Testing Cross-Section Correlation in Panel Data Using Spacings

Serena NG
Department of Economics, University of Michigan, Ann Arbor, MI 48109 (*Serena.Ng@umich.edu*)

This article provides tools for characterizing the extent of cross-section correlation in panel data when we do not know a priori how many and which series are correlated. Our tests are based on the probability integral transformation of the ordered correlations. We first split the transformed correlations by their size into two groups, then evaluate the variance ratio of the two subsamples. The problem of testing cross-section correlation thus becomes one of identifying mean shifts and testing nonstationarity. The tests can be applied to raw data and regression errors. We analyze data on industrial production among 12 OECD countries, as well as 21 real exchange rates. The evidence favors a common factor structure in European real exchange rates but not in industrial production.

KEY WORDS: Group effects; Purchasing power parity; Unit root; Variance ratio.

1. INTRODUCTION

Existing tests of cross-section correlation are concerned primarily with the null hypothesis that all units are uncorrelated against the alternative that the correlation is nonzero for some unit. Other statistics test group correlation in an error component framework and thus maintain identical correlation within group as the null hypothesis, assuming that group membership is known. But perfect and zero correlation are extreme hypotheses, and rejections by conventional tests do not always reveal much information about the extent of the correlation. It is often useful for estimation, inference, and economic interpretation to know whether a rejection is due to, say, 10% or 80% of the correlations being nonzero.

The present analysis is concerned with the situation when possibly some, but not necessarily all, of the units are correlated. The correlation is not sufficiently prevalent to be judged common, but is sufficiently extensive so that testing the no correlation hypothesis will almost always lead to a rejection. Our objective is to characterize the correlation without imposing a structure. This means determining the number of correlated pairs, determining whether the correlations are homogeneous, and evaluating the magnitude of the correlations.

We develop tools to assess the extent of cross-section correlation in a panel of data with N cross-section units and T times series observations. Our analysis is based on the $n = N \cdot (N-1)/2$ unique elements above the diagonal of the sample correlation coefficient matrix, ordered from the smallest to the largest. We do not directly test whether the sample correlations (jointly or individually) are zero. Instead, we test whether the probability integral transformation of the ordered correlations, denoted by $\bar{\phi}_j$, are uniformly distributed. If the underlying correlations are 0, then the "uniform spacings," defined as $\bar{\phi}_j - \bar{\phi}_{j-1}$, is a stochastic process with well-defined properties, and it is these properties that we test.

Exploiting the duality between uniformity and no correlation in hypothesis testing is not new. Durbin (1961) considered using uniformity as a test for serial correlation in independent normally distributed data. The idea is that if the periodogram is evenly distributed across frequencies, then a suitably scaled periodogram is uniform on $[0, 1]$. Here we use uniformity to test cross-section correlation.

We partition the spacings into two groups, labeled S (small) and L (large), with $\hat{\theta} \in [0, 1]$ being the estimated fraction of the sample in S, and which we estimate using a breakpoint analysis. For each group, we test whether the variance of $\bar{\phi}_j - \bar{\phi}_{j-q}$ is linear in q. Essentially, the problem of testing cross-section correlation is turned into a problem of testing uniformity and nonstationarity. If we reject the no correlation hypothesis in the L sample but not in the S sample, then we can say that a fraction $\hat{\theta}$ of the correlation coefficients are not statistically different from 0. This is unlike existing tests when a rejection often reveals little about the extent of the correlation. Our procedures are valid when applied to the full or a subset of the correlations, a property that conventional tests usually do not have. Because the identity of the series generating small and large correlations can always be recovered, knowing which correlation pairs belong to S can sometimes reveal useful information for economic analysis.

The treatment of cross-correlation in the errors has important implications for estimation and inference. In recent work, Andrews (2003) showed that ordinary least squares, when applied to cross-section data, can be inconsistent unless the errors conditional on the common shock are uncorrelated with the regressors. The usual t test will no longer be asymptotically normal when there is extreme correlation, such as that induced by a common shock. Knowledge of the pervasiveness and size of the residual correlations is thus important.

Omitting cross-section correlation is known to create problems for inference. Richardson and Smith (1993) noted that evidence for cross-sectional kurtosis could be the result of omitted cross-section correlation in stock returns. The panel unit root tests developed by Levin, Lin, and Chu (2002) and others are based on the assumption that the units in the panel are uncorrelated. In a large T, small N setup, O'Connell (1998) found that the much-emphasized power gain of panel over univariate unit root tests could be a consequence of omitted cross-section correlation, causing the panel tests to be oversized. The tests proposed by Moon and Perron (2004) and Bai and Ng (2004) are valid when the correlation is driven by a pervasive source.

Journal of Business & Economic Statistics
January 2006, Vol. 24, No. 1
DOI 10.1198/073500105000000171

Which test is appropriate depends on how many units are correlated.

Recent years have seen much interest in using approximate factor models in econometric modeling. These differ from classical factor models by allowing weak cross-section correlation in the idiosyncratic errors, where "weak" is defined by a bound on the column sum of the $N \times N$ covariance matrix of the idiosyncratic errors (see Stock and Watson 2002; Bai and Ng 2002). However, this definition of weak cross-section correlation, although useful for the development of asymptotic theory, is not useful in guiding practitioners as to how much residual correlation is in the data. Toward this end, this article provides agnostic tools for identifying and characterizing correlation groups. The procedures can be applied to test correlations for which $\sqrt{T}$-consistent estimates are available. The article proceeds with a review of tests for cross-section correlation in Section 2. The procedures are developed in Sections 3 and 4, and simulations are reported in Section 5.

2. RELATED LITERATURE

Suppose that we have N cross-section units each with T times series observations. Denote the $T \times N$ data matrix by $\mathbf{z}$. This could be raw data or, in regression analysis, $\mathbf{z}$ would be the regression errors. We are interested in characterizing the cross-correlation structure of z_{it}. Write

$$z_{it} = \delta_i G_t + e_{it}, \tag{1}$$

where $E(e_{it}) = 0$ and $E(e_{it}e_{jt}) = \omega_{ij}$. The $N \times N$ population variance–covariance matrix of $\mathbf{z}$ with $\text{var}(G_t)$ normalized to unity is

$$\mathbf{\Sigma_z} = \boldsymbol{\delta}\boldsymbol{\delta}' + \mathbf{\Omega},$$

where $\mathbf{\Omega}$ (whose i,j element is ω_{ij}) will generally be nonspherical because of cross-section correlation and/or heteroscedasticity in e_{it}. If $\delta_i = 0$ for all i and $\omega_{ij} = 0$ if $i \neq j$, then $\mathbf{\Sigma_z}$ is a diagonal matrix, and the data are cross-sectionally uncorrelated. Strong-form cross-section correlation occurs when $|\delta_i| \neq 0$ for almost every i and the largest eigenvalue of $\mathbf{\Omega}$ is bounded so that all series are related through the common factor, G_t. The error-component structure occurs as a special case when $\delta_i = \bar{\delta} \neq 0$ for every i and $\mathbf{\Omega} = \omega^2\mathbf{I}_n$. The presence of a common factor does not imply that all bivariate correlations will be identical, because this will depend on the factor loading δ_i, as well as on the importance of the idiosyncratic error, e_{ti}. We are especially interested in better understanding the intermediate cases when the two extremes of zero and strong correlation are inappropriate characterizations of the data.

Let $c_{ij} = \text{cov}(z_i, z_j)/\sqrt{\text{var}(z_i)\,\text{var}(z_j)}$ be the population correlation coefficient between two random variables, z_i and z_j. Given observations $\{z_{it}\}$ and $\{z_{jt}\}$, $t = 1, \ldots, T$, the Pearson correlation coefficient is

$$\hat{c}_{ij} = \frac{s_{ij}}{\sqrt{s_{ii}s_{jj}}},$$

with $s_{ij} = T^{-1}\sum_{t=1}^{T}(z_{it} - \bar{z}_i)(z_{jt} - \bar{z}_j)$. If the data are normally distributed, then $\hat{c}_{it}\sqrt{(T-1)}/\sqrt{1-\hat{c}_{ij}^2}$ has a t distribution with $T-2$ degrees of freedom under the null hypothesis of no correlation, and when T is large,

$$\sqrt{T}(\hat{c}_{ij} - c_{ij}) \approx \text{N}\big(0, (1-c_{ij}^2)^2\big).$$

Exact tests for equality of a set of correlation coefficients are available for normally distributed data because under normality, expressions for the covariance of correlation coefficients are known. Otherwise, the asymptotic covariance of correlation coefficients can be approximated by the delta method and shown to be a function of the true underlying correlations (see Olkin and Finn 1990; Steiger 1980; Meng, Rosenthal, and Rubin 1992). Under normality and assuming that N is fixed, Breusch and Pagan (1980) showed that a test for the hypothesis that all correlation coefficients are jointly 0 is

$$\text{LM} = \binom{N}{2}^{-1}\sum_{i=2}^{N}\sum_{j=1}^{i-1}T\hat{c}_{ij}^2 \to \chi_n^2, \tag{2}$$

with $n = N(N-1)/2$. When N is large, the normalized test $\frac{\text{LM}-n}{\sqrt{2n}}$ is asymptotically N(0, 1) as $T \to \infty$ and then $N \to \infty$. A statistic that pools the p values is also asymptotically valid.

The LM test is asymptotically nuisance parameter-free when T is large, but it depends on the properties of the parent distribution even under the null hypothesis when T is fixed. Using rank correlations, Frees (1995) considered a nonparametric test of no cross-section correlation in panel data when N is large relative to T. A simulation based test was considered by Dufour and Khalaf (2002) for the null hypothesis that the error matrix from estimating a system of seemingly unrelated regressions with normally distributed errors is diagonal. Like the test of Frees and the LM test, the null hypothesis is $c_{ij} = 0$ for all (i,j) pairs, and the test rejects when $c_{ij} \neq 0$ for some (i,j) pair. At the other end of the spectrum, Moulton (1990) developed a test to detect group effects in panel data assuming that group membership is known (see also Moulton 1987). The maintained assumption is identical correlation within the group.

All of the aforementioned tests amount to testing zero or strong correlation. Only in rare cases can we test a correlation pattern other than the two extremes. The need to impose a rigid correlation pattern (i.e., $\mathbf{\Sigma_z}$ is diagonal or has a factor structure) in estimation and inference arises from that fact that unlike time series and spatial data, cross-section data have no natural ordering. Absent extraneous information, the correlation with the neighboring observation has no meaningful interpretation. One cannot impose some type of mixing condition to justify testing the correlation among just a few neighboring observations.

From a practical perspective, the shortcoming of these tests is that rejecting the null hypothesis does not reveal much information about the strength and prevalence of the cross-section correlation. Whether the rejection is due to 10% or 80% of the correlations being nonzero makes a difference not just in the way we handle the correlation econometrically, but also in the way we view the economic reason for the correlation.

