Economics Bulletin

Volume 29, Issue 3

Panel unit root testing and the martingale difference hypothesis for German stocks

Matei Demetrescu
Goethe University Frankfurt

Abstract

Several panel unit root tests based on different ways to account for cross-unit dependence are reviewed. The note then illustrates the tests by checking whether the martingale difference hypothesis is appropriate for stock prices on the German stock market: according to the martingale difference hypothesis, logarithmized stock prices follow an integrated process without short-run dynamics. Compared with usual tests for no autocorrelation, unit root tests do not require strong moment conditions and can cope with stock returns series exhibiting infinite kurtosis. Evidence against the martingale difference hypothesis is found in a panel of 30 DAX stocks observed daily between 2004 and 2007.

The author would like to thank Uwe Hassler, Vladimir Kuzin, Marco J. Lombardi, Helmut Lütkepohl, and Adina I. Tarcolea for very helpful comments and suggestions improving the paper. The research for this note was completed while the author was a Max Weber Fellow at the European University Institute in Florence, whose support is gratefully acknowledged.
Citation: Matei Demetrescu, (2009) "Panel unit root testing and the martingale difference hypothesis for German stocks", *Economics Bulletin*, Vol. 29 no.3 pp. 1749-1759.
Submitted: Mar 18 2009. **Published:** July 24, 2009.

1. Motivation

Modeling the behavior of stock market prices has always received much attention in economic research. The martingale difference hypothesis has been the preferred model, from early random walk models to today's GARCH-type models that allow one to quantify behavior of stock volatility as well. In its current formulation, the martingale difference hypothesis implies stock returns to have the property that the optimal prediction at time t, conditioned on all observations up to $t-1$, equals the unconditional expectation of the returns series. It implies that logarithms of stock prices follow an integrated process with uncorrelated increments.

Along these lines, Meese and Singleton (1982) or Baillie and Bollerslev (1989) applied unit root tests to exchange rates; neither study, however, was able to reject the null hypothesis for the daily exchange rates series they examined. This may well mean that the data follow a martingale difference data generating process, but it may equally be true that, although the martingale difference hypothesis is indeed wrong, the data generating process is, in some sense, close to the null hypothesis.

More powerful tests, increasing the chances of discriminating between null and alternative hypotheses, can be obtained when testing in panels. Since (single-unit) unit root tests are expected to have low power when the data generating process is close to the null hypothesis, panel unit root tests are potentially more useful than their individual counterparts, due to their resorting to additional information. An important issue with panel unit root tests, however, is having to deal with possible cross-dependence between the units of the panel (see Breitung and Pesaran, 2008).

This note employs a battery of panel unit root tests which account in different ways for the cross-unit correlation existing in the analyzed data. The standard test procedure for panel unit roots under cross-unit dependence is based on the PANIC methodology of Bai and Ng (2004); PANIC extracts common, and idiosyncratic, factors from the panel and tests these for unit roots. However, it is an approach requiring large time *and* cross-sectional dimension; see Westerlund and Larsson (2007) for a rigorous treatment of the joint test for unit roots in the extracted idiosyncratic factors. Therefore, alternative approaches to capture cross-unit dependence are used here.[1] Illustrating the alternative tests, evidence against the martingale difference

[1] The selection of methods in this note is not meant to be complete; see Breitung and Pesaran (2008) for a more exhaustive review.

hypothesis is found in a panel of 30 daily returns of the 30 components of DAX-30 index for the period January 1st, 2004 to December 31st, 2007.

(Panel) unit root tests have some advantages when testing the martingale difference hypothesis. One could test the null hypothesis of no correlation (implied by the martingale difference hypothesis). The portmanteau or LM tests for no autocorrelation are often used; but they require either an *iid* assumption or finite 4^{th} order moments, and are difficult to robustify against conditional heteroskedasticity, see e.g. Horowitz *et al.* (2006). In the frequency domain, McPherson and Palardy (2007) find evidence against the martingale difference hypothesis in five out of nine international daily stock index returns for the period 1988 to 2002. Still, frequency-domain tests require strong moment assumptions too, see McPherson and Palardy (2007). Since many financial returns series do not have finite kurtosis, the results should be confirmed by methods that are robust to such behavior. Unit root tests on the other hand require considerably weaker moment conditions, and are robust to conditional heteroskedasticity as well as, in the panel context, time-varying conditional correlation. This does come at a cost, as unit root tests may have reduced power against other departures from the martingale difference hypothesis, e.g. autocorrelation.

