/2.85	-3.07/3.05
1000	0.7	-2.27/2.31	-2.41/2.30	-2.44/2.39	-2.55/2.44	-2.65/2.62	-2.72/2.79	-2.97/2.89	-3.12/3.08	-3.16/3.29
1000	0.8	-2.17/2.17	-2.24/2.25	-2.36/2.30	-2.56/2.52	-2.61/2.53	-2.81/2.83	-3.00/3.03	-3.31/3.27	-3.60/3.48
1000	0.9	-2.08/2.02	-2.17/2.08	-2.28/2.20	-2.47/2.43	-2.61/2.61	-2.83/2.75	-3.15/3.31	-3.57/3.65	-4.33/4.48

This table reports dummy variable coefficient critical cut-off t-statistic values for dummy variable regression model, $r_t = \alpha + \beta\pi_t + u_t$, under different combinations of autocorrelation of the dependent variable and transition probability of the dummy variable. Autocorrelation of the dummy variable is equal to two times the dummy variable transition probability minus one ($2q - 1$). Panel A reports lower and upper bounds (i.e., critical cut-off values) of the 95% confidence interval of t-statistics adjusted for spurious regression bias. Panel B reports lower and upper bounds (i.e., critical cut-off values) of the 95% confidence interval of t-statistics adjusted for spurious regression bias and data mining with five variables examined. The t-statistics are adjusted for autocorrelation and heteroskedasticity using Newey and West (1987) with the lag length choice following Granger et al. (2001) who use a Newey West lag length equal to integer $(4T/100)^{1/4}$ in their simulations to deal with the effects of serial correlation. The monthly mean return of the simulated dependent variable series is zero and the monthly standard deviation is one.

popular press expressions that support alternative seasonal investment strategies. A recent *Wall Street Journal* article by Browning (2005), for example, instead argues that September (not May) is the month to sell since more bull markets “die” in September than any other month. On the other hand, investors are also told in the

press to anticipate a “summer rally” each year, which implies summer buying (not selling) opportunities. In reality, there are a potentially infinite number of seasonal dummies that could be tested for superior investment performance, so data mining should be an important consideration when evaluating the statistical significance

of the investment performance of the Halloween indicator or any other seasonal strategy.

An adjustment factor of five is used to estimate a set of “modified” critical t -statistics. This adjustment implicitly assumes that in situations where data mining is a potential concern, at least five combinations of dependent and independent variable series are examined in the search for statistically significant relationships.¹⁰ This modification is equivalent to requiring a 1% rather than a 5% level of significance due to the number of series being searched. Panel B of Table 2 contains the results. At 250 observations, for example, the modified critical t -statistics are (−3.63, 3.59) at autocorrelation of 0.1 in the dependent variable and transition probability of 0.99 in the independent variable, compared with (−2.32, 2.39) using no data mining consideration and (−1.96, 1.96) using conventional methods. In general, the differences introduced by data mining considerations are profound.

2.4. Simulation findings vis-à-vis analytical HAC asymptotic results

Kiefer and Vogelsang (2005) (hereafter, KV) provide analytical asymptotic distribution results that help explain this paper’s simulation findings. More specifically, they develop asymptotic theory for the situation where the regression error structure follows an ARMA process due to heteroskedastic–autocorrelation, and provide analytical asymptotic results for fixed lag length HAC adjustments. Their analysis applies to this paper’s simulation framework because of the presence of autocorrelated errors in both the dependent and independent regression variables. (See Eqs. (1) and (2) and section 2.1. See also Deng (2005), Ferson et al. (2003); and Nabeya and Perron (1994)).

KV derive the entire asymptotic distribution of the HAC estimator under the assumption that $M = bT$, where M is the lag length, T is the sample size, and b is a fixed constant in the interval (0, 1).¹¹ The variance matrix is a random variable rather than a constant under this assumption, so asymptotic variance estimators converge to limiting random matrices that depend upon the kernel (through its second derivative) and the fixed constant b . In contrast, under the standard approach, asymptotic distributions are derived under the assumption that M/T goes to zero as the sample size tends to infinity, in which case the asymptotic distribution does not depend upon b and the kernel. But, this assumption is unrealistic in finite samples where M/T cannot be zero. KV are therefore able to analyze the trade-off in band-width and kernel choice in size and power for the sampling variability of the HAC estimate of the asymptotic variance. They find that wider bandwidth provides fewer size distortions, especially when there is strong positive serial correlation, but smaller bandwidth provides greater power. At the same time, the Bartlett kernel is the most powerful, however, the quadratic spectral (QS) kernel leads to tests with the least size distortions. Thus, there is an important trade-off in band-width and kernel choice in size and power.

Asymptotic 95% confidence t -statistic interval critical values (cv) for fixed lag length HAC estimators are estimated using the cubic equation provided in KV,

$$cv(b) = \alpha_0 + \alpha_1 b + \alpha_2 b^2 + \alpha_3 b^3,$$

where b is a fixed constant and $\alpha_0 = 1.96$. Table 3 reports the asymptotic critical values (cv) for various values of the fixed constant b using Bartlett method estimation (with $\alpha_1 = 2.8986$, $\alpha_2 = 0.6970$, and $\alpha_3 = -0.7771$) and QS method estimation (with $\alpha_1 = 2.9880$,

$\alpha_2 = 10.6359$, and $\alpha_3 = -3.3974$). These coefficient values are obtained using the analytical cut-offs derived by KV.¹²

The Table 3 confidence intervals are smaller for all values of the fixed lag length coefficient b when estimated using the Bartlett procedure. The Table 3 confidence intervals can be compared to the Table 2 simulation results by relating the analytical versus simulated t -statistic cut-offs to the value of b used in each simulation (recall $M = bT$). With 250 observations, $b = .04$ and the Bartlett/QS cut-offs are 2.08/ 2.10. With 500 observations, $b = .02$ and the Bartlett/QS cut-offs are both 2.02, and, with 1000 observations, $b = .01$ and the Bartlett/QS cut-offs are both 1.99.¹³ It is important to note that the simulation cut-offs in Table 2 are much closer to the analytical cut-offs than they are to 1.96.

Deng (2005) recommends lag length fixed constants b in the interval [.05, .10] to minimize the potential for spurious regression in financial time-series regressions involving persistent explanatory variables. Consequently, we repeat the dichotomous dependent variable simulation procedure regression results table (Table 2) using the fixed lag length constant $b = .05$. The results are reported in Table 4 and indicate that this provides an improvement when autocorrelation of the dependent variable is high (e.g., .8 or .9) but widens confidence intervals for lower autocorrelation levels when compared to Table 2. This finding can be explained using the Table 3 results by noting that the Bartlett/QS 95% confidence t -statistic interval converge asymptotically to (−2.11, 2.11)/(−2.14, 2.14) when the fixed lag length coefficient b is set equal to .05. The simulation cut-offs in Table 2 are, therefore, a lot closer to the analytical cut-offs than they are to (−1.96, 1.96). Table 4 indicates that there is still considerable potential for spurious regression in small samples that must be analyzed using the simulation procedure, but it also reveals a clear improvement relative to Table 2, so we recommend fixed lag length HAC estimation with a fixed lag length constant $b = .05$ and consequently a cutoff of 2.11 for Bartlett weights and 2.14 for QS weights rather than the 1.96 table value.

3. Tests for regime return and return volatility differences using the simulation procedure

In the previous section, we illustrated how to perform statistical inference when the dependent variable has autocorrelation and the dummy variable regressor is persistent using arbitrary values for the level of autocorrelation in the dependent variable and the level of persistence in the regressor. We now turn to specific applications.

3.1. Presidential regimes and the stock market

Leblang and Mukherjee (2005) (hereafter LM) examine the question of whether political regimes affect the volatility of daily stock returns using a rational expectation model of inflation under left-wing governments.¹⁴ They hypothesize that traders anticipate the results of elections, thus leading to reduced stock market trading when Democrats are expected to win presidential elections as investors anticipate lower returns due to higher inflation. The LM volatility hypothesis is tested using daily capital returns for the Dow Jones Industrial Average for the time period May 1896 through December 2001. Various variables that control for interest rates (the 10-year government bond yield), divided governments between the presi-

¹² Note that there appears to be a coefficient transcription error in Kiefer and Vogelsang (2005) that has been corrected to obtain these values.

¹³ Note that the lag length used in the simulations is integer $4(T/100)^{1/4}$, so, with $T = 250$, $M = 9$ and, therefore, $b = 0.04$.

¹⁴ More recently, Kanas (2008) examines whether real interest rate regime changes are related to presidential regime changes.

¹⁰ Since this adjustment factor is arbitrary, we choose conservatively.

¹¹ Earlier work by Richardson and Stock (1989) derives fixed lag length asymptotic distribution theory for tests of mean reversion with multi-year returns that provides superior approximations to the relevant finite sample distributions.