3. THE ECONOMETRIC FRAMEWORK

We start with the premise that the panel of data are neither all uncorrelated nor all correlated with each other. As a matter of notation, we let $[x]$ denote the integer part of x. For a

series $x_j, j = 1, \ldots, n$, let $x_{[1:n]}, x_{[2:n]}, \ldots, x_{[n:n]}$ denote the ordered series, that is, $x_{[1:n]} = \min_j(x_j)$ and $x_{[n:n]} = \max_j(x_j)$. We assume for now that the data are normally distributed and discuss the case of nonnormal data in Section 5. Because the analysis applies to other definitions of cross-section correlation coefficients, we generically denote the vector of *unique* sample correlation coefficients by $\hat{\mathbf{p}}$ and the corresponding population correlation coefficients by $\mathbf{p}$. We also let $\bar{\mathbf{p}} = |\hat{\mathbf{p}}|$, the vector of absolute sample correlation coefficients. Thus if Pearson correlation is used for the analysis, then we have

$$\hat{\mathbf{p}} = (\hat{p}_1, \hat{p}_2, \ldots, \hat{p}_n) = \text{vech}(\hat{\mathbf{c}})$$

and

$$\bar{\mathbf{p}} = (|\hat{p}_1|, |\hat{p}_2|, \ldots, |\hat{p}_n|), \qquad n = N(N-1)/2.$$

We are interested in the severity of the correlation among the units. Our general strategy is to split the sample into a group of "small" and a group of "large" correlations, then test whether the small correlations are 0. The two steps allow us to understand how pervasive and how strong the cross-section correlation is.

The following lemma forms the basis of our methodology.

Lemma 1. Let $(u_1, u_2, \ldots, u_n)'$ be an $n \times 1$ vector of iid U[0, 1] variates and let $u_{[1:n]}, \ldots, u_{[n:n]}$ denote the ordered data. Let $D_1 = u_{[1:n]}$, $D_j = u_{[j:n]} - u_{[j-1:n]}, j = 2, \ldots, n$, and $D_{n+1} = 1 - u_{[n:n]}$ be the spacings. Then $E(D_j) = \frac{1}{n+1}$, $\text{var}(D_j) = n \times (n+1)^{-2}(n+2)^{-1}\ \forall j$, and $\text{cov}(D_i, D_j) = \frac{-1}{(n+1)^2(n+2)}\ \forall i \neq j$.

If u_j is uniformly distributed on the unit interval, then a plot of j against $u_{[j:n]}$ should be a straight line with slope $1/(n+1)$. The variable $D_j = u_{[j:n]} - u_{[j-1:n]}$, known as the uniform "spacings," has the property that $\sum_{j=1}^{n+1} D_j = 1$. This summing-up constraint also implies that $D_1, \ldots, D_n$ contains all of the information about the $n+1$ spacings. Because $E(D_j)$ is the same for all j, $E(D_j) = 1/(n+1)$. Furthermore, the variance for all units and the covariance between any two units are the same. (See Pyke 1965 for an excellent review on spacings.) We make use of the fact that if $u_j \sim \text{U}[.5, 1]$, then $2u_j - 1$ is also uniformly distributed on the interval [0, 1], and spacings for $2u_j - 1$ have the properties stated in Lemma 1. Because the spacings for $u_j \sim \text{U}[.5, 1]$ are half the spacings for $2u_j - 1$, Lemma 1 implies that if $u_j \sim \text{U}[.5, 1]$, then the corresponding spacings have mean $\frac{1}{2}\frac{1}{n+1}$ and variance $\frac{1}{4}n(n+1)^{-2}(n+2)^{-1}$.

3.1 Partitioning the Correlations Into Two Sets

Under $H_0 : p_j = 0$, $\sqrt{T}\hat{p}_j \sim \text{N}(0, 1)$, we have

$$\sqrt{T} \cdot \bar{p}_j = |\sqrt{T} \cdot \hat{p}_j| \sim \chi_1;$$

that is, $\bar{p}_j$ is asymptotically distributed as chi (not chi-squared) with 1 degree of freedom, or, equivalently, $\sqrt{T}\bar{p}_j$ is a half-normally distributed random variable with mean .7979 and variance .3634. Sort $\bar{\mathbf{p}}$ from the smallest to the largest, and denote the ordered series by $(\bar{p}_{[1:n]}, \ldots, \bar{p}_{[n:n]})'$. Taking absolute values ensures that large negative correlations are treated symmetrically as large positive correlations. Define $\bar{\phi}_j$ as $\Phi(\sqrt{T}\bar{p}_{[j:n]})$, where Φ is the cumulative distribution function of the standard normal distribution. We have

$$\begin{aligned}\bar{\boldsymbol{\phi}} &= (\bar{\phi}_1, \bar{\phi}_2, \ldots, \bar{\phi}_n)' \\ &= \big(\Phi\big(\sqrt{T}\bar{p}_{[1:n]}\big), \Phi\big(\sqrt{T}\bar{p}_{[2:n]}\big), \ldots, \Phi\big(\sqrt{T}\bar{p}_{[n:n]}\big)\big)'.\end{aligned}$$

Because $\bar{p}_{[j:n]}$ is ordered, $\bar{\phi}_j$ is also ordered with $\bar{\phi}_{j-1} < \bar{\phi}_j$. Because $\bar{p}_{[j:n]} \in [0, 1]$, it follows that $\bar{\phi}_j \in [.5, 1]$. The null hypothesis of $p_j = 0$ is now restated as $\bar{\phi}_j \sim \text{U}[.5, 1]$. If all of the correlations are 0, then $\frac{1}{n}\sum_{j=1}^{n} \bar{\phi}_j$ should be close to .75, and in principle a t test can be constructed. But like the LM test, a rejection reveals little about the extent of the correlation.

Let $\Delta\bar{\phi}_j = \bar{\phi}_j - \bar{\phi}_{j-1}$ be the spacings and, by Lemma 1, let $E(\Delta\bar{\phi}_j) = \frac{1}{2(n+1)}$. Given $\bar{\phi}_j$, $j = 1, \ldots, n$, consider estimating $E(\Delta\bar{\phi}_j)$ on partitioning the sample at arbitrary $\tilde{\theta} \in (0, 1)$,

$$\bar{\Delta}_S(\tilde{\theta}) = \frac{1}{[\tilde{\theta}n]} \sum_{j=1}^{[\tilde{\theta}n]} \Delta\bar{\phi}_j$$

and

$$\bar{\Delta}_L(\tilde{\theta}) = \frac{1}{[n(1-\tilde{\theta})]} \sum_{j=[\tilde{\theta}n]+1}^{n} \Delta\bar{\phi}_j,$$

where $[n\tilde{\theta}]$ is the integer part of $n\tilde{\theta}$. If $p_j = 0\ \forall j$, then we should have $\bar{\Delta}_S(\tilde{\theta}) \approx \bar{\Delta}_L(\tilde{\theta}) \approx \frac{1}{2(n+1)}\ \forall \tilde{\theta}$.

Now suppose that only $m \leq n$ of the correlation coefficients are 0 or small. We would expect $\bar{\phi}_j$, $j = 1, \ldots, m$, to be strictly less than 1. In contrast, we would expect the $\bar{\phi}_j$, $j = m+1, \ldots, n$, to be close to 1 because if $p_{[j+1:n]}$ is not small, then $\sqrt{T}\bar{p}_{[j+1:n]}$ will diverge. In a q–q plot, we would expect $\bar{\phi}_j$ to be approximately linear in j until $j = m$, then rise steeply, and eventually flatten out at the boundary of 1. Let $\theta = m/n$. Then in terms of $\Delta\bar{\phi}_j$, we should have

$$\bar{\Delta}_S(\theta) \approx \frac{1}{2m} \neq \bar{\Delta}_L(\theta).$$

The difference between $\bar{\Delta}_S$ and $\bar{\Delta}_L$ is thus informative about m, which we can estimate by locating a mean shift in $\Delta\bar{\phi}_j$ or by a slope change in $\bar{\phi}_j$. Because we will be analyzing the properties of $\Delta\bar{\phi}_j$ in the two subsamples, we look for a mean shift in $\Delta\bar{\phi}_j$.

Define the total sum of squared residuals evaluated at $\tilde{m} = [\tilde{\theta}n]$ as

$$Q_n(\tilde{\theta}) = \sum_{j=1}^{[\tilde{\theta}n]} \big(\Delta\bar{\phi}_j - \bar{\Delta}_S(\tilde{\theta})\big)^2 + \sum_{j=[\tilde{\theta}n]+1}^{n} \big(\Delta\bar{\phi}_j - \bar{\Delta}_L(\tilde{\theta})\big)^2. \quad (3)$$

In our analysis, the series of ordered absolute values of the population correlation coefficients, $|p_{[j:n]}|$, exhibits a slope shift at m. But $\bar{p}_{[j:n]}$ is $\sqrt{T}$-consistent for $|p_{[j:n]}|$, and $\bar{\phi}_j$ is monotone-increasing in $\bar{p}_{[j:n]}$. The criterion function $Q_n(\tilde{\theta})$ converges uniformly to a function with a minimum where the true change-point in $|p_{[j:n]}|$ occurs. Thus the global minimizer of (3), that is,

$$\hat{\theta} = \underset{\tilde{\theta} \in [\underline{\theta}, \bar{\theta}]}{\arg\min}\, Q_n(\tilde{\theta}),$$

consistently estimates the break fraction θ, and the convergence rate is n (see Bai 1997). Partitioning the sample at $\hat{m} = [\hat{\theta}n]$ is optimal in the sense of minimizing (3).

The foregoing analysis is designed for cases where a subset of the correlations are nonzero. But when all of the correlations

are large and close to unity, there will be no variation in $\bar{\phi}_j$, and hence we would not expect to detect a mean shift in $\Delta\bar{\phi}_j$. The foregoing breakpoint analysis, although informative, is incomplete. To learn more about the nature of the correlation, such as whether the correlations are homogeneous or heterogeneous, we need to go one step further. Before proceeding with such an analysis, we first illustrate the properties of the breakpoint estimator using some examples.

3.2 Illustration

We simulate data as follows. For each $i = 1, \ldots, N$, $t = 1, \ldots, T$,

$$z_{it} = \delta_i G_t + e_{it} \qquad \text{and}$$
$$\mathbf{e}_t = (e_{1t}, \ldots, e_{NT})' \sim \mathbf{\Omega}^{1/2}\mathrm{N}(\mathbf{0}_N, \mathbf{I}_N),$$

where $G_t \sim \mathrm{N}(0, 1)$ and $\mathbf{\Omega}^{1/2}$ is a $n \times n$ matrix. By varying the number of δ_i that are nonzero or the structure of $\mathbf{\Omega}^{1/2}$, we have 10 configurations representing different degrees of correlation in the data.