2. Models and testing procedures

Let us first review the models used in the analysis. Denote y_{it} the log price of stock i at time t, $t = 1, \ldots, T$, $i = 1, \ldots, N$. A time trend is allowed for,

$$y_{it} = \mu_i + \tau_i t + z_{it},$$

capturing a possible non-zero mean of the returns series. The stochastic components are, according to the martingale difference hypothesis, integrated,

$$z_{it} = z_{it-1} + \varepsilon_{it},$$

where the returns series are $\boldsymbol{\tau} + \boldsymbol{\varepsilon}_t$, where $\boldsymbol{\tau} = (\tau_1, \ldots, \tau_N)'$, and $\boldsymbol{\varepsilon}_t = (\varepsilon_{1t}, \ldots, \varepsilon_{Nt})'$ is a N-dimensional vector martingale difference sequence,

$$E\left(\boldsymbol{\varepsilon}_t | \boldsymbol{\varepsilon}_{t-1}, \boldsymbol{\varepsilon}_{t-2}, \ldots\right) = \mathbf{0}.$$

The increments $\boldsymbol{\varepsilon}_t$ are allowed to exhibit conditional heteroskedasticity or dynamic conditional correlation, as long as the unconditional covariance matrix, $E\left(\boldsymbol{\varepsilon}_t\right) = \Sigma$, is constant, implying constant unconditional variances σ_i^2.

Most unit root tests are embedded in autoregressive alternatives:

$$z_{it} = \rho_i z_{it-1} + \varepsilon_{it},$$

with $z_0 = 0$, or $\Delta z_{it} = (1 - L) z_{it} = \phi_i z_{it-1} + \varepsilon_{it}$, where $\phi_i = \rho_i - 1 = 0$ under the null hypothesis and L is the lag operator, $Lz_{it} = z_{it-1}$. Under the alternative it holds for at least one stock $\phi_i \neq 0$, explosive behavior ($\phi_i > 0$) not being excluded a priori.

Another class of models encompassing unit roots is represented by fractionally integrated processes. Granger and Joyeux (1980) suggest to use such models in order to bridge the gap between stationary processes and unit root processes; they exhibit long range dependence in form of slow decay of their autocorrelation function. Applications of fractionally integrated models to stock market data enjoy increasing popularity, see e.g. Krämer, Sibbertsen and Kleiber (2002) or Gil-Alana (2006). Such models are specified as

$$\Delta^{d_i} z_{it} = \varepsilon_{it},$$

where d is allowed to be non-integer and the so-called fractional difference $\Delta^d = (1 - L)^d$ is given by the binomial series expansion. Should the martingale difference hypothesis hold, it follows $d_i = 1$ for all $i = 1, \ldots, N$ (recovering $(1 - L)z_{it} = \varepsilon_{it}$). Under the alternative, at least one stock exhibits $d_i \neq 1$. Intuitively, fractional integration is closer to integration than the autoregressive model: autoregressive processes with $|\rho| < 1$ are stationary and ergodic, while fractionally integrated processes with $d \geq 0.5$ do not have finite variance and are not ergodic, although they are mean-reverting for $d < 1$, see Granger and Joyeux (1980) for details.