Table 3

Two-sided asymptotic critical values at 95% significance level using the Bartlett and Quadratic spectral methods.

b	Bartlett	Quadratic spectral
0.01	(-1.99, 1.99)	(-1.99, 1.99)
0.02	(-2.02, 2.02)	(-2.02, 2.02)
0.03	(-2.05, 2.05)	(-2.06, 2.06)
0.04	(-2.08, 2.08)	(-2.10, 2.10)
0.05	(-2.11, 2.11)	(-2.14, 2.14)
0.06	(-2.14, 2.14)	(-2.18, 2.18)
0.07	(-2.17, 2.17)	(-2.22, 2.22)
0.08	(-2.20, 2.20)	(-2.27, 2.27)
0.09	(-2.23, 2.23)	(-2.31, 2.31)
0.1	(-2.26, 2.26)	(-2.36, 2.36)
0.2	(-2.56, 2.56)	(-2.96, 2.96)
0.3	(-2.87, 2.87)	(-3.72, 3.72)
0.4	(-3.18, 3.18)	(-4.64, 4.64)
0.5	(-3.49, 3.49)	(-5.69, 5.69)
0.6	(-3.78, 3.78)	(-6.85, 6.85)
0.7	(-4.06, 4.06)	(-8.10, 8.10)
0.8	(-4.33, 4.33)	(-9.42, 9.42)
0.9	(-4.57, 4.57)	(-10.79, 10.79)
1.0	(-4.78, 4.78)	(-12.19, 12.19)

Asymptotic 95% confidence t -statistic interval critical values (cv) for fixed lag length HAC estimators are estimated using the cubic equation provided in Kiefer and Vogelsang (2005) $cv(b) = \alpha_0 + \alpha_1 b + \alpha_2 b^2 + \alpha_3 b^3$, where b is a fixed constant and $\alpha_0 = 1.96$. Asymptotic critical values (cv) are reported for various values of the fixed constant b using Bartlett method estimation (with $\alpha_1 = 2.8986$, $\alpha_2 = 0.6970$, and $\alpha_3 = -0.7771$) and QS method estimation (with $\alpha_1 = 2.9880$, $\alpha_2 = 10.6359$, and $\alpha_3 = -3.3974$). These coefficient values are obtained using the analytical cut-offs derived by Kiefer and Vogelsang (2005).

dent and Congress, and major events such as war are included as explanatory variables, together with presidential party affiliation.

Table 5 provides test results for presidential regime volatility differences using the LM Dow Jones Industrial Average sample. In the row below the coefficient estimates in Table 5 are the lower and upper confidence bounds for the coefficient estimates, t -statistics, and \bar{R}^2 's adjusting for spurious regression bias. The row below that contains the coefficient estimates, t -statistics, and \bar{R}^2 's adjusting for spurious regression bias as well as the number of regression variable series potentially examined. The 95% confidence bounds for the t -statistic reported in Table 5 are wide ((-6.27, 6.21), and (-8.21, 8.33) when adjusted for data mining) due to the extreme persistence properties of the daily presidential regime dummy variable (see Table 1). The presidential regime coefficient t -statistic, -4.59, falls within the 95% t -statistic confidence bounds, demonstrating that correctly accounting for regime dummy persistence fails to reject the hypothesis that volatility is no different across presidential regimes.¹⁵ Interestingly, the same conclusion that apparent presidential regime return volatility differences are due to serial correlation can be obtained using the analytical t -statistics provided by a Bartlett kernel fixed lag length HAC estimator with the fixed lag length constant b set equal to .05. The volatility difference t -statistic falls in absolute value from -4.19 (non-parenthesized) to -1.19 (parenthesized) in Table 5 when the HAC lag length constant b is set equal to .05!

3.2. Seasonalities and return differences

The use of seasonal dummies in market anomaly time-series studies has a history that is virtually as long as the study of market efficiency itself. Whenever seasonal dummies are employed in situations where the season lasts longer than the observation frequency (e.g., daily observations are used in a study of monthly

seasonalities), the dichotomous variable is, by construction, persistent. An interesting example of this phenomenon and its consequences is the "Halloween Indicator" study of Bouman and Jacobsen (2002) (hereafter, BJ).¹⁶ The Halloween Indicator investment strategy is based upon the Wall Street adage "Sell in May and go away". A selling strategy needs a point in time where the selling is reversed, with this being deemed to happen six months later (hence the "Halloween Indicator"). As noted earlier, there are a potentially infinite number of beginning and ending points of seasonal dummies that could be tested for superior investment performance.¹⁷ The issue of data mining raised in BJ is, therefore, an important consideration when evaluating the statistical significance of the Halloween indicator investment strategy. Consequently, this study is of particular importance in showing how data mining can combine with dichotomous variable autocorrelation to affect statistical inference.

BJ test their investment strategy of selling stocks in summer and buying them back six months later using dummy variable regression model (1), with the minor modification that monthly returns are regressed against the concurrent value of the seasonal dummy variable. The dependent variable in the study is the monthly value-weighted local currency stock market rates of return for 19 developed countries over the time period January 1970 to August 1998 and 18 developing countries over shorter time intervals. The independent variable is a dummy variable that is 1 if month t is November through April and is 0 otherwise. The data set is obtained from Morgan Stanley Capital International (MSCI). BJ report that the investment strategy of selling (buying) stocks in May (Halloween) leads to statistically significant superior performance in 20 of the 37 countries examined.

The BJ study test for seasonal return differences is reproduced in Table 6. The regression coefficient estimates obtained are very close to those in BJ, with the only noticeable difference being for the Philippines and Malaysia (presumably due to changes in construction of the MSCI indexes). Table 6 also reports simulated confidence intervals for the BJ study results in the same format as used in Table 5. In this case, the simulation procedure must be modified to account for the fact that the dummy variable is deterministic – a seasonal determined by the calendar. To do so, we use the same π_t in each simulated sample. The inferences using the different stock market indexes change from those of the BJ study. For the United States, for example, the 95% confidence bounds are (-1.97, 2.03) when spurious regress bias is considered alone and (-2.64, 2.77) when both spurious regression bias and data mining are considered. The reported t -statistic is 1.95 and is contained within both sets of confidence bounds. Again, similar conclusions can be drawn from the general results reported in Table 2. The BJ sample size, 344, is closest to the 250 observation panels of Table 2. If the dependent variable has autocorrelation of 0.1 and the dummy variable transition probability is 0.9, the cut-off t -statistic values are (-2.07, 2.05) when spurious regress bias is considered alone and (-2.72, 2.82) when both spurious regression bias and data mining are considered. These are essentially the same as those obtained for the United States results with the customized simulation procedure. A bootstrap procedure provides (unreported) results that are essentially the same as those obtained using the simulation procedure.

These results are another good example of how taking account of autocorrelation can affect statistical inference since they go from being reported as significant in BJ to being insignificant when the cut-off t -statistic is adjusted for autocorrelation. Indeed, the differ-

¹⁵ As a robustness check, we repeated the test using a bootstrap procedure and reached the same conclusion. The bootstrap procedure uses a block sampler to retain the data's autocorrelation properties.

¹⁶ More recently, Jacobsen and Marquering (2008) re-examine competing explanations for temporal variation in stock returns. Among other things, they confirm the Bouman and Jacobsen (2002) findings.

¹⁷ See, in particular, Sullivan et al. (2001).

Table 4
Cut-off dummy variable t -statistics for dependent and dummy variable autocorrelation combinations using Bartlett and Quadratic spectral methods.

Method/sample size (n)	Auto-correlation of dependent variable	Transition probability of dummy variable (q)								
		0.01	0.05	0.10	0.25	0.50	0.75	0.90	0.95	0.99
<i>Bartlett</i>										
250	0.1	-2.04/2.05	-2.09/2.10	-2.12/2.13	-2.12/2.11	-2.12/2.13	-2.12/2.12	-2.14/2.18	-2.22/2.25	-2.61/2.74
250	0.2	-2.04/2.10	-2.07/2.10	-2.09/2.10	-2.11/2.06	-2.13/2.13	-2.11/2.13	-2.16/2.18	-2.26/2.17	-2.62/2.67
250	0.8	-1.82/1.82	-1.86/1.84	-1.89/1.92	-2.01/2.05	-2.15/2.09	-2.20/2.18	-2.40/2.41	-2.57/2.58	-3.59/3.49
250	0.9	-1.57/1.59	-1.75/1.78	-1.84/1.86	-2.00/2.02	-2.08/2.11	-2.23/2.17	-2.46/2.44	-2.74/2.78	-4.05/3.93
<i>Quadratic spectral</i>										
250	0.1	-2.20/2.18	-2.19/2.18	-2.14/2.16	-2.17/2.18	-2.16/2.17	-2.21/2.17	-2.17/2.22	-2.18/2.30	-3.08/3.18
250	0.2	-2.14/2.20	-2.15/2.13	-2.21/2.15	-2.15/2.15	-2.13/2.20	-2.25/2.19	-2.19/2.22	-2.28/2.23	-2.97/3.05
250	0.8	-2.16/2.19	-2.13/2.14	-2.14/2.18	-2.20/2.14	-2.18/2.15	-2.18/2.21	-2.28/2.31	-2.51/2.45	-3.41/3.44
250	0.9	-2.16/2.15	-2.18/2.15	-2.17/2.14	-2.12/2.16	-2.14/2.17	-2.23/2.18	-2.38/2.34	-2.57/2.60	-3.91/3.73
<i>Bartlett</i>										
500	0.1	-2.12/2.10	-2.10/2.13	-2.06/2.10	-2.08/2.08	-2.10/2.09	-2.09/2.11	-2.08/2.10	-2.14/2.20	-2.34/2.39
500	0.2	-2.05/2.07	-2.09/2.15	-2.16/2.08	-2.14/2.13	-2.13/2.05	-2.07/2.10	-2.15/2.08	-2.13/2.15	-2.28/2.48
500	0.8	-1.93/1.91	-1.96/1.97	-2.02/1.99	-2.08/2.14	-2.12/2.10	-2.17/2.10	-2.24/2.25	-2.34/2.36	-2.68/2.69
500	0.9	-1.78/1.81	-1.86/1.89	-2.01/1.96	-2.00/2.06	-2.11/2.04	-2.22/2.17	-2.31/2.26	-2.38/2.48	-2.95/2.97
<i>Quadratic spectral</i>										
500	0.1	-2.13/2.13	-2.11/2.16	-2.16/2.17	-2.24/2.14	-2.18/2.14	-2.20/2.13	-2.21/2.18	-2.17/2.17	-2.56/2.52
500	0.2	-2.15/2.13	-2.19/2.19	-2.18/2.14	-2.17/2.14	-2.16/2.22	-2.19/2.19	-2.16/2.19	-2.18/2.22	-2.62/2.56
500	0.8	-2.14/2.21	-2.15/2.20	-2.15/2.16	-2.23/2.14	-2.21/2.15	-2.22/2.21	-2.19/2.25	-2.27/2.24	-2.68/2.64
500	0.9	-2.17/2.13	-2.12/2.08	-2.19/2.19	-2.20/2.21	-2.12/2.17	-2.18/2.17	-2.24/2.23	-2.31/2.31	-2.84/2.85
<i>Bartlett</i>										
1000	0.1	-2.13/2.11	-2.10/2.01	-2.09/2.06	-2.10/2.10	-2.12/2.10	-2.09/2.08	-2.10/2.11	-2.14/2.17	-2.22/2.23
1000	0.2	-2.07/2.07	-2.10/2.15	-2.14/2.05	-2.12/2.05	-2.10/2.11	-2.11/2.14	-2.11/2.11	-2.03/2.12	-2.23/2.18
1000	0.8	-2.02/2.01	-2.00/2.01	-2.03/2.05	-2.04/2.04	-2.18/2.08	-2.20/2.11	-2.19/2.15	-2.20/2.23	-2.31/2.31
1000	0.9	-1.94/1.97	-2.04/2.01	-2.00/2.03	-2.11/2.00	-2.09/2.08	-2.15/2.14	-2.21/2.21	-2.28/2.27	-2.44/2.44
<i>Quadratic spectral</i>										
1000	0.1	-2.17/2.12	-2.14/2.18	-2.16/2.17	-2.14/2.13	-2.13/2.11	-2.08/2.13	-2.18/2.19	-2.15/2.17	-2.31/2.27
1000	0.2	-2.15/2.12	-2.18/2.14	-2.14/2.12	-2.17/2.15	-2.17/2.23	-2.19/2.15	-2.18/2.12	-2.18/2.19	-2.34/2.34
1000	0.8	-2.19/2.14	-2.11/2.17	-2.11/2.19	-2.18/2.15	-2.20/2.13	-2.11/2.12	-2.17/2.15	-2.22/2.28	-2.34/2.31
1000	0.9	-2.21/2.16	-2.18/2.12	-2.19/2.12	-2.17/2.21	-2.23/2.23	-2.19/2.20	-2.21/2.19	-2.27/2.22	-2.34/2.38

