Journal of Empirical Finance 16 (2009) 734–744

Contents lists available at ScienceDirect

Journal of Empirical Finance

journal homepage: www.elsevier.com/locate/jempfin

On the usefulness of the contrarian strategy across national stock markets: A grid bootstrap analysis ☆

Hyeongwoo Kim *

Department of Economics, Auburn University, Auburn, AL 36849, United States

ARTICLE INFO

Article history:
Received 10 January 2008
Received in revised form 13 June 2009
Accepted 16 June 2009
Available online 26 June 2009

JEL classification:
G10
G14
G15

Keywords:
Stock index price deviation
Median-unbiased estimator
Nonparametric grid bootstrap
Cross-section dependence

ABSTRACT

This paper statistically evaluates the usefulness of the contrarian investment strategy across the national stock markets of 18 developed countries. The contrarian strategy implicitly assumes that asset prices tend toward a fundamental value path over time. Conventional bootstrap analyses and panel unit root tests are often consistent with such a hypothesis. However, these results might be contaminated by small-sample bias and/or by not controlling cross-section dependence. Correcting for small-sample bias nonparametrically, I find extremely slow mean reversion rates, which provide strong evidence against the usefulness of the contrarian strategy.

1. Introduction

One implicit assumption of the contrarian investment strategy (DeBondt and Thaler, 1985) is the mean reversion property of asset prices toward a fundamental value path. That is, when asset prices are mean-reverting, one may obtain excess returns by short-selling assets that have performed well and buying assets with relatively poor past performance.

Empirical evidence on mean reversion in US stock prices is mixed. Fama and French (1988) found significantly negative autocorrelation coefficients for long-horizon real US stock returns. Poterba and Summers (1988) found that a substantial part of the variance of the US stock returns is due to a transitory component. Many other researchers, however, question the validity of the mean reversion hypothesis. Kim et al. (1991) report very weak evidence of mean reversion in the post-war era. Richardson and Stock (1989) and Richardson (1993) point out that the previous empirical evidence in favor of mean reversion might be spurious, with the results caused by not controlling for small-sample bias. McQueen (1992) also shows that mean reversion results are not robust to distributional assumptions.

An array of researchers has also investigated mean reversion in the context of the international stock markets, again producing mixed evidence of mean reversion. Kasa (1992) argues that national stock indices for 5 industrialized countries are cointegrated around a common world component, which implies short-lived excess returns for these countries. Richards (1995), however, finds no evidence of cointegration using stock index data for 16 OECD countries.[1] Balvers et al. (2000) implemented a Levin et al. (2002) type panel unit root test for 18 countries with well-developed stock markets, and reported strong evidence in favor of mean reversion. More recently, Bhojraj and Swaminathan (2006) found some evidence in favor of the long-run contrarian strategy using stock index data for 38 countries that included both developed and less developed economies.

☆ I thank Randy Beard, Masao Ogaki, Young-Kyu Moh, Donggyu Sul, Michael Stern, and in particular, two anonymous referees for their helpful comments and suggestions.
* Tel.: +1 334 844 2928; fax: +1 334 844 4615.
E-mail address: gmmkim@gmail.com.
[1] Notwithstanding such results, he provided some evidence for predictability of stock returns relative to a world index from the same data set.

0927-5398/$ – see front matter
doi:10.1016/j.jempfin.2009.06.005

Given such mixed evidence, I attempt here to shift the focus from the mean reversion tests to a more practical issue. That is, this paper statistically evaluates the usefulness of the contrarian investment strategy using a nonparametric grid bootstrap analysis. I am especially interested in measuring how fast stock price deviations from fundamental value disappear, because findings of persistent deviations would imply that the practical usefulness of the contrarian strategy is limited.

Assuming that convergence rates are the same for all countries, Balvers et al. (2000) report 3 to 3 and half year half-life estimates for deviations after correcting for median-bias by Andrews' (1993) method.[2] They also reported very compact confidence intervals that range around 2 to 5 years.

Their results, however, are subject to potentially serious problems. First, they assume that deviations from fundamentals are caused solely by idiosyncratic shocks. Put it differently, they employ an assumption of cross-section independence in estimating half-lives of stock price deviations, which can be too strong in empirical work dealing with macroeconomic panels. Recently, new panel unit root tests that utilize common factor structures to take into account of the cross-section dependence have been put forward, among others, Pesaran (2007), Bai and Ng (2004), Moon and Perron (2004), and Phillips and Sul (2003). One related finding is that panel unit root tests with *no* cross-section dependence consideration suffer from serious size distortion (Phillips and Sul 2003), which casts doubt on the results of Balvers et al. (2000).

Second, they neglect the fact that the median-bias is *random* in the presence of nuisance parameters. When one is dealing with higher order autoregressive, AR(p), processes, the median-bias should be estimated for each realization of other coefficients (nuisance parameters).[3] Using bias estimates from an AR(1) specification, therefore, yields inaccurate correction.[4]

In this paper, I employ Hansen's (1999) grid bootstrap technique for bias-correction in higher order AR processes. His method does not require any distributional assumption such as Gaussianity, but provides excellent first-order asymptotic coverage properties.[5] For panel Seemingly Unrelated Regression (SUR) estimation following Balvers et al. (2000), I relax the assumption of cross-section independence as I find strong evidence of cross-section dependence.

Using the Morgan Stanley Capital International (MSCI) stock index data for 18 developed countries from December of 1969 to September of 2007, I estimate half-lives of national stock price deviations from fundamental values in univariate and in panel regression models. From univariate regressions, I obtain a 3- to 4-year median half-life estimate with the conventional bootstrap confidence intervals ranging around 1 to 8 years. When I nonparametrically correct for small-sample bias, however, some of the half-life point estimates become infinity. More surprisingly, I find finite confidence intervals for only a few countries. This result seriously casts doubt on the usefulness of the contrarian investment strategy, because an experiment that demonstrates the contrarian strategy outperforms others might be a singular event. When I implement panel-SUR estimations with cross-section dependence consideration, I find much longer half-life estimates compared with those of Balvers et al. (2000) especially when the US serves as a reference country. This again provides strong empirical evidence against the use of the contrarian strategy.