In an autoregressive framework, the null hypothesis $\rho_i = 1$ is tested by means of the Dickey-Fuller [DF] test in the following test equation:

$$\Delta y_{it} = \widehat{\mu}_i + \widehat{\tau}_i t + \widehat{\phi}_i y_{it-1} + \widehat{\varepsilon}_{it}, \tag{1}$$

and the usual t statistic of ϕ_i follows the so-called DF-distribution under the null. Here, an advantage of the DF test is that it has power against $\phi_i < 0$ as well as against $d_i < 1$ (Krämer, 1998), although the power against fractional alternatives can be low, see Hassler and Wolters (1994). Moreover, in the absence of short-run dynamics, the moment conditions required to establish its asymptotic distribution are fairly weak: Phillips (1987) shows $E\left|\varepsilon_{it}\right|^{\alpha} < \infty$ for some $\alpha > 2$ to be a sufficient condition; conditional heteroskedasticity can also be allowed for. This accommodates fat-tailed increments ε_{it} with volatility clustering, and is particularly useful, since the data employed here do not have finite 4^{th} order moments (see the following section).

The panel character of stock market data (i.e. several stock prices observed over a period of time) takes us along the path to *joint* panel testing of relevant hypotheses. Not only does the panel setup enable us to test hypotheses that should, according to economic theory, hold for all units of the panel, but one can also expect an increase in power in many circumstances. The more homogenous (under the alternative) the panel is, the larger the power gain would be. But panel tests are more difficult to conduct under cross-unit dependence, which can hardly be dismissed for stock returns.

An intuitive way to account for cross-dependence is to orthogonalize – GLS-wise – the units prior to the application of a DF-type test; see Breitung and Das (2005); but their procedure requires weak correlation among the units of the panel. The more popular technique to account for cross-unit dependence is the factor-based PANIC method due to Bai and Ng (2004); but there are other methods as well, perhaps better suited to the size of the sample we work with.

Thus, the empirical analysis in Section 3 uses alternative unit root tests robust to cross-sectional dependence. First, the panel nonlinear instrumental variable test due to Chang (2002);[2] second, the panel test resulting from combining the significance of individual DF tests (Demetrescu, Hassler and Tarcolea, 2006); third, the panel test of Hassler, Demetrescu and Tarcolea (2007), based on the univariate fractional integration test of Demetrescu, Kuzin and Hassler (2008) with orthogonalization in the manner of Breitung and Das (2005); and fourth, the panel stationarity test proposed by Tarcolea, Hassler and Demetrescu (2009) which can be used as a panel unit root test as well, see Tarcolea, Hassler and Demetrescu (2009) for details. The four tests are denoted S_N, $t(\hat{\rho}^*, 0.2)$, χ^2, and $\widetilde{\kappa}$, respectively.

3. Empirical findings

The examined data consist of logarithms of closing prices of the 30 stocks composing the DAX-30 index on December 31st, 2007. They cover the period

[2]The asymptotics allowing for martingale difference innovations have not been worked out yet, but Monte Carlo experiments (available upon request) show Chang's test to work reliably under the null even for innovations with extreme conditional heteroskedasticity.

from January 1st, 2004 to December 31st, 2007, or a total of 1022 trading days. The Deutsche Postbank is in the DAX 30 only since June 22nd, 2004; it only exhibits a number of 901 observations, and the panel is unbalanced. Under the null, it is expected that there is no cross-cointegration of levels and no autocorrelation of increments. The data is depicted in Figure 1.