Using the Bartlett and Quadratic spectral methods, this table reports dummy variable coefficient critical cut-off t -statistic values for dummy variable regression model, $r_t = \alpha + \beta\pi_t + u_t$ under different combinations of autocorrelation of the dependent variable and transition probability of the dummy variable. Autocorrelation of the dummy variable is equal to two times the dummy variable transition probability minus one ($2q - 1$). Lower and upper bounds (i.e., critical cut-off values) of the 95% confidence interval of t -statistics adjusted for spurious regression bias are reported. The monthly mean return of the simulated dependent variable series is zero and the monthly standard deviation is one. For the Bartlett method, the fixed lag length HAC standard error estimation sets the lag length (M) as a fixed function of the sample size T ($M = bT$), where b is a constant and is set at 0.05.

Table 5
Political regime volatility differential regression results.

Variable	Sample size (n)	Parameter estimates and lower/upper cut-off levels				
		α	$t(\alpha)$	β	$t(\beta)$	\bar{R}^2
<i>DJIA volatility</i>	28,987	0.97	52.68 (10.62)	-0.11	-4.59 (-1.19)	0.96%
	Spurious regression bias	0.80/1.03	56.77/81.75	-0.12/0.12	-6.27/6.21	1.07%
	Spurious regression bias and data mining	0.76/1.07	53.82/86.32	-0.15/0.16	-8.21/8.33	1.89%

This reports OLS regression of daily return volatility on Democratic partisanship in Leblang and Mukherjee (2005), $Y_t = \alpha + \beta D_t + \mu_t$, where y_t denotes *DJIA volatility* at day t , *DJIA volatility* is the twenty-day moving standard deviation of daily return on the DJIA. D_t is 1 if a democratic president is in power and 0 otherwise. The sample period begins in May 26, 1896 and extends through December 31, 2001. The data are obtained from Global Financial Data. \bar{R}^2 denotes adjusted R^2 . Regressions are estimated by OLS and t -statistics are adjusted for autocorrelation and heteroskedasticity using Newey and West (1987). t -statistics in parentheses are adjusted using fixed lag length HAC standard error estimation by setting the lag length (M) as a fixed function of the sample size T ($M = bT$), where b is a constant set at 0.05.

ence in seasonal returns is found to be significant in only 6 (in contrast to 20) of the 37 countries examined by BJ once spurious regression and data mining are taken into account.

3.3. Business cycle regimes and momentum profits

Two recent studies by Chordia and Shivakumar (2002) and Cooper et al. (2004) examine how momentum profits are affected by business cycle regimes. On face appearance, we have the necessary ingredients for a spurious regression bias problem. Business cycles (i.e., the dichotomous independent variable) are known to be highly persistent, and measured momentum returns are positively autocorrelated due to the use of overlapping observations.

Chordia and Shivakumar (2002) (hereafter, CS) explore whether momentum profits can be explained by macroeconomic variables in an attempt to provide a “risk-based” explanation of momentum profits in contrast to the prevailing “investor irrationality” explanation. The CS explanation of momentum profits builds upon their finding that momentum profits are only positive during periods of economic expansion. To reach this conclusion, they (implicitly) use dummy variable regression model (1). The dependent variable in their study is cumulative abnormal returns to the past winner minus past loser momentum strategy for all CRSP stocks during the sample period July 1926 to December 1994, where cumulative abnormal returns are measured over the six-month period subsequent to each month t and are, therefore, overlapping. The dichot-

Table 6
Halloween indicator regression results for countries with significant t-statistics reported by Bouman and Jacobsen (2002).

Country	n	Mean	Standard deviation	Auto-correlation	Parameter estimates and lower/upper cut-off levels				
					α	$t(\alpha)$	β	$t(\beta)$	R^2
Austria	344	0.66	5.41	0.153	-0.13	-0.30	1.57	2.71	1.82%
	Spurious regression bias				-0.26/1.58	-0.07/3.94	-1.25/1.27	-2.17/2.18	1.07%
	Spurious regression bias and data mining				-0.51/1.86	-1.26/4.68	-1.66/1.64	-2.86/2.86	2.04%
Belgium	344	1.18	4.73	0.157	0.02	0.06	2.31	4.67	5.72%
	Spurious regression bias				0.38/1.99	1.07/5.69	-1.09/1.12	-2.18/2.21	1.09%
	Spurious regression bias and data mining				0.12/2.23	0.33/6.45	-1.45/1.51	-2.87/2.99	2.17%
Brazil	128	15.38	21.9	0.172	11.61	4.73	7.53	1.97	2.21%
	Spurious regression bias				9.23/21.41	3.39/8.37	-8.12/8.22	-2.14/2.17	2.73%
	Spurious regression bias and data mining				7.29/23.22	2.64/9.25	-10.81/10.75	-2.89/2.87	5.30%
Canada	344	0.83	4.99	0.015	0.26	0.67	1.14	2.12	1.01%
	Spurious regression bias				0.08/1.58	0.22/4.25	-1.05/1.04	-1.98/1.98	0.84%
	Spurious regression bias and data mining				-0.15/1.8	-0.40/4.93	-1.36/1.43	-2.53/2.68	1.63%
France	344	1.03	6.02	0.078	-0.13	-0.27	2.31	3.62	3.41%
	Spurious regression bias				0.08/2	0.17/4.48	-1.32/1.36	-2.07/2.12	0.96%
	Spurious regression bias and data mining				-0.24/2.29	-0.52/5.17	-1.75/1.84	-2.78/2.85	1.98%
Germany	344	0.78	5.3	0.06	0.08	0.19	1.38	2.44	1.42%
	Spurious regression bias				-0.04/1.58	-0.10/4.02	-1.15/1.15	-2.04/2.03	0.90%
	Spurious regression bias and data mining				-0.29/1.85	-0.75/4.75	-1.49/1.55	-2.63/2.71	1.77%
Greece	128	2.12	10.8	0.114	0.45	0.32	3.34	1.77	1.64%
	Spurious regression bias				-0.57/4.85	-0.46/4.01	-3.72/3.75	-2.14/2.15	2.70%
	Spurious regression bias and data mining				-1.51/5.68	-1.21/4.83	-4.89/4.92	-2.80/2.86	5.21%
Ireland	128	1.25	5.83	0.111	-0.05	-0.06	2.6	2.57	4.24%
	Spurious regression bias				-0.28/2.77	-0.39/3.98	-2.18/2.14	-2.16/2.13	2.71%
	Spurious regression bias and data mining				-0.75/3.25	-1.08/4.71	-2.89/2.83	-2.88/2.85	5.25%
Italy	344	0.91	7.13	0.074	-0.44	-0.75	2.7	3.56	3.30%
	Spurious regression bias				-0.22/2.06	-0.41/3.88	-1.57/1.61	-2.06/2.11	0.95%
	Spurious regression bias and data mining				-0.59/2.4	-1.11/4.56	-2.07/2.18	-2.77/2.84	1.96%
Japan	344	0.7	5.42	0.057	-0.06	-0.14	1.52	2.62	1.69%
	Spurious regression bias				-0.13/1.52	-0.33/3.78	-1.17/1.18	-2.03/2.03	0.89%
	Spurious regression bias and data mining				-0.40/1.79	-0.98/4.5	-1.52/1.58	-2.62/2.71	1.76%
Malaysia	128	0.17	8.67	0.033	-1.12	-1.01	2.59	1.70	1.47%
	Spurious regression bias				-1.95/2.33	-1.90/2.29	-3.06/3.16	-2.05/2.09	2.46%
	Spurious regression bias and data mining				-2.63/3.06	-2.52/3.05	-4.08/4.16	-2.73/2.76	4.82%
Netherlands	344	1.12	4.95	0.071	0.18	0.45	1.88	3.58	3.33%
	Spurious regression bias				0.34/1.92	0.91/5.21	-1.08/1.12	-2.05/2.11	0.95%
	Spurious regression bias and data mining				0.08/2.15	0.24/5.92	-1.44/1.51	-2.76/2.83	1.96%
Philippines	128	0.98	9.23	0.146	-0.05	-0.04	2.07	1.27	0.48%
	Spurious regression bias				-1.44/3.46	-1.29/3.21	-3.47/3.42	-2.18/2.18	2.81%
	Spurious regression bias and data mining				-2.27/4.24	-2.06/4	-4.55/4.51	-2.87/2.93	5.34%
Singapore	344	0.67	8.39	0.123	-0.25	-0.40	1.84	2.05	0.92%
	Spurious regression bias				-0.71/2.07	-1.14/3.32	-1.90/1.95	-2.13/2.17	1.04%
	Spurious regression bias and data mining				-1.16/2.49	-1.88/4.02	-2.52/2.63	-2.84/2.92	2.08%
Spain	344	1.06	6.04	0.115	0.12	0.26	1.88	2.92	2.15%
	Spurious regression bias				0.08/2.05	0.18/4.58	-1.37/1.36	-2.13/2.11	0.99%
	Spurious regression bias and data mining				-0.22/2.36	-0.47/5.28	-1.80/1.8	-2.78/2.83	1.93%
Sweden	344	1.39	6.15	0.116	0.31	0.66	2.17	3.32	2.84%
	Spurious regression bias				0.38/2.41	0.83/5.29	-1.39/1.42	-2.11/2.16	1.03%
	Spurious regression bias and data mining				0.06/2.72	0.12/6.01	-1.84/1.93	-2.83/2.91	2.07%