DGP 1 simulates a panel of uncorrelated data. DGPs 2 and 3 assume the presence of a common factor but with different assumptions about the importance of the common to idiosyncratic component. DGP 4 and 5 assume that $\mathbf{\Omega}^{1/2}$ is a Toeplitz matrix but the δ_i's are all 0. (A Toeplitz matrix is a band-symmetric matrix.) Under DGP 5, the diagonal elements of $\mathbf{\Omega}^{1/2}$ are 1, the elements above and below the diagonals are .8, and all other elements are 0. Under DGP 6, the diagonal elements of $\mathbf{\Omega}^{1/2}$ are 1, those above and below the diagonal elements are $-.5$, the elements that are two positions from the diagonal are .3, and all other elements are 0. DGPs 6 and 7 assume equal correlation within a group but with varying group size. DGPs 8–10 make different assumptions about the number of δ_i's that are nonzero. Because δ_i varies across i, there is heterogeneity in the magnitude of the correlation.

Define θ_0 as the fraction of zero entries above the diagonal of the $N \times N$ matrix $\mathbf{\Sigma_z}$. In general, θ_0 will be different than θ, the fraction of correlated pairs that are small. As a benchmark for how correlated the data are, Table 1 reports θ_0 for $N = 10, 20$, and 30. Notice that under DGP 5 and 6, θ_0 varies with N. This is because the number of nonzero entries in $\mathbf{\Omega}$ increases with the number of nonzero entries in $\mathbf{\Omega}^{1/2}$ in a nonlinear way. The "max"column reports $\max_i \frac{1}{N}\sum_{j=1}^{N} |\hat{c}_{ij}|$ in 1,000 replications. This statistic is an upper bound on the average correlation in the data because from matrix theory, the largest eigenvalue of $\mathbf{\Sigma_z}$ is bounded by $\max_j \sum_{i=1}^{n} |c_{ij}|$. The bound is often used as a condition on permissible cross-section correlation in approximate factor models (see, e.g., Bai and Ng 2002; Stock and Watson 2002). However, this statistic has shortcomings as a measure of severity of the correlations in the panel of data. For one thing, "max" is not monotonic in the number of nonzero correlations; for example, DGP 3 has more nonzero correlations than DGP 10, yet "max" is larger in DGP 10 than in DGP 3. For another, two similar values of "max" can be consistent with very different degrees of correlatedness; for example, DGPs 5 and 6 both have similar "max" values when $N = 20$, but θ_0 is much higher under DGP 6. The point to highlight is that when there is heterogeneity in the correlations, it will be difficult to characterize the severity of the correlation with a single statistic. This is unlike the situation for time series data, where the largest root of the autocorrelation matrix is informative about the extent of serial correlation.

Many of the features about $\bar{\phi}_j$ can be illustrated by looking at the q–q plot for a particular draw of the data. Figure 1(a) considers DGP1 (uncorrelated data) for $T = 200$ and $N = 20$ (or $n = 190$); Figure 1(b) considers $T = 400$ and $N = 30$ (or $n = 435$). Figure 1 depicts the line $.5 + .5j/(n + 1)$ on which all $\bar{\phi}_j$ should lie if they are exactly uniformly distributed. In finite samples, $\bar{\phi}_j$ should be approximately linear in j if they are transformed from normally distributed variables. Because the data are uncorrelated, the quantile function does not exhibit any abrupt change in slope, and the average of $\Delta\bar{\phi}_j$ is approximately $\frac{1}{2(n+1)}$.

Now consider the case of correlated data generated by DGPs 9 and 10. In both cases, about one-third of the correlations are 0, but the correlations are much smaller (and hence $\bar{\phi}_j$) under DGP 9 because of the larger idiosyncratic variance. Because many correlations are nonzero, Figure 2 shows that $\bar{\phi}_j$ no longer evolves around the straight line with slope $\frac{1}{2(n+1)}$. Instead, a subset of them vary around a straight line drawn with an x-axis that is appropriately truncated to the number of correlation coefficients that are 0 or small.

For the sample of data simulated using DGP 9, 82 of the $\bar{\phi}_j$'s are $<.95$. This means that if we used the asymptotic standard error under no correlation of $\frac{1}{\sqrt{T}}$ to test the 190 correlations one by one, then we would end up with a group S consisting of 82 correlations deemed insignificant at the 5% level. A different significance level will produce a different group size. Our breakpoint analysis does not depend on the choice of the significance level; rather, it lets the data determine the sharpest difference between the two groups.

Table 1. DGP

				$\theta_0(N)$			max(N)		
DGP	$\#\delta_i \neq 0$	δ_i	$\mathbf{\Omega}^{1/2}$	10	20	30	10	20	30
1	0	0	$\mathbf{I}_n$	1.0000	1.0000	1.0000	.1707	.1217	.1048
2	N	N(0, 1)	$.2\mathbf{I}_n$	0	0	0	.7084	.7110	.4104
3	N	N(0, 1)	$\mathbf{I}_n$	0	0	0	.4884	.4880	.4886
4	0	0	$T([1, .8, \mathbf{0}_{n-2}])$	.6222	.8053	.8690	.3509	.2183	.1723
5	0	0	$T([1, -.5, .3, \mathbf{0}_{n-3}])$	.3333	.6316	.7471	.4312	.2585	.1995
6	$.4N$	1	$\mathbf{I}_n$	.8667	.8526	.8483	.2977	.2775	.2709
7	$.8N$	1	$\mathbf{I}_n$	.3778	.3684	.3655	.4935	.4767	.4733
8	$.4N$	N(0, 1)	$\mathbf{I}_n$	.8667	.8526	.8483	.2444	.2293	.2278
9	$.8N$	N(0, 1)	$\mathbf{I}_n$	.3778	.3684	.3655	.4019	.4024	.3990
10	$.8N$	N(0, 1)	$.2\mathbf{I}_n$	.3778	.3684	.3655	.5778	.5795	.5821

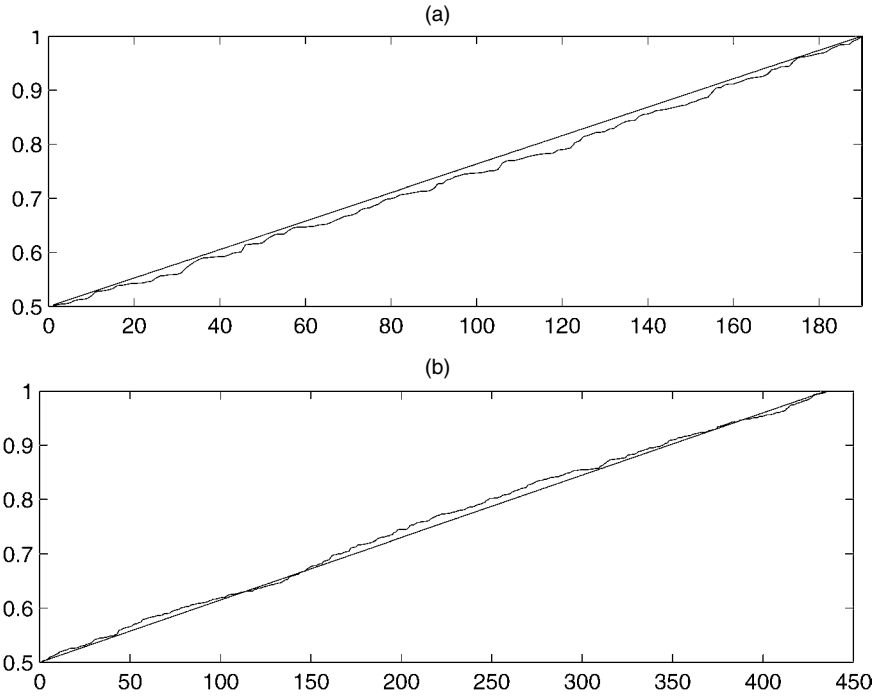

Figure 1. $\bar{\phi}(j)$: DGP 1. (a) $T = 200$, $N = 20$; (b) $T = 400$, $N = 30$.

Figure 3 presents $\bar{\phi}_j$ for DGP 2 (small idiosyncratic error) and 3 (large idiosyncratic error). Because DGP 2 has a stronger factor structure than DGP3, there are more $\bar{\phi}_j$'s at unity. In both cases, the minimum $\bar{\phi}_j$ is $>.5$, indicating even the small correlations are nonzero. The feature to highlight in Figure 3 is that p_j can be nonzero and yet $\bar{\phi}_j < 1$. We cannot always use the boundary of 1 to split the sample. Instead, we use the least squares criterion to separate the small and large correlations.

Figures 1, 2, and 3 show that the q–q plot indeed reveals much information about the extent of cross-correlation in the data. If all correlations are nonzero, then the q–q plot will be shifted upward with an intercept exceeding .5. If there is homogeneity in a subset of the correlations, then the q–q plot should be flat over a certain range, because there is no dispersion in the corresponding $\bar{\phi}_j$'s. The more prevalent and the stronger the correlation, the further away are $\bar{\phi}_j$ from the straight line with slope $\frac{1}{2(n+1)}$.

4. TESTING THE CORRELATION IN THE SUBSAMPLES

So far, we have used the breakpoint estimator to split the sample of n observations into two groups, one of size $\hat{m}$ and the other of size $n - \hat{m}$. It is possible that the $n - \hat{m}$ correlations

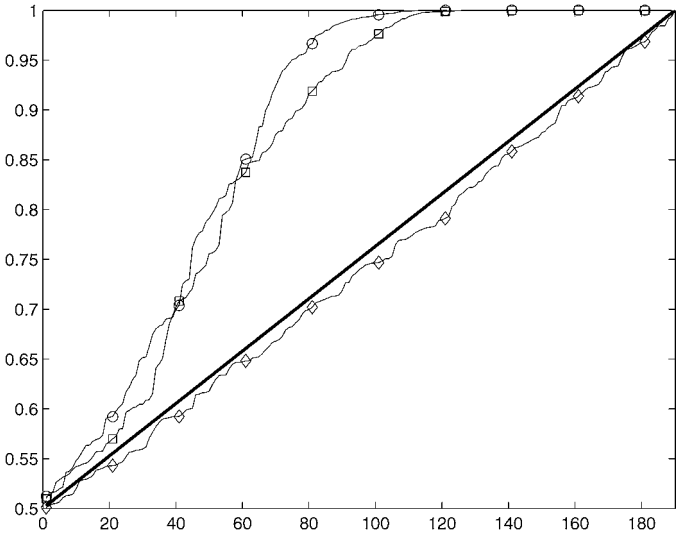

Figure 2. $\bar{\phi}(j)$: DGPs 1 (◇), 9 (○), and 10 (□). Here $z_{it} = \delta_i G_t + e_{it}$, $\delta_i \sim N(0, 1)$, $i = 1, \ldots, .8N = 152$, and $\delta_i = 0$, $i > .8N$. For DGP 9, $e_{it} \sim N(0, 1)$. For DGP 10, $e_{it} \sim N(0, .2)$.