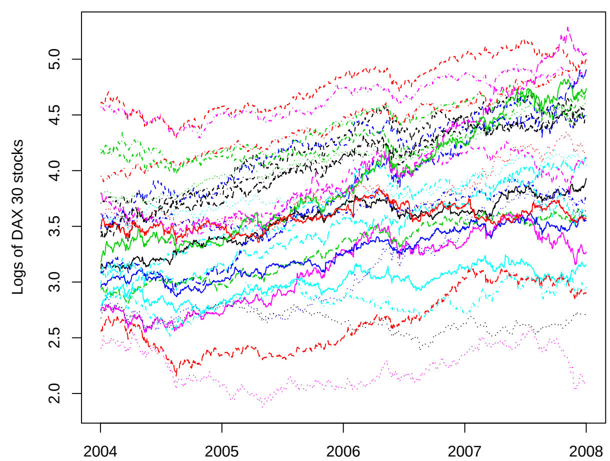

Figure 1: Logs of 30 DAX stock prices, 1/1/2004 to 12/31/2007

Some of the log-levels seem to exhibit quite strong deterministic trends, although the variation around these appears to be very persistent. The returns series have extreme conditional heteroskedasticity, almost IGARCH (when fitting individual GARCH(1,1) processes to single returns series, the sum of the respective GARCH parameters is above 0.9, but below 1, for each of the 30 stocks). While this implies finite variance of the returns, it definitely does not represent evidence for finite kurtosis. To assess the degree of cross-unit dependence, two estimates of the correlation matrix of the stock returns are examined, one computed with the Deutsche Postbank (but covering only 901 trading days) and one computed without the Deutsche Postbank (but with all time observations). For the first correlation matrix, the 435 cross-correlations range between 0.1524 and 0.7463, with a median of 0.3439 and a mean of 0.3565. The first and third quartiles are 0.2871 and 0.4145, respectively, and the semi-interquartile range is 0.0637, indicating a low spread of the correlations and thus some support for a model with constant correlation.[3] The correlation matrix of the returns series exhibits eigenvalues ranging from 11.62 to 0.2091, the second largest being 1.3004. For the second correlation matrix, the figures are practically the same (relative differences are at most 0.05). To sum up, there is strong evidence in favor of a one-factor correlation structure. (As would be implied by the CAPM, by the way.) The evidence in favor of equal weights however, leading to constant correlation across the panel, is somewhat weaker. A weak correlation framework, let alone a independent-units one, is inappropriate for the analysis.

In a first step, 30 individual DF tests are computed based on Equation (1). The resulting overall image is given in Table 1. Four individual tests (for BMW, Bayer, E.ON and Linde) reject the unit root null hypothesis at the 5% level. If the stock prices were independent, this would indeed be evidence against the martingale difference hypothesis. But one needs to confront the data with tests at the panel level to make sure the finding is substantiated.

-2.5528 (0.30246)	-2.5805 (0.28942)	-3.0317 (0.12400)
-3.5970 (0.03050)	-3.8021 (0.01675)	-1.2563 (0.89733)
-1.5284 (0.81947)	-2.1920 (0.49304)	-1.7894 (0.70946)
-2.3618 (0.39956)	-3.0032 (0.13175)	-1.6159 (0.78630)
-2.1183 (0.53443)	-4.0310 (0.00812)	-2.7648 (0.21084)
-2.9252 (0.15483)	-0.9311 (0.95074)	-1.8359 (0.68660)
-4.4523 (0.00185)	-2.0254 (0.58614)	-2.6629 (0.25246)
-2.6141 (0.27404)	-2.9855 (0.13676)	-3.0557 (0.11772)
-3.1432 (0.09689)	-2.6490 (0.25853)	-2.3765 (0.39170)
-2.3275 (0.41802)	-2.6228 (0.27012)	-2.8177 (0.19109)

Table 1: Individual DF test statistics (p values after MacKinnon, 1996)

To this end, the robust panel unit root tests named in Section 2 are used. For Chang's test, the function F for generating the nonlinear instruments is $F(u) = c_i\, u \exp\left(-c_i\, |u|\right)$ for each unit, where the scaling constant c_i equals the inverse of the sample standard deviation of the respective returns series. The residual variance for each test regression is estimated under the null $\rho_i = 0$.[4] Equal weights are used when combining p values, see Demetrescu,

[3] Note that the variance of the sample correlations may not be finite, as the kurtosis of the stock returns may itself be infinite.

[4] This is necessary since the instrumental variable estimate of ϕ does not converge fast

Hassler and Tarcolea (2006) for details. Since the χ^2-test requires a balanced panel, we report two test statistics, one without the Postbank, and one with all stocks, but with only 901 time observations; both are computed with panel-wide White standard errors. The results are given in Table 2.