(continued on next page)

Table 6 (continued)

Country	n	Mean	Standard deviation	Auto-correlation	Parameter estimates and lower/upper cut-off levels			R ²	
					α	$t(\alpha)$	β		$t(\beta)$
Switzerland	344	0.82	5	0.063	0.28	0.66	1.08	2.01	0.88%
	Spurious regression bias				0.05/1.58	0.13/4.26	-1.09/1.09	-2.04/2.03	0.90%
	Spurious regression bias and data mining				-0.20/1.83	-0.52/4.99	-1.41/1.47	-2.63/2.72	1.78%
Taiwan	128	0.78	12.24	0.15	-2.01	-1.33	5.57	2.63	4.46%
	Spurious regression bias				-2.59/4.19	-1.76/2.84	-4.60/4.67	-2.19/2.20	2.86%
	Spurious regression bias and data mining				-3.74/5.25	-2.56/3.61	-6.04/6.12	-2.85/2.91	5.40%
United Kingdom	344	1.17	6.12	0.083	0.16	0.36	2.02	3.1	2.46%
	Spurious regression bias				0.19/2.12	0.43/4.69	-1.35/1.36	-2.04/2.09	0.94%
	Spurious regression bias and data mining				-0.12/2.42	-0.25/5.42	-1.75/1.78	-2.62/2.71	1.77%
United States	344	0.96	4.42	0.012	0.50	1.38	0.92	1.95	0.81%
	Spurious regression bias				0.30/1.64	0.89/4.98	-0.93/0.96	-1.97/2.03	0.87%
	Spurious regression bias and data mining				0.08/1.84	0.25/5.65	-1.24/1.30	-2.64/2.77	

1.79%

OLS regression of monthly MSCI value-weighted gross index log returns for various countries on the Halloween dummy variable, $R_t = \alpha + \beta HW_t + u_t$, where R_t denotes monthly return at time t for a country and is expressed in % form, HW_t is 1 if month t falls during the period November through April and 0 otherwise. The sample period is January 1970 to August 1998 for developed markets and January 1988 to August 1998 for emerging markets. R^2 denotes adjusted R^2 . Like Bouman and Jacobsen (2002), all data in the table are obtained from MSCI. Also like Bouman and Jacobsen (2002), regressions are estimated by OLS and t -statistics are adjusted for heteroskedasticity using White's correction. Since Bouman and Jacobsen (2002) only adjust t -statistics using White's correction, there is no need to adjust t -statistics using fixed lag length HAC standard error estimation. Lower and upper cut-off levels are 95% confidence bounds obtained by simulation. Results for the Philippines and Malaysia differ from those reported in Bouman and Jacobsen (2002), presumably due to data revisions by MSCI.

omous variable used in their study is 1 for months of economic expansion as defined by the National Bureau of Economic Research (and reproduced in Table 2 of CS) and 0 for months of economic retraction.¹⁸ Chordia and Shivakumar (2002, p. 992, Table II) provide a summary.

CS report that the 1.25% difference in momentum return between business cycle upturns and downturns is economically and statistically significant. For purposes of comparison, their results are reported in Panel A of Table 7. Panel B provides the cut-off t -statistics adjusted for spurious regression bias and spurious regression bias and data mining. The transition probabilities for the simulated dummy variable series are determined from the monthly NBER expansionary/retractionary definitions. The simulated dependent variable series has the same mean, standard deviation and autocorrelation as the momentum profit cumulative abnormal return series, *UMD6M*, reported in Panel B of Table 7.¹⁹ Panel B also contains the lower/upper 95% confidence bounds for the t -statistic. As the results show, the reported t -statistic of 2.10 for the test of differences in momentum return across the business cycle falls comfortably within the bounds of (-2.45, 2.48) and (-3.35, 3.41) for the spurious regression bias and the spurious regression bias and data mining simulations, respectively. The return difference between business cycle upturns and downturns is therefore insignificant from a statistical standpoint. Again, it is worthwhile to note that the matrix of critical t -statistics reported in Table 2 confirms this conclusion. Using a dependent variable autocorrelation of 0.8, a dummy variable transition probability of 0.95, and a sample size of 1000, the cut-off t -statistic values in Panels A and B are (-2.50, 2.48) and (-3.31, 3.27), respectively. A bootstrap procedure also provides (unreported) results that are essentially the same as those obtained using the simulation procedure.

Cooper et al. (2004) (hereafter, CGH) examine whether momentum profits are affected by the state of the market. Like CS, the dependent variable in the CGH study includes cumulative abnormal returns to the past winner minus past loser momentum strategy for all CRSP stocks over the six-month period subsequent to each month t . CGH use the sample period January 1929 through December 1995. Again, since monthly data are used in the time-series tests, the dependent variable will be autocorrelated due to overlapping observations. The independent variable is a state-of-the-market dummy variable whose value is 1 in month t if the total return to the CRSP value-weighted index is positive during the preceding three years and 0 otherwise. Cooper et al. (2004, p. 1352) report that momentum profits following UP markets are statistically greater than DOWN markets. For the six-month cumulative return, for example, the reported t -statistic is 2.26, well above the critical level in conventional tests. The monthly returns and t -statistic are reported in Panel A of Table 8.

To determine whether spurious regression bias has affected statistical inference in the CGH study, we again develop 95% confidence bounds that explicitly account for the autocorrelation in the dependent variable and the transition probabilities of the state-of-the-market dummy. The transition probabilities for the simulated state-of-the-market dummy variable series are computed using the CRSP value-weighted index. The simulated dependent variable series has the same mean, standard deviation and

¹⁸ In a recent, related study, Basintha and Kurov (2008) examine cyclical variation in the effect of Fed policy on the stock market.

¹⁹ We use the monthly Fama and French momentum factor to calculate the overlapping six-month momentum return properties. The data are available at http://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library.html. The Fama and French momentum portfolio is not identical to the decile past winner minus past loser momentum portfolio, so we also convert the daily autocorrelation of the decile momentum factor reported in Yan (2005) to monthly autocorrelation and get an autocorrelation estimate of 0.7 which is quite close to the estimate we obtained using the Fama and French momentum factor.

Table 7
Results for Chordia and Shivakumar (2002).

Panel A: Chordia and Shivakumar (2002) result					
		Expansionary	Contractionary	Difference	t-statistic
Average monthly profit		0.53	−0.72	1.25	2.10
Panel B: Simulation results					
Variable	n	Mean	Standard deviation	Auto-correlation	Lower/upper cut-off t-statistics
UMD6M	811	3.64	13.03	0.845	
		Spurious regression bias			−2.45/2.48
		Spurious regression bias and data mining			−3.35/3.41

Panel A contains results reported in Chordia and Shivakumar (2002) that correspond to an OLS regression of six-month momentum profits on a business cycle dummy variable, $UMD6M_t = \alpha + \beta_t BC_t + u_t$, where $UMD6M_t$ denotes cumulative six-month return, expressed in percentage form, of the winner minus loser momentum strategy for all CRSP stocks. Panel B contains simulation results for the OLS regression of six-month momentum profits where $UMD6M_t$ is the cumulative six-month return of the Fama–French monthly momentum factor at time t and is expressed in percentage. BC_t in both panels is 1 if month t falls into expansionary periods and 0 otherwise, and expansionary and contractionary periods are classified based on the NBER published business cycle dates. While the Chordia and Shivakumar (2002) sample period is July 1926 to December 1994, the Fama and French monthly momentum factor is available only from January 1927. Cut-off t -statistics are adjusted for autocorrelation and heteroskedasticity using Newey and West (1987).