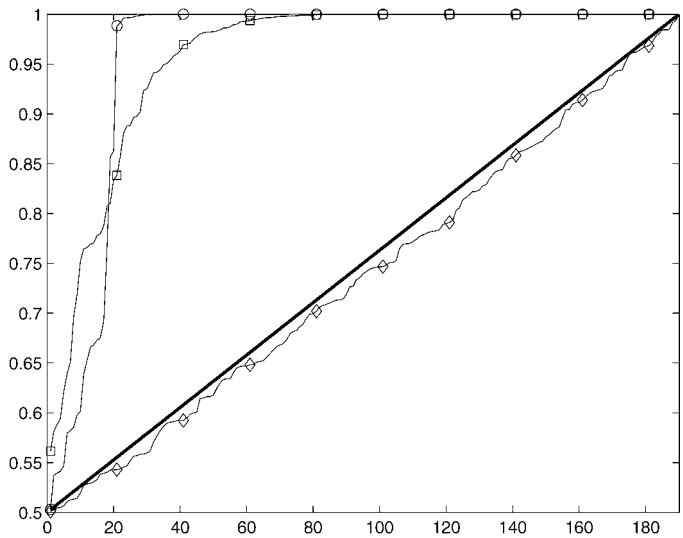

Figure 3. $\bar{\phi}(j)$: DGPs 1 (◇), 2 (○), and 3 (□). $\bar{\phi}_j$ can be < 1 even though DGPs 2 (small idiosyncratic errors) and 3 (large idiosyncratic errors) both have a factor structure.

in L are, in fact, not statistically different from the $\hat{m}$ correlations in S. Verifying this by testing the null hypothesis of no break is uninformative, because an error component structure would be consistent with the no break hypothesis as all correlations are identical, and yet incompatible with the no correlation assumption. In theory, it is also possible to formulate an order-statistic–oriented test based on the idea that all nonzero correlations, when multiplied by $\sqrt{T}$, are "large" with probability 1. But, as we have seen, some nonzero correlations may not be large enough to render $\bar{\phi}_j$ exactly 1. Applying an LM test to the subsamples is also problematic, because $\bar{p}_{[j:n]}, j = 1, \ldots, \hat{m}$, is a censored sample whenever $\hat{m} < n$. The subsample test will no longer be asymptotically chi-squared.

Here we consider a different approach that aims to test whether the subset of $\hat{m}$ correlations are 0. If the small correlations are found to be different than 0, then the correlations in L must also be different than 0.

Lemma 2. Let $u_j \sim \mathrm{U}[0, 1]$ with $D_j = u_{[j:n]} - u_{[j-1:n]}$ being uniform spacings. Then (a) $nD_j \sim \Gamma(1, 1)$ and (b) $\mathrm{corr}(nD_j, nD_k) = \frac{-1}{n}\ \forall j \neq k$. Let $D_j^q = u_{[j:n]} - u_{[j-q:n]}$ be q-order spacings. Then (c) $D_j^q \sim \mathrm{beta}(q, n-q+1)$ and (d) $nD_j^q \sim \Gamma(q, 1)$.

The simple spacings, multiplied by n, are distributed as $\Gamma(1, 1)$. Although the spacings are not independent, they are asymptotically uncorrelated. Furthermore, the structure of dependence is the same between D_j and D_k for any j and k. Spacings are thus "exchangeable." The quantity $u_{[j:n]} - u_{[j-q:n]}$ is referred to in the statistics literature as a q-order spacing. Properties of q-order spacings have been given by Holst (1979) and Arnold, Balakrishman, and Nagaraja (1992).

The features that motivate the test that follows are (a) and (d). Property (a) implies that if we define

$$\bar{\phi}_j^n = n \cdot \bar{\phi}_j,$$

and again using the fact that the spacings of variables that are uniformly distributed on [.5, 1] are half of the uniform spacings, then we can represent the scaled spacings process $\bar{\phi}_j^n$ as

$$\bar{\phi}_j^n - \bar{\phi}_{j-1}^n = .5\frac{n}{n+1} + \epsilon_j, \qquad \epsilon_j \sim (0, \sigma_\epsilon^2),$$

where $\sigma_\epsilon^2 = \frac{1}{4}\frac{n^3}{(n+1)^2(n+2)}$, $\frac{\mathrm{cov}(\epsilon_j, \epsilon_k)}{\mathrm{var}(\epsilon_j)} = \frac{-1}{n}$. Viewed in this light, $\bar{\phi}_j^n$ is a unit root process with a nonzero drift that tends to .5 when $N \to \infty$.

The observation that $\bar{\phi}_j^n$ is a differenced stationary process is reinforced by (d), which states that the mean and variance of q-order spacings are $\frac{q}{n+1}$ and $\frac{q(n+1-q)}{(n+1)^2(n+2)}$. But this implies that the mean and variance of q order spacings are q times the mean and variance of the scaled first-order spacings. A first-order integrated process also has this variance ratio property.

The feature that $\bar{\phi}_j^n$ is differenced stationary originates from the property that $\bar{\phi}_j$ is uniformly distributed under the no correlation assumption. Evidence against nonstationarity of $\bar{\phi}_j^n$ is thus evidence against no cross-section correlation. Thus we can use methods for analyzing integrated data to test cross-section correlation. Consider testing the no correlation hypothesis using a vector of $\bar{\phi}_k^n$ of dimension η; for example, $\eta = \hat{m}$ if we were to test the hypothesis of no correlation in the subsample S. Let

$$\hat{\mu}_1 = \frac{1}{\eta - 1}\sum_{k=1}^{\eta}(\bar{\phi}_k^n - \bar{\phi}_{k-1}^n),$$

$$\hat{\mu}_q = \frac{1}{\eta - q}\sum_{k=q+1}^{\eta}(\bar{\phi}_k^n - \bar{\phi}_{k-q}^n),$$

$$\hat{\sigma}_1^2 = \frac{1}{\eta}\sum_{k=1}^{\eta}(\bar{\phi}_k^n - \bar{\phi}_{k-1}^n - \hat{\mu}_1)^2,$$

and

$$\hat{\sigma}_q^2 = \frac{1}{q \cdot m_q}\sum_{k=q+1}^{\eta}(\bar{\phi}_k^n - \bar{\phi}_{k-q}^n - \hat{\mu}_q)^2,$$

where $m_q = (\eta - q)$. Alternatively, we can let $m_q = (\eta - q) \times (1 - q/\eta)$, which will yield an unbiased estimate of the variance. Define

$$SVR(\eta) = \frac{\hat{\sigma}_q^2}{\hat{\sigma}_1^2} - 1.$$

To gain more intuition, let $q = 2$ and consider

$$\frac{\hat{\sigma}_2^2}{\hat{\sigma}_1^2} - 1 \approx \frac{\mathrm{cov}(\Delta\bar{\phi}_j^n, \Delta\bar{\phi}_{j-1}^n) + (2\hat{\mu}_1 - \hat{\mu}_2)^2}{\mathrm{var}(\Delta\bar{\phi}_j^n)}.$$

The first ratio on the right side is the first-order autocorrelation coefficient for $\Delta\bar{\phi}_j^n$, whereas the second term in the numerator is the squared difference between the mean of q-order and simple spacings. But by (b) of Lemma 2, $\Delta\bar{\phi}_j^n$ is asymptotically uncorrelated, and by (a) of Lemma 2, the mean of q order spacings is linear in q. Both terms on the right side should be close to 0. A comparison of $\hat{\sigma}_q^2$ to $\hat{\sigma}_1^2$ thus allows us to test both properties of spacings that should hold if the underlying correlations are indeed 0.

Theorem 1. Consider the null hypothesis that a subset of $p_{[j:n]}$ of size η are jointly 0. Then, as $\eta \to \infty$, $\sqrt{\eta} \times SVR(\eta) \xrightarrow{d} \mathrm{N}(0, \omega_q^2)$, where $\omega_q^2 = \frac{2(2q-1)(q-1)}{3q}$.

Various authors have used functions of simple and $q > 1$ order spacings to test uniformity. Rao and Kuo (1984), among others, showed that using squares of the spacings in testing is optimal in terms of maximizing local power. Our spacings variance ratio (SVR) test is also based on the second moment of spacings. In the statistics literature, $\hat{\sigma}_1^2$ and $q \cdot \hat{\sigma}_q^2$ are referred to as simple and generalized Greenwood spacing tests for uniformity. The original Greenwood statistic was developed to test whether the spread of disease is random in time. Wells, Jammalamadaka, and Tiwari (1992) showed that spacing tests have the same asymptotic distribution when parameters of the distribution to be tested must be estimated as when the parameters are known. We combine the simple Greenwood test with a q-order test and apply our test to functions of estimated correlations, not the raw data. As shown in the Appendix, we can treat the $\bar{\phi}_j$'s as though they are uniformly distributed on [.5, 1], provided that $T \to \infty$.

The limiting distribution of the SVR test is obtained by noting that $\hat{\sigma}_1^2$ and $\hat{\sigma}_q^2$ are both consistent estimators for σ_ϵ^2,

but $\hat{\sigma}_1^2$ is efficient. Thus the Hausman principle applies, and $\text{var}(\hat{\sigma}_q^2 - \hat{\sigma}_1^2) = \text{var}(\hat{\sigma}_q^2) - \text{var}(\hat{\sigma}_1^2)$, from which $\text{var}(\hat{\sigma}_q^2/\hat{\sigma}_1^2 - 1)$ is obtained. In actual testing, we use the standardized SVR test, that is,

$$svr(\eta) = \frac{\sqrt{\eta}SVR(\eta)}{\sqrt{\omega_q^2}}, \tag{4}$$

which in large samples is asymptotically standard normal.

Some might observe that the SVR test has the same limiting distribution as the variance ratio (VR) test of Lo and MacKinlay (1988) for testing the null hypothesis of a random walk, that is, with $\epsilon_j \sim \text{N}(0, \sigma^2)$. The ϵ_j's considered here are nonnormal, however. Although the kurtosis of $\bar{\phi}_j$ appears in the variance of $\hat{\sigma}_1^2$ and $\hat{\sigma}_q^2$, the two kurtosis terms offset one another because we consider the variance of the difference. At an intuitive level, the reason that the SVR has the same distribution as the VR is that we are testing whether the variance of $\Delta^q \bar{\phi}_j^n$ is linear in q. This is a property of an integrated process, not just that of a random walk with Gaussian innovations.

The SVR test can be constructed in slightly differently ways. For example, one can follow the Greenwood statistic and set $\hat{\mu}_1$ to the population mean, which amounts to .5 in our setup, or one can set $\hat{\mu}_q$ to $q\hat{\mu}_1$ as in the VR test. Because $\hat{\mu}_1$ consistently estimates the mean of .5, and $\hat{\mu}_q$ and $q\hat{\mu}_1$ are both consistent for the mean of q-spacings under the null, the limiting distribution of the SVR test is the same for all implementations.