Test	Value	Crit.val. [5%]
S_N	0.1798	-1.645
$t(\hat{\rho}^*, 0.2)$	-2.4015*	-1.645
χ^2	52.231*/44.792*	43.772
$\widetilde{\kappa}$	-0.1049*	0.0366

Table 2: Results of panel tests (* denotes significance at the 5% level)

The one test not rejecting at the 5% level is the nonlinear IV test of Chang (2002). This can be explained by its low power against local alternatives, see the Appendix. All remaining tests reject at the 5% level, although the significance of χ^2 is lower when computed for the time span common to all stocks. The test statistic $\widetilde{\kappa}$ is even negative, which is asymptotically impossible under the null, but likely under the alternative. This is consistent with the findings of McPherson and Palardy (2007), who find evidence against the martingale difference hypothesis for the DAX index series.

Checking the robustness of the results, note that there is an apparent simultaneous drop of the level for some series around mid-2006. This drop may point toward a level shift in the series. As well known in the unit root testing literature, see e.g. Perron (1989), this may lead to over-rejection under the null and reduced power under the alternative. But two of the employed tests are robust to such breaks, namely χ^2 and $\widetilde{\kappa}$,[5] both of which do reject the examined null. Also, before finally rejecting the martingale difference hypothesis, note that $\widetilde{\kappa}$ assumes constant correlation, so the strength of its rejecting the null should be seen in this perspective.

4. Concluding remarks

There are several explanations for the rejection of the martingale difference hypothesis that do not invalidate the economic reasoning behind the theory

enough to ensure consistent residuals; see Demetrescu (2009).

[5]They are based on *differences*, and thus not affected by an outlier after differencing.

completely; among others, it can be the case that the martingale difference hypothesis fails only for a limited number of stocks in the panel. Discriminating between this situation and a 'homogenous' rejection of the null might be an interesting venue for further research.

Appendix

Let us study the local power of the nonlinear IV unit root test for a single unit. To ease notation, drop the index i. To simplify the argument, assume the increments ε_t to follow an *iid* standard normal process and ignore the deterministic components. Under the local alternative $\rho = 1 - c/T$, we have

$$\Delta z_t = -\frac{c}{T} z_{t-1} + \varepsilon_t.$$

Since, under the simplified assumptions, $\left(T\sigma^2\right)^{-0.5} \sum_{j=0}^{t} \varepsilon_j$ is a standard Wiener process discretized on an equally spaced grid t/T, $\left(T\sigma^2\right)^{-0.5} z_t$ will be the discretized solution $X(s)$ of the stochastic differential equation

$$dX(s) = -cX(s)ds + dW(s),$$

i.e. a standard Ornstein-Uhlenbeck process. The single-unit test statistic is

$$t_\phi^{IV} = \frac{\sum_{t=2}^{T} F(z_{t-1})\,\varepsilon_t}{\widehat{\sigma}\sqrt{\sum_{t=2}^{T} F^2(z_{t-1})}} - \frac{c}{T}\frac{\sum_{t=2}^{T} G(z_{t-1})}{\widehat{\sigma}\sqrt{\sum_{t=2}^{T} F^2(z_{t-1})}},$$

where $G(u) = F(u) \cdot u$. The function G is integrable, bounded and continuous, as is F^2, so $\sum G(z_{t-1})$ has the same behavior as $\sum F^2(z_{t-1})$. But Jacod (1998, Section 2) shows that diffusion processes (of which the Ornstein-Uhlenbeck process is one) behave essentially like the Wiener process with respect to such sums. Hence, $\sum G(z_{t-1}) = O_p(T^{0.5})$ and $\sum F^2(z_{t-1}) = O_p(T^{0.5})$, since $\sum F^2\left(\sum_{j=0}^{t} \varepsilon_j\right) = O_p(T^{0.5})$ follows e.g. from Demetrescu (2009, Lemma 1 *a)*). Thus, t_ϕ^{IV} behaves under the local alternative as follows

$$t_\phi^{IV} \stackrel{d}{=} \frac{\sum_{t=2}^{T} F(z_{t-1})\,\varepsilon_t}{\widehat{\sigma}\sqrt{\sum_{t=2}^{T} F^2(z_{t-1})}},$$

with '$\stackrel{d}{=}$' standing for equivalence in distribution. It can be shown using the arguments given in Chang (2002, Section 4) that the quotient on the right-hand side converges to a standard normal distribution. The result follows.