Table 8
Results for Cooper et al. (2004).

Part A: Cooper et al. (2004)					
		UP	DOWN	Difference	t-statistic
Average monthly profit		0.93	−0.37	1.30	2.26
Panel B: Simulation results					
Variable	n	Mean	Standard deviation	Auto-correlation	Lower/upper cut-off t-statistics
UMD6M	804	3.55	13.04	0.844	
		Spurious regression bias			−2.70/2.58
		Spurious regression bias and data mining			−3.68/3.58

Panel A contains results reported in Cooper et al. (2004) that correspond to an OLS regression of six-month momentum profits on a state-of-the-market dummy variable, $UMD6M_t = \alpha + \beta_t SOM_t + u_t$, where $UMD6M_t$ denotes cumulative six-month return, expressed in percentage form, of the winner minus loser momentum strategy for all CRSP stocks. Panel B contains simulation results for the OLS regression of six-month momentum profits where $UMD6M_t$ is the cumulative six-month return of Fama–French monthly momentum factor at time and is expressed in percentage form, SOM in both panels is 1 if the lagged three-year market return is non-negative and 0 otherwise, UP is when the lagged three-year CRSP value-weighted index return is non-negative, and DOWN is when the lagged three-year CRSP value-weighted index return is negative. The sample period is January 1929 and December 1995. Cut-off t -statistics are adjusted for autocorrelation and heteroskedasticity using Newey and West (1987).

autocorrelation as the momentum profit cumulative abnormal return series, $UMD6M$, reported in Panel B of Table 8. The simulated confidence bounds reported in Panel B of Table 8 indicate that the t -statistics obtained by CGH are less than would be expected by chance. The reported t -statistic for the CGH test for equality between cumulative abnormal returns in up and down market states, 2.26 for the six-month momentum strategy, is well within the simulated bounds (−2.70, 2.58) for spurious regression bias for dependent variable autocorrelation of .8, a dummy variable transition probability of .95, and a sample size of 750. Once again, the statistical inference is reversed. Contrary to CGH, the observed return difference between up and down markets is less than would be expected by chance. The critical confidence bounds reported in Table 2 provide a convenient means of confirming this conclusion. For dependent variable autocorrelation of 0.8, a dummy variable transition probability of 0.95, and a sample size of 1000, the bounds are (−2.50, 2.48) and (−3.31, 3.27) for spurious regression bias and spurious regression bias and data mining, respectively. Once again, a bootstrap procedure also provides (unreported) results that are essentially the same as those obtained using the simulation procedure.

3.4. Monetary aggregates and interest rate regimes

Another interesting issue to consider in exploring the influence of regression variable autocorrelation on tests for regime differences is raised by Jensen et al. (1996) (hereafter, JMJ) when they examine differences in the behavior of monetary aggregates across interest rate regimes. The interest rate regime dummy variable used in the study equals 1 if the last change in the discount rate is positive and 0 otherwise. Since the number of changes in the

discount rate each year is generally few, a dummy variable based on monthly intervals is persistent by construction. Indeed, as Table I shows, the autocorrelation of the dummy variable used in the JMJ study is .900. At the same time, monetary aggregates are also strongly persistent. None of the four JMJ dependent variables considered in Table 1 has autocorrelation below .872.

JMJ test the difference between the median levels of each monetary aggregate under expansive (discount rate decrease) and restrictive (discount rate increase) monetary policy regimes during the period 1954 through 1992 and conclude that each difference is statistically significant using conventional test methodologies. To examine the effect of autocorrelation on their test, we regress the monthly levels of monetary aggregates on a dummy variable whose value is 1 if the most recent interest rate change was an increase and 0 otherwise. We use only four of the five economic series examined by JMJ – seasonally-adjusted money supply ($M1SL$), seasonally-adjusted monetary base ($AMBSL$), excess reserves ($EXCRESNS$), and the federal funds premium ($FFPREM$). We do not include the seasonally-adjusted Fed credit variable since the series has been discontinued and is no longer available in its entirety. All data are from the Federal Reserve Bank of Saint Louis. The results are reported in Table 9.

The results in Table 9 are interesting in a number of respects. First, two of the four series, seasonally-adjusted money supply ($M1SL$) and seasonally-adjusted monetary base ($AMBSL$), are not different in a statistical sense under expansive (discount rate decrease) and restrictive (discount rate increase) monetary policy regimes. Both of the reported t -statistics are within the conventional bands (−1.96, 1.96). Apparently the fact that we test the difference between the means in the two regimes while JMJ test the difference between medians produces different results. Second,

Table 9
Simulation results for Jensen et al. (1996).

Variable	Sample size (<i>n</i>)	Mean	Standard deviation	Auto-correlation	Parameter estimates and lower/upper cut-off levels				
					α	$t(\alpha)$	β	$t(\beta)$	R^2
M1SL	408	385.29	244.88	0.974	423.40	9.79 (5.90)	-72.65	-1.32 (-0.87)	1.96%
	Spurious regression bias				174.78/579.45	4.05/21.72	-168.99/172.52	-3.49/3.6	13.23%
	Spurious regression bias and data mining				101.97/668.96	2.27/25.9	-231.64/241.25	-4.69/4.78	21.79%
AMBSL	467	117.62	89.01	0.974	132.33	8.72 (5.10)	-27.48	-1.47 (-0.97)	2.17%
	Spurious regression bias				45.03/192.42	3.19/19.57	-58.73/60.45	-3.51/3.6	11.80%
	Spurious regression bias and data mining				19.92/215.91	1.32/23.48	-78.79/81.47	-4.70/4.62	18.76%
EXCRESNS	467	0.5	0.3	0.938	0.57	11.57 (6.92)	-0.12	-2.01 (-1.38)	3.74%
	Spurious regression bias				0.32/0.68	6.09/19.29	-0.19/0.19	-3.23/3.2	9.20%
	Spurious regression bias and data mining				0.26/0.74	4.73/22.28	-0.24/0.25	-4.26/4.23	15.10%
FFPREM	462	0.58	0.84	0.872	0.37	4.46 (2.93)	0.39	2.66 (-1.90)	5.22%
	Spurious regression bias				0.21/0.96	1.57/9.1	-0.45/0.45	-2.88/2.84	6.56%
	Spurious regression bias and data mining				0.07/1.09	0.57/10.58	-0.59/0.6	-3.86/3.79	11.39%

OLS regression of economic variables on a discount rate change dummy variable, $EV_t = \alpha + \beta DIR_t + u_t$, where EV_t denotes monthly economic variables (*M1SL*, *AMBSL*, *EXCRESNS* and *FFPREM*) at time t , *M1SL* is seasonally-adjusted money supply, *AMBSL* is seasonally-adjusted, monetary-base, *EXCRESNS* is excess reserves, *FFPREM* is the federal funds premium and is calculated as the difference between the federal funds rate and the three-month T-bill rate, and *DIR_t* is 1 if the previous change in the discount rate is an increase and 0 otherwise. The sample period is February 1954 through December 1992, except *M1SL* and *FFPREM* which begin January 1959 and August 1954, respectively. R^2 denotes adjusted R^2 . All data are obtained from the Federal Reserve Bank at St. Louis. Regressions are estimated by OLS and t -statistics are adjusted for autocorrelation and heteroskedasticity using Newey and West (1987). t -statistics in parentheses are adjusted using fixed lag length HAC standard error estimation by setting the lag length (M) as a fixed function of the sample size T ($M = bT$), where b is a constant set at 0.05. Note that we do not include the seasonally-adjusted Fed-credit (*ACFS*) variable used in Jensen et al. (1996). This series has been discontinued and is no longer available in its entirety.

two of the four series, excess reserves (*EXCRESNS*) and the federal funds premium (*FFPREM*), are significantly different under expansive and restrictive monetary policy regimes using standard OLS inference. Once we account for spurious regression bias, however, the difference becomes insignificant. The 95% confidence bounds are (-3.23, 3.20) and (-4.26, 4.23) for the spurious regression bias and the spurious regression bias and data mining simulations, respectively, for *EXCRESNS*, and (-2.88, 2.84) and (-3.86, 3.79) for the spurious regression bias and the spurious regression bias and data mining simulations, respectively, for *FFPREM*. The final column of Table 9 provides, perhaps, the sharpest perspective, since it indicates that an R^2 of close to 10% or higher would be expected to be obtained by chance when regressing uncorrelated dependent and independent series that have the same extreme persistence as the data, yet the actual R^2 's obtained are 5% or less. Again, the critical t -statistics reported in Table 2 further reinforce this conclusion. The Federal funds premium series (*FFPREM*) has autocorrelation of 0.9, the interest rate transition dummy variable transition probability of 0.95, and the sample size of nearly 500. The 95% confidence bounds are (-3.05, 3.10) and (-4.20, 4.15) for spurious regression bias and spurious regression bias and data mining, respectively. The fixed lag length HAC t -statistics reported in parentheses in Table 9 strongly reinforce the conclusion that the reported differences are due to serial correlation, since the reported t -statistics fall by at least a third and clearly become insignificant when the HAC estimator lag length constant is set equal to .05.