The SVR test can be applied to the full sample, to S, or to L, with $\eta = n$, $\eta = \hat{m}$, or $\eta = n - \hat{m}$. An important appeal of the SVR test is that it is based on spacings, and spacings are exchangeable. That is, we can test any subset of the spacings between adjacent order statistics. As seen in Figure 1 for uncorrelated data, the $\bar{\phi}_j$'s all lie along the straight line. This allows us to use any partition of the full sample to test the slope. Obviously, if we reject the uniformity hypothesis in S, then testing whether the hypothesis holds in L is not too interesting because, by construction, S is the sample with the smaller correlations. But in principle we can reapply the breakpoint estimator to the partition S to yield two subsamples within S, say SS and SL. We then can perform the SVR test to see whether the subsample SS consists of uncorrelated coefficients. In contrast, the LM test loses its asymptotic chi-squared property once the sample is censored. But there are two caveats to repeatedly applying the proposed procedures. First, if there are already too few observations in S, then the subsample SS may be too small to make the tests precise. Second, if the SVR is applied to the SS subsample after the S sample rejects uniformity, then the sequential nature of the test would need to be taken into account.

Because the identity of the two series underlying any given $\bar{p}_j$ can always be recovered, the procedure can be used to see whether those highly correlated pairs share common characteristics. It should be noted, however, that correlations are not transitive; that is, $\text{cov}(z_i, z_j) \neq 0$ and $\text{cov}(z_j, z_k) \neq 0$ do not necessarily imply $\text{cov}(z_i, z_k) \neq 0$. Our analysis provides a classification of correlation pairs into groups. Additional economic analysis is needed to determine whether the series underlying the correlation pairs can be grouped.

5. FINITE–SAMPLE PROPERTIES

For each of the 10 models given in Table 1, we consider $(N, T) = (10, 100), (20, 100), (20, 200)$, and $(30, 200)$. In results not reported here, an LM test for the null hypothesis that $p_j = 0$ for all j as well as a t test for $E(\Delta\phi_j) = .75$ have a rejection rate of 1 when tested on DGPs 2–10. These rejections are uninformative about the extent of the cross-section correlation, however.

Besides the Pearson correlation, we also use other measures for $\hat{p}_j$. Let $\mathbf{R}_i$ be a $T \times 1$ vector of rankings of z_{it}. The Spearman rank correlation coefficient is defined as

$$\hat{r}_{ij} = \frac{\mathcal{R}_{ij}}{\sqrt{\mathcal{R}_{ii}\mathcal{R}_{jj}}},$$

where $\mathcal{R}_{ij} = \text{cov}(\mathbf{R}_i, \mathbf{R}_j) = \frac{1}{T}\sum_{t=1}^{T}(R_{it} - (T+1)/2)(R_{jt} - (T+1)/2)$ is the sample covariance in the rankings of two series, z_{it} and z_{jt}. It is well known that $\sqrt{(T-2)} \cdot \hat{r}_{ij}/\sqrt{1-\hat{r}_{ij}^2} \sim t_{T-2}$. The statistic has mean 0 and variance $\frac{T-3}{T-5}$, which becomes approximately unity as $T \to \infty$. We also consider Fisher's z-transformation of the Pearson correlation coefficient,

$$\hat{c}_{ij}(z) = \frac{1}{2}\ln\left(\frac{1+\hat{c}_{ij}}{1-\hat{c}_{ij}}\right).$$

If $c_{ij} = 0$, then $\hat{c}_{ij}(z) \approx \text{N}(0, \frac{1}{T-3})$. The transformation is especially useful in stabilizing the variance when T is small. Analogously, a transformation can also be applied to the Spearman rank correlations and Kendall's τ.

To summarize, given N cross-section units over T time periods with $n = N(N-1)/2$, the following variables are constructed:

$\hat{c}_{ij}(z)$:	sample cross-section correlations
$\hat{p}_j$:	vech(x), where x is one of $\hat{c}$, $\hat{c}(z)$, $\hat{r}$, or $\hat{r}(z)$
$\bar{p}_j$:	$\lvert\hat{p}_j\rvert$
$\bar{p}_{[j:n]}$:	$\bar{p}_j$ ordered from smallest to largest
$\bar{\phi}_j$:	$\Phi(\sqrt{T}\bar{p}_{[j:n]})$
$\bar{\phi}_j^n$:	$n \cdot \bar{\phi}_j$
$\Delta^q\bar{\phi}_j^n$:	$\bar{\phi}_j^n - \bar{\phi}_{j-q}^n$.

Tables 2 and 3 report results for the Pearson correlation coefficients assuming that ϵ_{it} is normally distributed. Results for Fisher's z correlations are given in Tables 4 and 5. Results for the Spearman rank correlations are similar and hence are not reported here. The sample correlations are estimated more precisely the larger the T, although the sample-splitting procedure is more accurate the larger the N. Because it is difficult to identify mean shifts occurring at the two ends of the sample, we use 10% trimming; that is, the smallest and largest 10% of $\bar{\phi}_j$ are not used in the breakpoint analysis. Recall that $\hat{\theta}$ estimates the fraction of "small" correlations. The average and the standard deviation of $\hat{\theta}$ over 1,000 replications are reported in columns three and four. Given that we use 10% trimming, having an average $\hat{\theta}$ of around .1 for DGP 2 is clear evidence of extensive correlation. The average $\hat{\theta}$ is around .3 for DGP 3. The reason that DGPs 2 and 3 have rather different $\hat{\theta}$'s even though the θ_0's are the same is that DGP 3 has higher idiosyncratic error variance. Consequently, more correlations are nonzero, albeit small. This again highlights the point that θ_0 (the fraction of zero correlations) need not be the same as θ (the fraction of small correlations). Although θ_0 is a useful benchmark, θ is more informative from a practical perspective, and an estimate of θ is provided by our breakpoint analysis.

Table 2. Pearson Correlation Coefficients: Normal Errors

		First split				Second split		
DGP	θ_0	$\hat{\theta}$	std($\hat{\theta}$)	svr_S	svr_L	$\hat{\theta}$	std($\hat{\theta}$)	svr_{SS}
$N = 10,\ T = 100$								
1	1.000	.475	.252	.091	.095	.549	.264	.048
2	0	.171	.092	.272	.938	.392	.233	.021
3	0	.279	.166	.095	.764	.452	.247	.043
4	.622	.524	.161	.059	.837	.539	.271	.038
5	.333	.416	.158	.088	.883	.480	.260	.045
6	.867	.728	.212	.066	.767	.538	.278	.057
7	.378	.327	.080	.074	.953	.553	.267	.052
8	.867	.539	.261	.092	.263	.557	.264	.049
9	.378	.489	.182	.064	.725	.497	.266	.054
10	.378	.385	.147	.079	.877	.524	.284	.052
$N = 20,\ T = 100$								
1	1.000	.483	.271	.062	.068	.562	.272	.049
2	0	.150	.076	.200	.998	.355	.251	.049
3	0	.302	.138	.060	.990	.418	.236	.053
4	.805	.762	.115	.051	.947	.493	.261	.043
5	.632	.692	.131	.063	.975	.465	.248	.063
6	.853	.831	.062	.046	.984	.505	.273	.055
7	.368	.345	.046	.071	1.000	.522	.278	.061
8	.853	.760	.192	.058	.603	.485	.272	.056
9	.368	.540	.129	.047	.963	.445	.243	.057
10	.368	.419	.112	.093	.998	.483	.264	.060
$N = 20,\ T = 200$								
1	1.000	.482	.275	.065	.059	.571	.271	.062
2	0	.128	.051	.285	.999	.314	.243	.030
3	0	.237	.117	.096	.997	.385	.229	.054
4	.805	.774	.086	.059	.983	.534	.269	.051
5	.632	.666	.121	.058	.977	.483	.256	.053
6	.853	.827	.073	.048	.980	.509	.273	.057
7	.368	.349	.042	.079	1.000	.554	.277	.062
8	.853	.796	.169	.042	.774	.490	.269	.052
9	.368	.507	.119	.058	.990	.454	.246	.060
10	.368	.400	.100	.087	.998	.492	.259	.051
$N = 30,\ T = 200$								
1	1.000	.501	.270	.055	.055	.571	.279	.055
2	0	.127	.046	.332	1.000	.342	.246	.061
3	0	.240	.103	.094	1.000	.435	.216	.056
4	.869	.857	.037	.063	.997	.543	.260	.057
5	.747	.790	.071	.051	.999	.513	.253	.044
6	.848	.841	.024	.067	.999	.537	.267	.056
7	.366	.355	.026	.081	1.000	.564	.270	.057
8	.848	.858	.075	.060	.938	.488	.264	.055
9	.366	.525	.090	.072	1.000	.510	.239	.060
10	.366	.417	.081	.081	1.000	.495	.257	.058

NOTE: $\hat{\theta}$ is the estimated fraction of the sample in *S* (small correlations). std($\hat{\theta}$) denotes the standard deviation of $\hat{\theta}$ in 1,000 replications. svr_S and *svr* are the rejection rates of the standardized spacings variance-ratio test applied to the subsamples *S* and *L*.

Table 3. Fisher-Transformed Pearson Correlation Coefficients: Normal Errors

		First split				Second split		
DGP	θ_0	$\hat{\theta}$	std($\hat{\theta}$)	svr_S	svr_L	$\hat{\theta}$	std($\hat{\theta}$)	svr_{SS}
$N = 10,\ T = 100$								
1	1.000	.478	.253	.091	.097	.546	.264	.048
2	0	.170	.092	.271	.941	.391	.233	.022
3	0	.280	.166	.095	.776	.452	.247	.044
4	.622	.525	.160	.060	.841	.540	.272	.038
5	.333	.416	.157	.087	.882	.481	.259	.046
6	.867	.729	.210	.064	.770	.538	.277	.059
7	.378	.327	.079	.074	.954	.552	.267	.054
8	.867	.545	.261	.092	.276	.556	.263	.048
9	.378	.489	.180	.065	.736	.497	.267	.054
10	.378	.385	.146	.080	.884	.525	.283	.053
$N = 20,\ T = 100$								
1	1.000	.485	.272	.064	.071	.560	.272	.049
2	0	.150	.076	.199	.999	.355	.251	.049
3	0	.302	.138	.060	.990	.420	.238	.053
4	.805	.762	.112	.051	.951	.490	.258	.041
5	.632	.691	.128	.064	.976	.463	.248	.061
6	.853	.829	.062	.046	.986	.506	.272	.054
7	.368	.344	.046	.069	1.000	.517	.277	.060
8	.853	.767	.186	.059	.633	.481	.271	.056
9	.368	.539	.129	.048	.966	.446	.244	.057
10	.368	.418	.111	.093	.998	.484	.264	.062
$N = 20,\ T = 200$								
1	1.000	.485	.275	.064	.059	.568	.272	.063
2	0	.128	.051	.285	.999	.314	.243	.030
3	0	.237	.117	.094	.997	.385	.228	.054
4	.805	.773	.086	.060	.984	.535	.269	.052
5	.632	.665	.121	.058	.978	.486	.256	.054
6	.853	.826	.073	.050	.982	.510	.274	.058
7	.368	.349	.042	.077	1.000	.551	.276	.062
8	.853	.798	.166	.042	.786	.489	.269	.052
9	.368	.507	.118	.059	.990	.454	.247	.058
10	.368	.400	.099	.089	.998	.493	.259	.049
$N = 30,\ T = 200$								
1	1.000	.505	.268	.053	.056	.574	.279	.055
2	0	.127	.046	.332	1.000	.343	.247	.061
3	0	.240	.103	.093	1.000	.435	.216	.055
4	.869	.857	.037	.065	.997	.543	.260	.058
5	.747	.788	.071	.050	.999	.510	.253	.046
6	.848	.840	.025	.067	.999	.537	.266	.054
7	.366	.355	.025	.080	1.000	.560	.269	.057
8	.848	.859	.072	.062	.947	.488	.264	.054
9	.366	.526	.089	.072	1.000	.510	.240	.061
10	.366	.417	.081	.079	1.000	.497	.257	.058

NOTE: $\hat{\theta}$ is the estimated fraction of the sample in *S* (small correlations). std($\hat{\theta}$) denotes the standard deviation of $\hat{\theta}$ in 1,000 replications. svr_S and *svr* are the rejection rates of the standardized spacings variance-ratio test applied to the subsamples *S* and *L*.