References

Bai, J. and S. Ng (2004) "A PANIC Attack on Unit Roots and Cointegration" *Econometrica* **72**, 1127-1177.

Baillie, R.T. and T. Bollerslev (1989) "Common Stochastic Trends in a System of Exchange Rates" *Journal of Finance* **44**, 167-181.

Breitung, J. and S. Das (2005) "Panel Unit Root Tests Under Cross Sectional Dependence" *Statistica Neerlandica* **59**, 1-20.

Breitung, J. and M.H. Pesaran (2008) "Unit Roots and Cointegration in Panels" in *The Econometrics of Panel Data: Fundamentals and Recent Developments in Theory and Practice* by Matyas, L. and P. Sevestre, Eds., Kluwer Academic Publishers, 279-322.

Chang, Y. (2002) "Nonlinear IV Unit Root Tests in Panels with Cross-Sectional Dependency" *Journal of Econometrics* **110**, 261-292.

Demetrescu, M. (2009) "Panel Unit Root Testing with Nonlinear Instruments for $AR(\infty)$ Processes" *Mimeo.*

Demetrescu, M., U. Hassler and A.I. Tarcolea (2006) "Combining Significance of Correlated Statistics with Application to Panel Data" *Oxford Bulletin of Economics and Statistics* **68**, 647-663.

Demetrescu, M., V. Kuzin and U. Hassler (2008) "Long Memory Testing in the Time Domain" *Econometric Theory* **24**, 176-215.

Gil-Alana, L.A. (2006) "Fractional Integration in Daily Stock Market Indexes" *Review of Financial Economics* **15**, 28-48.

Granger, C.W.J. and R. Joyeux (1980) "An Introduction to Long Memory Time Series Models and Fractional Differencing" *Journal of Time Series Analysis* **1**, 15-39.

Hassler, U., M. Demetrescu and A.I. Tarcolea (2007) "Asymptotically χ^2 Unit Root Testing in Small Cross-Dependent Panels" *Mimeo.*

Hassler, U. and J. Wolters (1994) "On the Power of Unit Root Tests against Fractionally Integrated Alternatives" *Economics Letters* **45**, 1-5.

Horowitz, J.L., I.N. Lobato, J.C. Nankervis, and N.E. Savin (2006) "Bootstrapping the BoxPierce Q test: A robust test of uncorrelatedness" *Journal of Econometrics* **133**, 841-862.

Jacod, J. (1998) "Rates of Convergence to the Local Time of a Diffusion" *Annales de l'Institut Henri Poincare (B) Probability and Statistics* **34**, 505-544.

Krämer, W. (1998) "Fractional Integration and the Augmented Dickey-Fuller Test" *Economics Letters* **61**, 269-272.

Krämer, W., P. Sibbertsen and C. Kleiber (2002) "Long memory versus structural change in financial time series" *Allgemeines Statistisches Archiv* **86**, 83-96.

MacKinnon, J.G. (1996) "Numerical distribution functions for unit root and cointegration tests" *Journal of Applied Econometrics* **11**, 601-618.

McPherson, M.Q. and J. Palardy (2007) "Are International Stock Returns Predictable? An examination of linear and non-linear predictability using generalized spectral tests" *Journal of International Financial Markets, Institutions & Money* **17**, 452-464.

Meese, R.A. and K.J. Singleton (1982) "On Unit Roots and the Empirical Modelling of Exchange Rates" *Journal of Finance* **37**, 1029-1034.

Perron, P. (1989) "The Great Crash, the Oil Price Shock, and the Unit Root Hypothesis" *Econometrica* **57**, 1361-1401.

Phillips, P.C.B. (1987) "Time Series Regression with a Unit Root" *Econometrica* **55**, 277-301.

Tarcolea, A.I., U. Hassler and M. Demetrescu (2009) "Testing for Stationarity in Large Panels with Cross-Dependence, and US Evidence on Unit Labor Cost" forthcoming *Journal of Applied Statistics*.

Westerlund, J. and R. Larsson (2007) "A Note on the Pooling of Individual PANIC Unit Root Tests" forthcoming *Econometric Theory*.