Again, the critical t -statistics generated by our simulation procedure reverse the statistical inferences that monetary aggregates behave differently across interest rate regimes, thus undermining the premise upon which JMJ's subsequent analysis of return predictability across monetary regimes is based. This provides the sharpest illustration so far of the need to take into account regression variable autocorrelation in dichotomous variable regression analysis, since doing so apparently reverses the study's conclusions even though the results appear to be highly significant under conventional tests of significance.

4. Binary regime dependent variables

The last two sections focus on testing for regime differences when regimes are persistent. While the most common use of re-

gime dummy variables is as regressors, a natural question to ask at this point is whether problems can exist when testing hypotheses in situations in which the dependent variable in a time-series regression is a regime dummy. Dichotomous dependent variables require regression analysis within the binary dependent variable (i.e., probit or logit) framework. In this section, we alter the simulation procedure to answer the question.

The binary dependent variable Probit regression model is

$$P(\pi_t = 1 | r_t) = N(\alpha + \beta r_t), \quad (3)$$

where P denotes probability, π_t is a binary dependent variable equal to 1 if a particular event or regime is occurring and 0 otherwise, N is the cumulative normal distribution, and r_t is a continuous explanatory variable. The binary variable and explanatory variable are independently simulated following the procedure outlined in Sections 2.1 and 2.2. The only difference is that the simulated binary variable series now plays the role of the dependent variable in the simulation procedure and the simulated continuous variable is the explanatory variable. The mean and standard deviation of the explanatory variable series are 0 and 1, respectively.

Again, 10,000 simulation runs are used to create 95% critical values. Table 10 contains the critical z -statistic values. The transition probability of the binary dependent variable is represented on the x -axis of the table. Recall the autocorrelation of this binary variable may be computed as $2q - 1$. The autocorrelation of the continuous explanatory variable is represented on the y -axis of the table. The critical z -statistics are reported for three sample sizes - 250, 500, and 1000. Panel A of Table 10 contains the critical z -statistic interval adjusting for spurious regression, and Panel B contains the critical z -statistic interval adjusted for data mining as well as spurious regression.

To illustrate the use of the table, assume a binary dependent variable has a transition probability of 0.99 (i.e., autocorrelation of 0.98), the explanatory variable has autocorrelation of 0.1, and the sample size is 250. The 95% confidence bounds are (-2.13, 2.11) and (-2.75, 2.75) for spurious regression bias and spurious regression bias and data mining, respectively, substantially wider than the conventional level. Further increases in explanatory variable autocorrelation have a huge influence on critical z -statistics when the binary dependent variable is also positively autocorrelated. An increase in inde-

Table 11
Simulation results for Resnick and Shoesmith (2002).

Independent variable, $SPREAD_{t-k}$	Parameter estimates			
	α	$Z(\alpha)$	β	$t(\beta)$
k				
1	-0.33	-3.69	-0.31	-5.58
2	-0.34	-3.81	-0.30	-5.49
3	-0.36	-4.00	-0.29	-5.31
4	-0.39	-4.27	-0.27	-5.00
5	-0.42	-4.65	0.24	-4.55
8	-0.48	-5.21	-0.21	-3.96
	Lower/upper cut-off levels			
Spurious regression bias	-1.83/-0.08	-13.30/-0.77	-0.44/0.42	-6.50/6.30
Spurious regression bias and data mining	-2.28/0.20	-14.27/1.76	-0.61/0.61	-8.10/8.13

Probit analysis of predictive power of lagged term structure spread on bear market binary variable, $P(BEAR_t = 1 | SPREAD_{t-k}) = N(\alpha + \beta SPREAD_{t-k})$, where P indicates probability, $BEAR_t$ is a binary dependent variable at month t and is 1 if the six month trend in monthly returns is negative and 0 otherwise, N is the standard normal distribution, and $SPREAD_{t-k}$ is the slope of the yield curve $t-k$ months earlier ($k = 1, 2, 3, 4, 5$, and 8) and is calculated as the difference between the monthly 10-year government bond rate and the monthly three-month T-Bill rate (expressed in % form). Data are obtained from the Federal Reserve Bank St. Louis.

The differences in sample size indicate that high cut-off intervals do not diminish as the sample size increases when the dependent variable is persistent, a finding that is consistent with Ferson et al. (2003) and illustrates the difference between autocorrelation in the dependent variable versus the independent variable (as indicated by a comparison of Table 10 with Table 2). Interestingly, the simulated confidence intervals actually tighten sharply when the binary variable autocorrelation becomes increasingly negative (i.e., transition probabilities less than 0.5), thus implying that conventional confidence intervals are considerably overstated. Finally, the simulated critical values almost exactly match conventional cut-offs when the dependent variable is not autocorrelated (the transition probability equals 0.5).

To illustrate the effect of regression variable autocorrelation for binary dependent variables, the simulation procedure is used to test the hypothesis that the term structure can be used to forecast bear markets. Clearly, data mining is a concern, since the goal of predicting bear markets has been an important and potentially lucrative area of endeavor.

Resnick and Shoesmith (2002) use the difference between the ten-year government bond yield and the three-month treasury bill yield, lagged by various time intervals from one month to eight months, to predict bear markets (where the bear market binary variable is defined as being equal to 1 if the six month trend in monthly returns is negative, and 0 otherwise.²⁰ They find that a sharp reduction in the yield spread foreshadows bear markets in their sample period January 1960 through December 1999, with the yield curve lagged by either one month or two months having the best predictive ability out of sample.²¹ Table 11 provides results for the performance of their model within sample. Table 11 reveals that z-scores for the independent yield spread variable coefficient range from -3.96 for an eight-month lag of the yield spread to -5.58 for the shortest lag (one month), all of which imply a highly significant relation between the yield spread and subsequent recessions by conventional standards. Table 11 indicates, however, that the z-scores for the independent yield spread variable coefficient sit well within the critical z-statistic intervals of (-6.50, 6.30) and (-8.10, 8.13). The simulation procedure z-statistic critical values provided in Table 10 ($n = 500$) also lead to a conclusion of no significant predictability, since explanatory variable autocorrelation of 0.9 and a dependent variable transition probability of 0.95 correspond to critical cut-off z-statistic intervals of (-5.79, 5.82) and (-7.37, 7.27) for spurious regression bias and spurious regression bias and data

mining, respectively. Regression variable autocorrelation, therefore, has an important influence on statistical inference in this probit model and raises serious doubts about whether changes in the yield curve predict bull and bear markets.

5. Advice to researchers

The results of this study show the dramatic effects that persistent dummy variables can have on standard OLS tests of significance. This section offers researchers some guidance on how to avoid the trap when using dichotomous random variables in a time-series regression framework. The first, and most obvious, suggestion is to report summary statistics for the dichotomous regression variables in the same manner as continuous regression variables. In this way, readers will be alerted to the prospect of a spurious regression problem. With the levels of autocorrelation and persistence known (or estimated), the critical t -statistics provided in Tables 2 and 10 can be used to correctly assess statistical inference.

A second suggestion is to use fixed lag length HAC standard estimation (with a fixed lag length constant $b = .05$) since it provides an improvement relative to other HAC standard error adjustment procedures recommended in the literature. Whenever HAC estimators are employed to estimate t -statistics, statistical significance should be assessed analytically using the asymptotic cut-off values obtained from Table 3 (e.g., for a lag equal to a fixed function of the sample size, bT , where $b = 0.05$ and a Bartlett kernel, the cut-off is 2.11 rather than 1.96), thus analytically taking into account the lag length that is employed.

A third suggestion for situations where the dependent variable is, itself, highly autocorrelated is to include the lagged value of the dependant variable as a regressor.²² Table 12 shows that this provides a close approximation to the asymptotic cutoffs from Table 3 (2.11 in the case of the Bartlett kernel and sample size = $.05T$) even in cases where the dichotomous random variable is negatively correlated.

A fourth suggestion is to try to avoid situations where the dichotomous variable is, by construction, autocorrelated. In presidential regime studies, for example, only one value of the presidential dummy variable is observed each four years. Consequently, only the total return over the presidential term is relevant. A simpler, more intuitive, approach is to run the regression model using one observation every four years, thus greatly reducing the dependent variable autocorrelation. Similarly, the use of overlapping dependent variable observations should be avoided whenever possible. Finally, where data mining concerns are potentially relevant,

²⁰ Candelon et al. (2008) examine bull and bear markets and co-movements in market returns in five Asian stock markets.

²¹ In a related paper, Chen (2009) investigates whether macroeconomic announcements can predict recessions in the stock market.

²² We thank the referee for offering this alternative approach.

Table 12
Cut-off dummy variable t -statistics for dependent and dummy variable autocorrelation combinations with lagged dependent variable as a regressor using the Bartlett Method.