Next, we apply the *svr* test to the two subsamples with $q = 2$. The critical value for a two-tailed 5% test is 1.96. The results are based on 1,000 replications. The rejection rates are labeled svr_S and svr_L. The rejection rate should be .05 in both subsamples of DGP 1, because the data are uncorrelated. The test is oversized when $(N, T) = (10, 100)$ but is close to the nominal size of 5% for larger (N, T). The size distortion when $N = 10$ is due to the small size of the subsamples (which have <25 observations given that $\hat{\theta}$ is, on average, around .5). For DGPs 2–10, the svr_L tends to reject uniformity with probability close to 1 when $N = 20$, demonstrating that the test has power. Because the correlations in the *L* subsample are large by selection, the ability to reject uniformity in *L* is not surprising. It is more difficult correctly reject uniformity in *S* when the underlying correlations are small but not necessarily 0. The size of the test is generally accurate, especially considering the average size of the subsamples. We also split *S* into two samples, *SS* and *SL*, and then apply the *svr* to *SS*. Table 3 shows that the rejection rates are generally similar to those using the larger sample, *S*, demonstrating that censoring poses no problem for the spacings-based *svr*.

Strictly speaking, Lemmas 1 and 2 apply to iid uniform variables. Except when the data are normally distributed, the sample correlations are not independent. Nonetheless, even when the data are nonnormal, $\hat{c}_{ij}$ is at least asymptotically normally distributed and hence asymptotically independent. The spacings thus should be independently uniformly distributed in large samples. We can use simulations to check robustness of the results against departures from normality in finite samples. Table 4 reports results for Pearson correlations assuming that ϵ_{it} is a chi-squared variable with 1 degree of freedom. Table 5 assumes that ϵ_{it} is a ARCH(1) process (with parameter .5). As we can see, the breakpoint analysis and the *svr* test have proper-

Table 4. Pearson Correlation Coefficients: ARCH(1) Errors

		First split				Second split		
DGP	θ_0	$\hat{\theta}$	std($\hat{\theta}$)	svr_S	svr_L	$\hat{\theta}$	std($\hat{\theta}$)	svr_{SS}
$N = 10$, $T = 100$								
1	1.000	.487	.255	.106	.084	.533	.260	.045
2	0	.170	.091	.272	.939	.399	.239	.017
3	0	.278	.164	.091	.793	.441	.248	.031
4	.622	.511	.174	.075	.822	.515	.268	.058
5	.333	.420	.159	.071	.887	.496	.260	.059
6	.867	.726	.211	.057	.764	.572	.269	.059
7	.378	.330	.078	.078	.956	.540	.269	.061
8	.867	.556	.255	.075	.305	.559	.264	.036
9	.378	.484	.184	.074	.733	.494	.263	.044
10	.378	.377	.149	.089	.880	.519	.270	.043
$N = 20$, $T = 100$								
1	1.000	.512	.273	.072	.083	.580	.272	.062
2	0	.150	.076	.212	.995	.355	.252	.043
3	0	.299	.137	.066	.984	.409	.237	.056
4	.805	.736	.133	.071	.939	.497	.258	.056
5	.632	.689	.132	.052	.976	.485	.256	.057
6	.853	.830	.052	.056	.987	.519	.272	.055
7	.368	.344	.049	.081	.999	.528	.276	.073
8	.853	.765	.189	.053	.650	.511	.271	.040
9	.368	.537	.132	.057	.956	.443	.246	.068
10	.368	.418	.111	.082	.995	.476	.258	.049
$N = 20$, $T = 200$								
1	1.000	.513	.267	.069	.067	.559	.276	.058
2	0	.128	.052	.255	.999	.316	.247	.032
3	0	.239	.123	.094	.996	.396	.229	.058
4	.805	.756	.103	.048	.974	.521	.269	.059
5	.632	.671	.120	.055	.991	.489	.253	.045
6	.853	.829	.063	.056	.990	.521	.270	.056
7	.368	.348	.045	.079	.999	.548	.276	.066
8	.853	.796	.168	.060	.758	.495	.269	.054
9	.368	.510	.117	.060	.985	.462	.245	.076
10	.368	.401	.103	.078	1.000	.511	.269	.047
$N = 30$, $T = 200$								
1	1.000	.496	.274	.059	.066	.576	.276	.058
2	0	.127	.048	.335	1.000	.347	.243	.066
3	0	.239	.100	.097	1.000	.418	.222	.060
4	.869	.842	.068	.050	.992	.537	.258	.042
5	.747	.788	.073	.058	.999	.516	.248	.045
6	.848	.840	.019	.046	1.000	.538	.271	.047
7	.366	.355	.028	.079	1.000	.577	.271	.058
8	.848	.859	.064	.061	.941	.500	.267	.056
9	.366	.524	.088	.068	1.000	.517	.241	.049
10	.366	.416	.084	.098	1.000	.505	.255	.054

NOTE: $\hat{\theta}$ is the estimated fraction of the sample in S (small correlations). std($\hat{\theta}$) denotes the standard deviation of $\hat{\theta}$ in 1,000 replications. svr_S and svr are the rejection rates of the standardized spacings variance-ratio test applied to the subsamples S and L.

Table 5. Pearson Correlation Coefficients: χ^2_1 Errors

		First split				Second split		
DGP	θ_0	$\hat{\theta}$	std($\hat{\theta}$)	svr_S	svr_L	$\hat{\theta}$	std($\hat{\theta}$)	svr_{SS}
$N = 10$, $T = 100$								
1	1.000	.500	.263	.094	.092	.539	.260	.033
2	0	.200	.114	.187	.883	.412	.241	.015
3	0	.358	.188	.071	.681	.460	.255	.046
4	.622	.523	.178	.087	.845	.523	.272	.036
5	.333	.424	.157	.070	.887	.500	.260	.060
6	.867	.680	.236	.074	.633	.561	.274	.052
7	.378	.321	.090	.080	.853	.505	.257	.047
8	.867	.515	.264	.084	.211	.556	.258	.052
9	.378	.511	.212	.072	.565	.494	.266	.046
10	.378	.418	.164	.074	.841	.518	.267	.041
$N = 20$, $T = 100$								
1	1.000	.529	.289	.055	.066	.561	.272	.052
2	0	.186	.102	.138	.995	.383	.237	.061
3	0	.416	.162	.064	.928	.413	.229	.046
4	.805	.761	.130	.049	.937	.504	.263	.052
5	.632	.697	.121	.061	.978	.486	.256	.058
6	.853	.824	.074	.068	.945	.509	.278	.060
7	.368	.346	.053	.047	.996	.467	.254	.059
8	.853	.675	.261	.050	.441	.547	.280	.061
9	.368	.613	.140	.057	.853	.450	.242	.065
10	.368	.473	.111	.095	.994	.463	.255	.066
$N = 20$, $T = 200$								
1	1.000	.516	.282	.057	.065	.551	.267	.059
2	0	.155	.076	.169	.998	.372	.239	.047
3	0	.341	.145	.065	.978	.407	.235	.072
4	.805	.770	.090	.061	.983	.527	.269	.041
5	.632	.672	.117	.051	.985	.494	.257	.044
6	.853	.837	.037	.056	.996	.530	.277	.059
7	.368	.353	.034	.061	1.000	.533	.275	.071
8	.853	.749	.220	.052	.626	.535	.281	.050
9	.368	.581	.121	.058	.943	.440	.241	.051
10	.368	.445	.100	.076	.998	.475	.255	.073
$N = 30$, $T = 200$								
1	1.000	.508	.280	.055	.065	.565	.280	.053
2	0	.156	.073	.187	1.000	.390	.233	.069
3	0	.354	.121	.075	.998	.459	.217	.063
4	.869	.857	.046	.061	.996	.545	.263	.052
5	.747	.796	.061	.068	.996	.539	.252	.053
6	.848	.841	.011	.054	1.000	.538	.276	.059
7	.366	.357	.019	.072	1.000	.543	.270	.048
8	.848	.843	.127	.054	.807	.487	.278	.050
9	.366	.601	.095	.061	.995	.508	.228	.042
10	.366	.458	.081	.068	1.000	.505	.244	.046

NOTE: $\hat{\theta}$ is the estimated fraction of the sample in S (small correlations). std($\hat{\theta}$) denotes the standard deviation of $\hat{\theta}$ in 1,000 replications. svr_S and svr are the rejection rates of the standardized spacings variance-ratio test applied to the subsamples S and L.

ties similar to those reported in Tables 2 and 3. This robustness to departures from normality is likely due to the fact that the breakpoint analysis is valid under very general assumptions, and the SVR tests a generic feature of integrated data. Thus both the breakpoint analysis and the SVR test continue to have satisfactory finite-sample properties.

5.1 Applications

We consider two applications. The first analyzes the correlation in real economic activity in the Euro area, and the second considers a panel of real exchange rate data. To isolate cross-section from serial correlation, we compute the correlation coefficients for residuals from a regression of each variable of interest on a constant and its own lags.

Industrial Production. Let z_{it}, $i = 1, 12$, $t =$ 1982:1–1997:8, be monthly data on industrial production for Germany, Italy, Spain, France, Austria, Luxembourg, the Netherlands, Finland, Portugal, Belgium, Ireland, and the United States. These data are the same as given by Stock, Watson, and Marcellino (2003). With 12 countries, we have $n = 66$ for $T = 186$ observations. The variable of interest is the growth rate of industrial production. Figure 4 plots $\bar{\phi}_j$. Although most of the $\bar{\phi}_j$'s are far from unity, the observations do not lie along a straight line, indicating substantial heterogeneity in the correlations. The breakpoint analysis yields $\hat{\theta} = .468$, 31 correlations in S, and 35 correlations in L. The svr test statistics are $-.234$ and 2.673. Industrial production in 31 of the 66 country pairs appear to be contemporaneously uncorrelated. Because there are many zero correlation pairs, a common factor structure does not appear to be a suitable characterization of the output of these 12 countries.

A listing of the correlations in the two groups is given in Table 6. The largest correlation is the France–Germany pair

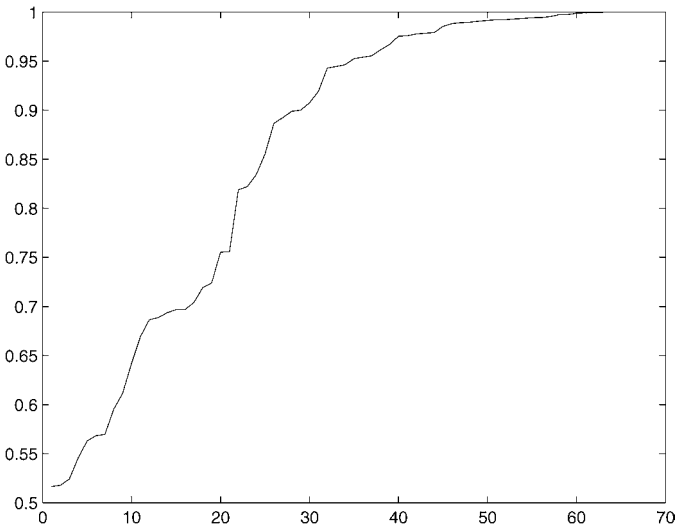

Figure 4. $\bar{\phi}(j)$: Growth Rate of Industrial Production for 12 OECD Countries.

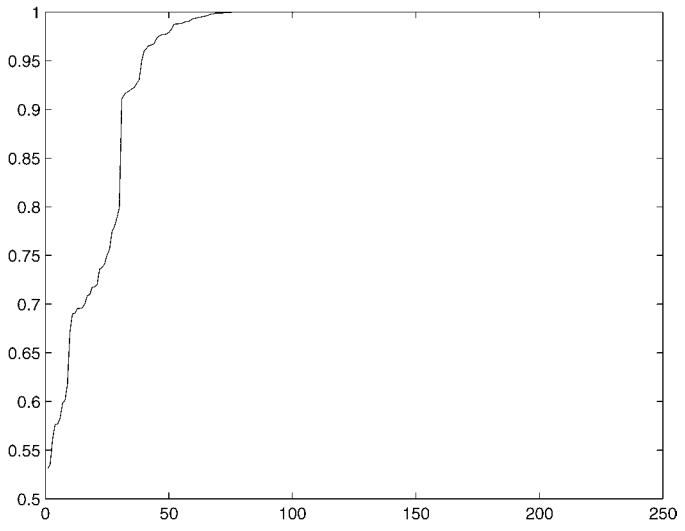

Figure 5. $\bar{\phi}(j)$: Real Exchange Rates for 21 Countries.

($\bar{p}_j = .33$), and the weakest is between Austria and Luxembourg ($\bar{p}_j = .006$). As pointed out by a referee, France is highly correlated with Belgium, which is highly correlated with Spain, but the correlation between Spain and Belgium is in the low group. This result, which is of some economic interest, clearly shows that our analysis provides a way to classify correlation coefficients into groups. Economic analysis will be necessary to see whether the series underlying the correlations can be grouped meaningfully.

Table 6. Correlations in Industrial Production for 12 OECD Countries

Group S			Group L		
Country		$\hat{p}_j$	Country		$\hat{p}_j$
FRA	BEL	.003	FRA	FIN	.116
GER	PORT	−.003	AUS	FIN	.117
NETH	BEL	.004	GER	LUX	.118
BEL	US	−.008	LUX	PORT	.123
AUS	IRE	.012	IRE	US	.124
FIN	IRE	−.013	ITA	AUS	.125
PORT	BEL	−.013	ITA	BEL	.130
ITA	FIN	−.018	FRA	PORT	.135
FRA	IRE	.021	BEL	IRE	.144
AUS	PORT	−.027	FRA	LUX	.145
AUS	LUX	.032	GER	NETH	.147
SPA	FIN	.036	GER	BEL	.148
FRA	AUS	.036	LUX	US	.149
SPA	BEL	.037	LUX	BEL	.160
LUX	NETH	.038	GER	SPA	.166
NETH	FIN	−.038	FRA	NETH	.169
NETH	US	.039	SPA	US	.170
GER	US	.043	FIN	US	.173
GER	AUS	−.044	FIN	BEL	.175
LUX	IRE	.051	ITA	PORT	.177
GER	FIN	.051	GER	IRE	.178
PORT	IRE	.067	SPA	IRE	.180
PORT	US	.068	SPA	LUX	.182
NETH	IRE	.071	AUS	NETH	.185
GER	ITA	.078	SPA	PORT	.186
LUX	FIN	.089	ITA	NETH	.191
NETH	PORT	.091	SPA	NETH	.205
SPA	AUS	.094	ITA	IRE	.205
FIN	PORT	.094	ITA	SPA	.226
AUS	US	.097	ITA	US	.233
ITA	FRA	.103	FRA	US	.250
			AUS	BEL	.293
			ITA	LUX	.299
			SPA	FRA	.364
			GER	FRA	.372

NOTE: The industrial production data are for Germany, Italy, Spain, France, Austria, Luxembourg, the Netherlands, Finland, Portugal, Belgium, Ireland, and the United States for 1982:1–1997:8.

Real Exchange Rates. Quarterly data for nominal exchange rates and the consumer price indices were obtained from the International Finance Statistics. We use data from 1974:1–1997:4 for 21 countries: Canada, Austria, New Zealand, Australia, Belgium, Denmark, Finland, France, Germany, Ireland, Italy, Netherlands, Norway, Spain, Sweden, Switzerland, U.K., Japan, Korea, Singapore, and Thailand. The United States is used as the numeraire country. The variable of interest is the log real exchange rate.

A plot of $\bar{\phi}_j$ is presented in Figure 5. The algorithm finds that $\hat{\theta} = .142$, so .85 of the sample belongs to L. Because there are a total of $n = 210$ correlations in this example, Table 7 gives only the 30 correlations in S and the 30 largest correlations in L. Evidently, correlations involving the real exchange rates for Canada and Korea do not display much correlation with those of the European countries. However, the European real exchange rates are strongly correlated with one another. The results suggest that one cannot simply attribute the source of the correlations to the use of a common numeraire (the U.S. dollar). If this were the sole source of the correlation, then all real exchange rates should be equally correlated.

The *svr* test statistic with $q = 2$ is 1.884 for the S sample of size 30 and 8.574 for the L sample. With $q = 4$, these statistics test are 1.4585 for S and 10.8271 for S. Further partitioning the L sample yields a *svr* for LS of 8.5724. Evidence against no correlation in the L sample is compelling. Because L constitutes .85 of the correlation pairs, the cross-section correlation in the panel is extensive. Not accounting for these correlations when constructing panel unit root tests can indeed lead to misleading inference about persistence.

Putting the two sets of results together suggests stronger and more extensive correlation for the European real exchange rates than for industrial production. Indeed, the first principal component explains more than 60% of the variation in real exchange rates. Because a common factor exists in European real exchange rates but not in real output, the evidence points to the presence of a nominal or a monetary factor.

Table 7. Real Exchange Rates for 21 Countries

Group S			Largest 30 correlations in group L		
Country		$\hat{p}_j$	Country		$\hat{p}_j$
CAN	FR	.008	NETH	NOR	.847
CAN	NETH	.009	AUS	IRE	.851
CAN	SPA	.016	FR	SWTZD	.857
CAN	GER	.019	GER	IRE	.860
CAN	ITA	.020	BEL	SWTZD	.862
CAN	BEL	.021	BEL	IRE	.867
CAN	NOR	.025	DEN	SWTZD	.867
UK	KOR	−.026	DEN	IRE	.869
CAN	AUS	.030	IRE	NETH	.869
SWTZD	KOR	.045	FR	IRE	.870
CAN	SWED	.050	GER	NOR	.871
FIN	KOR	.050	AUS	NOR	.872
DEN	KOR	.052	AUS	SWTZD	.876
CAN	SWTZD	.052	NETH	SWTZD	.877
CAN	SG	.052	GER	SWTZD	.879
NETH	KOR	.053	DEN	FR	.913
SWED	KOR	.056	AUS	FR	.921
CAN	IRE	.056	BEL	FR	.922
CAN	THAI	.058	FR	GER	.922
ITA	KOR	.058	FR	NETH	.927
FR	KOR	.059	AUS	DEN	.950
BEL	KOR	.064	DEN	GER	.956
CAN	JP	.064	DEN	NETH	.959
GER	KOR	.065	BEL	DEN	.963
AUS	KOR	.068	AUS	BEL	.963
CAN	DEN	.070	BEL	GER	.964
IRE	KOR	.076	BEL	NETH	.970
NOR	KOR	.078	AUS	NETH	.978
CAN	UK	.081	GER	NETH	.982
SPA	KOR	.085	AUS	GER	.985

NOTE: The real exchange rates are for Canada, Australia, New Zealand, Australia, Belgium, Denmark, Finland, France, Germany, Ireland, Italy, Netherlands, Norway, Spain, Sweden, Switzerland, U.K., Japan, Korea, Singapore, and Thailand. The United States is used as the numeraire country. The sample is for 1974:1–1997:4.

6. CONCLUSION

This article has presented an agnostic way of testing cross-section correlation when possibly a subset of the sample is correlated. This is a situation in which traditional testing of the no correlation hypothesis in the full sample is rejected, but for which the rejection provides little understanding about the extent of the correlation. Our analysis makes use of the fact that if the population correlations are 0, then the spacings of the transformed sample correlations should have certain properties. We use a breakpoint analysis and a variance ratio test to assess whether the properties of the spacings are consistent with the underlying assumption of zero correlation. We take a Gaussian transformation of $|\sqrt{T}\bar{p}_{[j:n]}|$. An alternative is to define $\tilde{p}_{[j:n]}$ as the sorted series of $\hat{p}_j^2$. Then a chi-squared transformation of $T\tilde{p}_{[j:n]}$ should be approximately uniformly distributed on the unit interval. The results for uniform spacings can be applied immediately, and our main results will continue to hold.

APPENDIX: PROOFS

We need to show that the tests, which are based on transformations of the sample correlations, are unaffected by the sampling variability of estimated correlation coefficients. Let $\Phi(\cdot)$ denote the standard normal distribution and z_α be the α-level standard normal quantile, that is, $\Phi(z_\alpha) = \alpha$. It is known that if a statistic $\hat{\tau}$ is $\sqrt{T}$ consistent and asymptotically normal, then $\Pr(\hat{\tau} \le z_\alpha) = \Phi(z_\alpha) + T^{-1/2}\Phi'(z_\alpha)v_1(z_\alpha) + T^{-1}v_2(z_\alpha)\Phi''(x) + \cdots$, where v_j is a polynomial. Furthermore, if $\Pr(\hat{\tau} \le q_\alpha) = \alpha$, then the quantile q_α admits the expansion $q_\alpha = z_\alpha + T^{-1/2}\Phi'(z_\alpha)w_1(z_\alpha) + \cdots$, where w_j is a polynomial. Thus, by the mean value theorem, $\Phi(q_\alpha) = \Phi(z_\alpha) + \Phi'(\bar{q}_\alpha)(q_\alpha - z_\alpha)$ for $\bar{q}_\alpha \in [q_\alpha, z_\alpha]$.