Sample size (n)	Auto-correlation of dependent variable	Transition probability of dummy variable (q)								
		0.01	0.05	0.10	0.25	0.50	0.75	0.90	0.95	0.99
<i>Panel A: Cut-offs for spurious regression bias</i>										
250	0.1	-2.10/2.01	-2.06/2.06	-2.03/2.07	-2.11/2.12	-2.13/2.11	-2.14/2.07	-2.11/2.16	-2.20/2.18	-2.59/2.58
250	0.2	-2.11/2.08	-2.04/2.12	-2.04/2.06	-2.13/2.09	-2.09/2.08	-2.14/2.13	-2.11/2.17	-2.13/2.18	-2.66/2.56
250	0.3	-2.11/2.11	-2.07/2.05	-2.09/2.11	-2.14/2.08	-2.14/2.07	-2.17/2.10	-2.12/2.10	-2.18/2.22	-2.53/2.66
250	0.4	-2.11/2.09	-2.09/2.09	-2.11/2.13	-2.12/2.10	-2.17/2.09	-2.13/2.04	-2.12/2.14	-2.17/2.17	-2.55/2.61
250	0.5	-2.09/2.10	-2.12/2.07	-2.11/2.14	-2.08/2.08	-2.13/2.12	-2.17/2.10	-2.13/2.17	-2.17/2.09	-2.65/2.58
250	0.6	-2.07/2.09	-2.14/2.08	-2.10/2.14	-2.13/2.09	-2.16/2.10	-2.17/2.06	-2.14/2.09	-2.18/2.16	-2.58/2.59
250	0.7	-2.11/2.08	-2.12/2.14	-2.13/2.13	-2.08/2.16	-2.15/2.09	-2.15/2.14	-2.15/2.19	-2.15/2.17	-2.54/2.67
250	0.8	-2.13/2.11	-2.06/2.12	-2.08/2.10	-2.03/2.12	-2.13/2.11	-2.11/2.13	-2.22/2.15	-2.22/2.14	-2.67/2.58
250	0.9	-2.09/2.07	-2.09/2.14	-2.06/2.14	-2.11/2.07	-2.13/2.05	-2.12/2.16	-2.19/2.17	-2.24/2.18	-2.70/2.74
500	0.1	-2.08/2.09	-2.14/2.05	-2.07/2.07	-2.11/2.16	-2.10/2.09	-2.08/2.05	-2.04/2.11	-2.11/2.12	-2.26/2.28
500	0.2	-2.12/2.13	-2.10/2.05	-2.13/2.10	-2.10/2.11	-2.07/2.06	-2.08/2.08	-2.12/2.13	-2.16/2.13	-2.30/2.33
500	0.3	-2.12/2.11	-2.08/2.07	-2.11/2.10	-2.12/2.10	-2.03/2.06	-2.11/2.09	-2.12/2.10	-2.11/2.10	-2.34/2.31
500	0.4	-2.12/2.07	-2.10/2.06	-2.05/2.11	-2.14/2.11	-2.03/2.12	-2.12/2.11	-2.07/2.13	-2.14/2.11	-2.33/2.33
500	0.5	-2.12/2.10	-2.10/2.10	-2.04/2.09	-2.06/2.11	-2.04/2.13	-2.08/2.05	-2.11/2.18	-2.16/2.14	-2.27/2.29
500	0.6	-2.12/2.05	-2.12/2.11	-2.13/2.13	-2.07/2.11	-2.07/2.09	-2.08/2.09	-2.12/2.12	-2.15/2.20	-2.33/2.24
500	0.7	-2.12/2.10	-2.10/2.08	-2.11/2.15	-2.03/2.08	-2.06/2.08	-2.09/2.16	-2.11/2.12	-2.14/2.18	-2.21/2.28
500	0.8	-2.07/2.12	-2.12/2.08	-2.12/2.06	-2.10/2.05	-2.12/2.11	-2.10/2.09	-2.12/2.09	-2.12/2.07	-2.33/2.33
500	0.9	-2.08/2.08	-2.09/2.06	-2.09/2.13	-2.06/2.09	-2.02/2.11	-2.11/2.15	-2.14/2.12	-2.12/2.15	-2.30/2.33
1000	0.1	-2.20/2.16	-2.09/2.09	-2.08/2.06	-2.1/2.08	-2.06/2.10	-2.12/2.07	-2.08/2.14	-2.02/2.06	-2.17/2.22
1000	0.2	-2.12/2.10	-2.12/2.13	-2.10/2.08	-2.05/2.03	-2.1/2.08	-2.15/2.09	-2.09/2.11	-2.10/2.12	-2.15/2.21
1000	0.3	-2.10/2.09	-2.11/2.11	-2.09/2.05	-2.10/2.10	-2.09/2.07	-2.11/2.09	-2.10/2.13	-2.02/2.09	-2.15/2.18
1000	0.4	-2.13/2.12	-2.06/2.16	-2.09/2.09	-2.10/2.10	-2.07/2.08	-2.06/2.10	-2.07/2.09	-2.12/2.09	-2.13/2.15
1000	0.5	-2.06/2.09	-2.13/2.10	-2.10/2.05	-2.06/2.07	-2.14/2.10	-2.10/2.08	-2.11/2.06	-2.12/2.10	-2.11/2.19
1000	0.6	-2.13/2.07	-2.04/2.09	-2.09/2.13	-2.05/2.12	-2.09/2.06	-2.12/2.08	-2.13/2.10	-2.12/2.05	-2.12/2.19
1000	0.7	-2.07/2.13	-2.13/2.05	-2.09/2.06	-2.07/2.06	-2.14/2.10	-2.13/2.07	-2.05/2.06	-2.06/2.14	-2.17/2.23
1000	0.8	-2.12/2.10	-2.15/2.13	-2.08/2.09	-2.09/2.05	-2.08/2.11	-2.08/2.05	-2.05/2.09	-2.09/2.08	-2.17/2.17
1000	0.9	-2.08/2.09	-2.11/2.07	-2.10/2.08	-2.09/2.08	-2.13/2.11	-2.11/2.08	-2.08/2.07	-2.12/2.08	-2.18/2.21
<i>Panel B: Cut-offs for spurious regression bias and data mining</i>										
250	0.1	-2.81/2.75	-2.78/2.72	-2.77/2.72	-2.92/2.87	-2.87/2.79	-2.87/2.76	-2.81/2.81	-2.90/2.88	-4.24/4.39
250	0.2	-2.87/2.84	-2.75/2.86	-2.08/2.78	-2.84/2.92	-2.85/2.84	-2.85/2.87	-2.92/2.86	-2.91/2.89	-4.21/4.29
250	0.3	-2.88/2.85	-2.74/2.83	-2.92/2.76	-2.82/2.88	-2.85/2.81	-2.93/2.77	-2.97/2.85	-2.93/2.95	-4.31/4.60
250	0.4	-2.82/2.80	-2.76/2.82	-2.82/2.77	-2.93/2.76	-2.86/2.76	-2.98/2.75	-2.94/2.88	-2.94/2.98	-4.13/4.42
250	0.5	-2.78/2.79	-2.87/2.85	-2.86/2.77	-2.81/2.75	-2.82/2.81	-2.92/2.79	-2.96/2.87	-2.90/2.85	-4.49/4.31
250	0.6	-2.82/2.76	-2.86/2.81	-2.91/2.83	-2.91/2.85	-2.79/2.84	-2.86/2.68	-2.89/2.81	-2.89/2.91	-4.23/4.03
250	0.7	-2.71/2.87	-2.79/2.86	-2.82/2.77	-2.95/2.86	-2.92/2.71	-2.98/2.87	-2.91/2.97	-3.01/2.89	-3.89/4.29
250	0.8	-2.81/2.83	-2.77/2.81	-2.71/2.76	-2.83/2.83	-2.93/2.90	-2.95/2.87	-3.02/2.89	-3.00/2.85	-4.42/4.15
250	0.9	-2.82/2.83	-2.82/2.90	-2.76/2.84	-2.84/2.89	-2.75/2.79	-2.89/2.85	-2.92/2.99	-2.92/3.06	-4.06/4.12
500	0.1	-2.81/2.78	-2.81/2.76	-2.82/2.89	-2.90/2.86	-2.82/2.76	-2.82/2.87	-2.73/2.85	-2.90/2.84	-3.17/3.22
500	0.2	-2.91/2.74	-2.80/2.73	-2.83/2.84	-2.72/2.94	-2.76/2.93	-2.80/2.85	-2.72/2.79	-3.06/2.76	-3.31/3.20
500	0.3	-2.85/2.82	-2.72/2.76	-2.82/2.87	-2.74/2.71	-2.79/2.78	-2.87/2.77	-2.77/2.80	-2.84/2.83	-3.44/3.39
500	0.4	-2.91/2.69	-2.78/2.77	-2.77/2.77	-2.89/2.83	-2.69/2.76	-2.85/2.82	-2.66/2.87	-2.85/2.83	-3.38/3.30
500	0.5	-2.86/2.79	-2.76/2.91	-2.68/2.74	-2.77/2.87	-2.76/2.88	-2.75/2.79	-2.94/2.86	-2.90/2.80	-3.34/3.13
500	0.6	-2.70/2.90	-2.88/2.86	-2.88/2.86	-2.77/2.83	-2.76/2.75	-2.81/2.79	-2.81/2.86	-2.86/2.89	-3.39/3.19
500	0.7	-2.88/2.85	-2.85/2.79	-2.81/2.92	-2.79/2.80	-2.79/2.89	-2.79/2.88	-2.85/2.94	-2.83/2.86	-3.25/3.21
500	0.8	-2.74/2.75	-2.87/2.85	-2.72/2.81	-2.79/2.80	-2.85/2.80	-2.79/2.73	-2.80/2.86	-2.87/2.76	-3.28/3.33
500	0.9	-2.88/2.90	-2.81/2.75	-2.81/2.89	-2.92/2.81	-2.72/2.84	-2.90/2.86	-2.83/2.78	-2.90/2.86	-3.37/3.36
1000	0.1	-2.88/2.91	-2.81/2.80	-2.74/2.79	-2.80/2.78	-2.80/2.82	-2.96/2.69	-2.78/2.87	-2.70/2.74	-2.97/2.90
1000	0.2	-2.84/2.86	-2.90/2.92	-2.80/2.80	-2.81/2.77	-2.88/2.82	-2.82/2.77	-2.75/2.85	-2.75/2.79	-3.05/2.98
1000	0.3	-2.74/2.74	-2.85/2.80	-2.87/2.82	-2.77/2.98	-2.78/2.73	-2.84/2.91	-2.79/2.85	-2.75/2.69	-2.92/2.99
1000	0.4	-2.92/2.80	-2.71/2.88	-2.74/2.82	-2.76/2.81	-2.81/2.79	-2.86/2.77	-2.81/2.96	-2.85/2.76	-2.9/2.94
1000	0.5	-2.79/2.86	-2.90/2.81	-2.77/2.68	-2.76/2.89	-2.81/2.72	-2.91/2.79	-2.74/2.70	-2.83/2.94	-2.97/3.00
1000	0.6	-2.73/2.92	-2.72/2.73	-2.84/2.87	-2.74/2.79	-2.91/2.66	-2.82/2.87	-2.90/2.78	-2.84/2.83	-2.85/2.96
1000	0.7	-2.79/2.89	-2.83/2.84	-2.73/2.78	-2.82/2.70	-2.84/2.83	-2.85/2.76	-2.84/2.79	-2.81/2.93	-2.96/2.92
1000	0.8	-2.81/2.83	-2.88/2.90	-2.82/2.80	-2.81/2.84	-2.82/2.84	-2.76/2.68	-2.85/2.91	-2.85/2.88	-2.92/2.91
1000	0.9	-2.77/2.79	-2.89/2.82	-2.82/2.71	-2.82/2.77	-2.77/2.86	-2.95/2.86	-2.84/2.75	-2.90/2.79	-2.99/3.04