Under the hypothesis of no correlation, our test statistic is $\bar{\tau}_j = \sqrt{T}\bar{p}_{[j:n]}$. Let $\bar{\phi}_j = \Phi(\sqrt{T}|\bar{p}_{[j:n]}|)$, $\phi_j = \Phi(|z_{[j:n]}|)$, with $z_{[j:n]}$ defined such that $\Phi(z_{[j:n]}) = \frac{j}{n+1}$. We may write

$$\bar{\phi}_j = \phi_j + \bar{u}_j,$$

where $\phi_j \sim \mathrm{U}[.5, 1]$ and $\bar{u}_j = O_p(T^{-1/2})$ for a given j and n. However, because $\phi_j = .5 + \frac{j}{2(n+1)}$, $\bar{u}_j = O_p(n^{-1}T^{-1/2})$ as n, T increases. Thus $n\Delta\bar{\phi}_j = n\Delta\phi_j + n\Delta\bar{u}_j$, where $n\Delta\bar{u}_j$ is $O_p(T^{-1/2})$. Similarly, $n\Delta^q\bar{\phi}_j = n\Delta^q\phi_j + O_p(T^{-1/2})$.

We have for $q = 1$, $\hat{\mu}_1 = \frac{1}{n}\sum_{j=1}^n n\Delta\bar{\phi}_j = \frac{1}{n}\sum_{j=1}^n n\Delta\phi_j + O_p(T^{-1/2})$. Thus $\frac{1}{n}\sum_{j=1}^n n\Delta\bar{\phi}_j = \frac{1}{n}\sum_{j=1}^n n\Delta\phi_j + o_p(1)$. Furthermore, $\hat{\sigma}_1^2 = \frac{1}{n}\sum_{j=1}^n (n\Delta\bar{\phi}_j)^2 - \hat{\mu}_1^2$. It is straightforward to verify that $\frac{1}{n}\sum_{j=1}^n (n\Delta\bar{\phi}_j)^2 = \frac{1}{n}\sum_{j=1}^n (n\Delta\phi_j)^2 + O_p(T^{-1/2})$. Analogous arguments show that for fixed q, the mean and variance of $n\Delta^q\bar{\phi}_j$ are the same as those of $n\Delta^q\phi_j$ if $T \to \infty$. This is verified in Table 8. Testing $\frac{1}{n}\sum_{j=1}^n \Delta\bar{\phi}_j$ is also same as testing the mean of a uniformly distributed series as $n \to \infty$, which is the basis of the breakpoint analysis.

Table 8. Properties of $n\Delta^q\phi_j$ versus $n\Delta^q\bar{\phi}_j$, $\phi_j \sim U[.5, 1]$, $\bar{\phi}_j = \Phi(\sqrt{T}|\bar{p}_{[j:n]}|)$

		$q = 1$				$q = 2$			
N	T	$\overline{n\Delta^q\bar{\phi}_j}$	$\overline{n\Delta^q\phi_j}$	var($n\Delta^q\bar{\phi}_j$)	var($n\Delta^q\phi_j$)	$\overline{n\Delta^q\bar{\phi}_j}$	$\overline{n\Delta^q\phi_j}$	var($n\Delta^q\bar{\phi}_j$)	var($n\Delta^q\phi_j$)
5	50	.452	.454	.187	.191	.910	.910	.321	.323
5	100	.451	.453	.186	.185	.907	.908	.331	.328
5	200	.454	.452	.186	.184	.914	.908	.342	.330
10	50	.489	.490	.234	.233	.980	.978	.457	.459
10	100	.489	.488	.234	.236	.978	.980	.447	.450
10	200	.489	.489	.234	.235	.979	.980	.457	.454
20	50	.497	.497	.245	.248	.995	.994	.489	.490
20	100	.497	.497	.247	.246	.995	.994	.490	.493
20	200	.497	.497	.246	.246	.995	.995	.489	.487
30	50	.499	.499	.248	.248	.998	.997	.495	.498
30	100	.499	.499	.248	.247	.998	.998	.493	.495
30	200	.499	.499	.248	.249	.998	.998	.494	.496

Proof of Theorem 1

Let D_j be a first-order uniform spacing, and let D_j^q be a q-order spacing. The statistic $G_{1n} = \frac{1}{n}\sum_{j=1}^{n}(nD_j)^2$ is referred to as the first-order Greenwood statistic, and $G_{3n} = \sum_{j=1}^{n}(n\Delta^q D_j)^2$ is a generalized Greenwood statistic based on overlapping spacings. As summarized by Rao and Kuo (1984),

$$\sqrt{n}(G_{1n} - 2) \xrightarrow{d} \mathrm{N}(0, 4)$$

and

$$\sqrt{n}\big(G_{3n} - q(q+1)\big) \xrightarrow{d} \mathrm{N}\left(0, \frac{2q(q+1)(2q+1)}{3}\right).$$

Now $E(D_j) = 1/(n+1)$. Thus $\mathrm{var}(nD_j) = \frac{1}{n}\sum_{j=1}^{n}(nD_j - 1)^2 = \frac{1}{n}\sum_{j=1}^{n}(nD_j)^2 - 1$ has the same limiting variance as G_{1n}. Likewise, $\mathrm{var}(D_i^q) = \frac{1}{n}\sum_{i=1}^{n}(nD_i^q - q)^2$ has the same limiting variance as G_{3n}.

As defined, $\hat{\sigma}_1^2 = \mathrm{var}(nD_j)$ and $\sigma_q^2 = \mathrm{var}(nD_j^q/q)$. Thus

$$\mathrm{var}(\hat{\sigma}_1^2) = 4$$

and

$$\mathrm{var}(\hat{\sigma}_q^2) = \frac{1}{q^2}\,\mathrm{var}(nD_j^q) = \frac{2(q+1)(2q+1)}{3q}.$$

Although $\hat{\sigma}_1^2$ and $\hat{\sigma}_q^2$ are both consistent estimates of σ_ϵ^2, $\hat{\sigma}_1^2$ is asymptotically efficient. Thus, following the argument of Lo and MacKinlay (1988), we have

$$\begin{aligned}\mathrm{var}(\hat{\sigma}_q^2 - \hat{\sigma}_1^2) &= \mathrm{var}(\hat{\sigma}_q^2) - \mathrm{var}(\hat{\sigma}_1^2)\\ &= \frac{2(q+1)(2q+1)}{3q} - 4 = \frac{2(2q-1)(q-1)}{3}.\end{aligned}$$

This is precisely the asymptotic variance of the SVR statistic of Lo and MacKinlay (1988).

ACKNOWLEDGMENTS

The author thanks Jushan Bai, Chang-Ching Lin, and Yuriy Gorodnichenko, and seminar participants at the 2004 MidWest Econometrics Group Meeting, UCLA, Rutgers University, the Federal Reserve Board, two anonymous referees, and the associate editor for many helpful discussions. Financial support from the National Science Foundation (grant SES-0345237) is gratefully acknowledged.

[Received June 2004. Revised July 2005.]

REFERENCES

Andrews, D. K. (2003), "Cross-Section Regression With Common Shocks," Discussion Paper 1428, Cowleds Foundation.

Arnold, B., Balakrishman, N., and Nagaraja, H. (1992), *A First Course in Order Statistics* (2nd ed.), New York: Wiley.

Bai, J. (1997), "Estimating Multiple Breaks One at a Time," *Econometric Theory*, 13, 551–564.

Bai, J., and Ng, S. (2002), "Determining the Number of Factors in Approximate Factor Models," *Econometrica*, 70, 191–221.

——— (2004), "A PANIC Attack on Unit Roots and Cointegration," *Econometrica*, 72, 1127–1177.

Breusch, T., and Pagan, A. (1980), "The Lagrange Multiplier Test and Its Applications to Model Specification in Econometrics," *Review of Economic Studies*, 47, 239–254.

Dufour, J. M., and Khalaf, L. (2002), "Exact Tests for Contemporaneous Correlation of Disturbances in Seemingly Unrelated Regressions," *Journal of Econometrics*, 106, 143–170.

Durbin, J. (1961), "Some Methods of Constructing Exact Tests," *Biometrika*, 48, 41–55.

Frees, E. (1995), "Assessing Cross-Sectional Correlation in Panel Data," *Journal of Econometrics*, 69, 393–414.

Holst, L. (1979), "Asymptotic Normality of Sum-Functions of Spacings," *The Annals of Probability*, 7, 1066–1072.

Levin, A., Lin, C. F., and Chu, J. (2002), "Unit Root Tests in Panel Data: Asymptotic and Finite-Sample Properties," *Journal of Econometrics*, 98, 1–24.

Lo, A., and MacKinlay, C. (1988), "Stock Market Prices Do Not Follow Random Walks: Evidence From a Simple Specification Test," *Review of Financial Studies*, 1, 41–66.

Meng, X. L., Rosenthal, R., and Rubin, D. (1992), "Comparing Correlated Correlation Coefficients," *Psychological Bulletin*, 111, 172–175.

Moon, R., and Perron, B. (2004), "Testing for a Unit Root in Panels With Dynamic Factors," *Journal of Econometrics*, 122, 81–126.

Moulton, B. R. (1987), "Diagnostics for Group Effects in Regression Analysis," *Journal of Business & Economic Statistics*, 5, 275–282.

——— (1990), "An Illustration of a Pitfall in Estimating the Effects of Aggregate Variables on Micro Units," *Review of Economics and Statistics*, 72, 334–338.

O'Connell, P. (1998), "The Overvaluation of Purchasing Power Parity," *Journal of International Economics*, 44, 1–19.

Olkin, I., and Finn, J. (1990), "Testing Correlated Correlations," *Psychological Bulletin*, 108, 330–333.

Pyke, R. (1965), "Spacings," *Journal of the Royal Statistical Society*, Ser. B, 27, 395–449.

Rao, J. S., and Kuo, M. (1984), "Asymptotic Results on the Greenwood Statistics and Some of Its Generalizations," *Journal of the Royal Statistical Society*, Ser. B, 46, 228–237.

Richardson, M., and Smith, T. (1993), "A Test for Multivariate Normality in Stock Returns," *Journal of Business*, 66, 295–321.

Steiger, J. (1980), "Tests for Comparing Elements of a Correlation Matrix," *Psychological Bulletin*, 87, 245–251.

Stock, J. H., and Watson, M. W. (2002), "Forecasting Using Principal Components From a Large Number of Predictors," *Journal of the American Statistical Association*, 97, 1167–1179.

Stock, J. H., Watson, M. W., and Marcellino, M. (2003), "Macroeconomic Forecasting in the Euro Area: Country-Specific versus Area-Wide Information," *European Economic Review*, 47, 1–18.

Wells, T., Jammalamadaka, S., and Tiwari, R. (1992), "Tests of Fit Using Spacing Statistics With Estimated Parameters," *Statistics and Probability Letters*, 13, 365–372.