Using the Bartlett method, this table reports dummy variable coefficient critical cut-off t -statistic values for dummy variable regression model, $r_t = \alpha + \beta_0 r_{t-1} + \beta_1 \pi_t + u_t$ under different combinations of autocorrelation of the dependent variable and transition probability of the dummy variable. Autocorrelation of the dummy variable is equal to two times the dummy variable transition probability minus one ($2q - 1$). Lower and upper bounds (i.e., critical cut-off values) of the 95% confidence interval of t -statistics adjusted for spurious regression bias are reported. The monthly mean return of the simulated dependent variable series is zero and the monthly standard deviation is one. For the Bartlett method, the fixed lag length HAC standard error estimation sets the lag length (M) as a fixed function of the sample size T ($M = bT$), where b is a constant and is set at 0.05.

they can be greatly ameliorated if regression model choices and the choice of regression variables are theory-driven.

6. Conclusion

A growing body of empirical literature indicates that long-lasting regimes such as political cycles, business cycles, and/or seasons

of the year produce significant market return or return volatility differentials. This paper argues that the lack of variability in an indicator variable representing a long-lasting regime in a time-series regression makes it hard for a test to distinguish the null from the alternative. While the severity of the problem becomes small as sample size grows large, the sample sizes in typical political-, seasonal-, and/or business-cycle studies are small enough that

the use of standard Newey and West (1987) significance tests is inappropriate. To remedy the problem, the paper provides a simulation procedure that allows researchers to properly test for regime differences in small samples. Among other things, the simulation procedure shows that the adjusted R^2 s and the coefficient estimates obtained in a number of financial time-series tests of political, seasonal, and business cycle return or return volatility differences are less than would be expected by chance. In other words, many documented regime market return or return volatility differentials are, in fact, illusory.

Based on the results of this study, researchers are encouraged to examine and report the properties of the dichotomous random variables, together with the continuous random variables, before performing hypothesis tests within a time-series regression framework. This study shows, for example, that the level of autocorrelation in the dichotomous variable contains key information about the level of persistence. Had these levels been reported in past work, the researcher and the reader would have been alerted to the fact that alternative testing procedures were warranted.

Acknowledgements

The authors are grateful to an anonymous referee, Cliff Ball, Nick Bollen, Michael Bradley, Wayne Ferson, Ben Jacobsen, Michael Munger and seminar participants at the University of Queensland, the University of Technology Sydney, Vanderbilt University and Canterbury University for helpful comments and suggestions.

References

- Andrews, D.W.K., 1991. Heteroskedasticity and autocorrelation consistent covariance matrix estimation. *Econometrica* 59, 817–858.
- Basintha, A., Kurov, A., 2008. Macroeconomic cycles and the stock market's reaction to monetary policy. *Journal of Banking and Finance* 32, 2606–2616.
- Bouman, S., Jacobsen, B., 2002. The Halloween indicator, "sell in May and go away": Another puzzle. *American Economic Review* 92, 1618–1635.
- Browning, E.S., 2005. For Wall Street, cruelest month is September. *Wall Street Journal Online*, August 15.
- Candelon, B., Piplack, J., Straetmans, S., 2008. On measuring synchronization of bulls and bears: The case of East Asia. *Journal of Banking and Finance* 32, 1022–1035.
- Chen, S., 2009. Predicting the bear market: Macroeconomic variables as leading indicators. *Journal of Banking and Finance* 33, 211–223.
- Chordia, T., Shivakumar, L., 2002. Momentum, business cycle, and time-varying expected returns. *Journal of Finance* 57, 985–1019.
- Cooper, M.J., Gutierrez Jr., R.C., Hameed, A., 2004. Market states and momentum. *Journal of Finance* 59, 1345–1365.
- Deng, A., 2005. Understanding Spurious Regression in Financial Economics. Working Paper, Boston University.
- Enders, W., 1995. *Applied Econometric Time-series*. Wiley, New York.
- Ferson, W.E., Sarkissian, S., Simin, T.T., 2003. Spurious regressions in financial economics? *Journal of Finance* 58, 1393–1413.
- Foster, F.D., Smith, T., Whaley, R.E., 1997. Assessing the goodness-of-fit of asset pricing models: The distribution of the maximal R -squared. *Journal of Finance* 52, 591–607.
- Granger, C.W.J., Hyung, N., Jeon, Y., 2001. Spurious regressions with stationary series. *Applied Economics* 33, 899–904.
- Granger, C.W.J., Newbold, P., 1974. Spurious regressions in economics. *Journal of Econometrics* 4, 111–120.
- Greene, W.H., 2003. *Econometric Analysis*. Prentice Hall, New Jersey.
- Hamilton, J.D., 1994. *Time-series Analysis*. Princeton University Press, Princeton.
- Harding, D., Pagan, A., 2006. The Econometric Analysis of Constructed Binary Time-series. Working Paper, Australian National University.
- Jacobsen, B., Marquering, W., 2008. Is it the weather? *Journal of Banking and Finance* 32, 526–540.
- Jensen, G.R., Mercer, J.M., Johnson, R.R., 1996. Business conditions, monetary policy, and expected security returns. *Journal of Financial Economics* 40, 213–237.
- Kanas, A., 2008. On real interest rate dynamics and regime switching. *Journal of Banking and Finance* 32, 2089–2098.
- Kiefer, N.M., Vogelsang, T.J., 2005. A new asymptotic theory for heteroskedasticity-autocorrelation robust tests. *Econometric Theory* 21, 1130–1164.
- Leblang, D., Mukherjee, B., 2005. Government partisanship, elections, and the stock market: Examining American and British stock returns, 1930–2000. *American Journal of Political Science* 49, 780–802.
- Lo, A.W., MacKinlay, C.A., 1990a. An econometric analysis of infrequent trading. *Journal of Econometrics* 45, 181–211.
- Lo, A.W., MacKinlay, C.A., 1990b. Data snooping in tests of financial asset pricing models. *Review of Financial Studies* 1, 41–66.
- Nabeya, S., Perron, P., 1994. Local asymptotic distribution related to the AR(1) model with dependent errors. *Journal of Econometrics* 62, 229–264.
- Newey, W.K., West, K.D., 1987. A simple, positive semi-definite, heteroskedasticity and autocorrelation consistent covariance matrix. *Econometrica* 55, 7003–7708.
- Pesaran, M.H., Timmerman, A.G., 2006. Testing Dependence Using Serially Correlated Multi-category Variables. Working Paper, University of Cambridge.
- Powell, J.G., Shi, J., Smith, T., Whaley, R.E., 2007. The persistent presidential dummy. *Journal of Portfolio Management* 33, 133–143.
- Resnick, B.G., Shoesmith, G.L., 2002. Using the yield curve to time the stock market. *Financial Analysts Journal* 58, 82–90.
- Richardson, M., Stock, J.H., 1989. Drawing inferences from statistics based on multiyear asset returns. *Journal of Financial Economics* 25, 323–348.
- Roll, R., 1984. Orange juice and weather. *American Economic Review* 74, 861–880.
- Santa-Clara, P., Valkanov, R., 2003. The presidential puzzle: Political cycles and the stock market. *Journal of Finance* 58, 1841–1872.
- Stoll, H.R., Whaley, R.E., 1990. The dynamics of stock index and stock index futures returns. *Journal of Financial and Quantitative Analysis* 25, 441–468.
- Sullivan, R., Timmermann, A., White, H., 1999. Data-snooping, technical trading rule performance, and the bootstrap. *Journal of Finance*, 1647–1691.
- Sullivan, R., Timmermann, A., White, H., 2001. Dangers of data-mining: The case of calendar effects in stock returns. *Journal of Econometrics*, 249–286.
- Yan, X., 2005. Are Momentum Profits Due to Common Factors or Idiosyncratic Returns? Evidence from the Volatility of Momentum Returns. Working Paper, University of Missouri-Columbia.
- Yule, G.U., 1926. Why do we sometimes get nonsense correlations between time-series? A study in sampling and the nature of time-series. *Journal of the Royal Statistical Society* 89, 1–64.