The remainder of the paper is organized as follows. In Section 2, I describe my estimation models of asset prices. Section 3 discusses the bias-correction methods for point estimates and confidence intervals based on Hansen's (1999) work. Section 4 describes the data and provides some preliminary statistics. In Section 5, I report my empirical findings. Section 6 concludes.

2. The econometric model

2.1. Univariate estimation

I start with a model of a stochastic process for national stock indices that is similar to the one by Balvers et al. (2000).

Let p_t^i and f_t^i be the log of the stock index and the log of its fundamental value for country i, respectively. When p_t^i is mean-reverting around f_t^i, its deviation should die out eventually, which implies the following error correction model.[6]

$$\Delta\left(p_{t+1}^i - f_{t+1}^i\right) = a^i - \lambda^i\left(p_t^i - f_t^i\right) + \varepsilon_{t+1}^i, \tag{1}$$

where $0<\lambda^i<1$ and ε_t^i is a mean-zero serially uncorrelated stochastic process from an *unknown* distribution. That is, the deviation of the country i's stock index away from its fundamental value, $p_t^i - f_t^i$, converges to its long-run equilibrium with the country-specific convergence rate λ^i.

It is interesting to compare our model (1) with the one by Balvers et al. (2000). They claim that the stochastic process of a mean-reverting (weakly stationary) asset price can be represented as,

$$\Delta p_{t+1}^i = a^i - \lambda^i\left(p_t^i - f_{t+1}^i\right) + \varepsilon_{t+1}^i, \tag{2}$$

which is motivated by forecasting models of stock returns (see, for example, Cochrane, 2008).[7] Unlike our error correction model, this equation does not require any cointegrating relation. However, it is hard to interpret the

[2] The term "half-life" refers to the time period sufficient for the deviations to decay to one-half.
[3] Bias-correction methods proposed by Andrews and Chen (1994) or Hansen (1999) are applicable to AR(p) processes.
[4] See their appendix (pp.769–770) for details.
[5] If we assume that errors are i.i.d. Gaussian, Hansen's method coincides with the method by Andrews and Chen (1994).
[6] If p_t^i and f_t^i are integrated processes, this assumption implicitly means that they are cointegrated with a known cointegrating vector $[1 \ -1]$.
[7] See their Eq. (1) on p.748.

stochastic property of Eq. (2), because the long-run equilibrium value of p^i is time-varying (f^i_{t+1}). That is, in their model, the stock index price is *not* weakly stationary (see, for example, Hamilton, 1994).[8] Thus, I conduct analysis with Eq. (1) rather than Eq. (2).

The fundamental value f^i_t is not directly observable, but is assumed to obey the following stochastic process:

$$f^i_t = c^i + p^w_t + \upsilon^i_t, \tag{3}$$

where c^i is a country-specific constant, p^w_t denotes a reference stock index price, and υ^i_t is a zero-mean, possibly *serially correlated* stationary process from an *unknown* distribution.

From Eqs. (1) and (3), I obtain the following equation.

$$\Delta\left(p^i_{t+1} - p^w_{t+1}\right) = \alpha^i - \lambda^i\left(p^i_t - p^w_t\right) + \eta^i_{t+1}, \tag{4}$$

where $\alpha^i = a^i + \lambda^i c^i$ and $\eta^i_{t+1} = \upsilon^i_{t+1} - (1-\lambda^i)\upsilon^i_t + \varepsilon^i_{t+1}$. Note that η_{t+1} is serially correlated even when υ_t is a white noise process.

In order to control for the serial correlation in η_{t+1}, I add lagged values of the excess return to Eq. (4), and obtain:

$$\Delta\left(p^i_{t+1} - p^w_{t+1}\right) = \alpha^i - \lambda^i\left(p^i_t - p^w_t\right) + \sum_{j=1}^{k} \beta_j \Delta\left(p^i_{t-j+1} - p^w_{t-j+1}\right) + \eta^i_{t+1} \tag{5}$$

or equivalently,

$$r^i_t = \alpha^i + \rho^i r^i_{t-1} + \sum_{j=1}^{k} \beta^i_j \Delta r^i_{t-j} + \eta^i_t, \tag{6}$$

where $r^i_t = p^i_t - p^w_t$ denotes the stock price deviation for country *i*. Note that $\rho^i = 1 - \lambda^i$ and $\rho^i \in (0, 1)$ is the persistence parameter of the stock index deviation for country *i*.

It is straightforward to verify that Eq. (6) is equivalent to Balvers et al.'s (2000) Eq. (4) (p.750). It should be noted, however, Eq. (6) does not require the homogeneity assumption for convergence rate λ.[9] Furthermore, I do not need to impose any distributional assumption on η_t.[10]

2.2. Panel seemingly unrelated estimation

I propose the following model for my panel-SUR estimation with cross-section dependence.

$$r^i_t = \alpha^i + \delta^i \theta_t + \rho r^i_{t-1} + \sum_{j=1}^{k} \beta^i_j \Delta r^i_{t-j} + \eta^i_t, \tag{7}$$

where δ^i denotes a country-specific factor loading coefficient for the country *i* and θ_t is the common factor that is orthogonal to η^i_t.[11] Note that I impose the same homogeneity assumption, $\rho^i = \rho$, $\forall i = 1, 2, ..., N$, as in Balvers et al. (2000) to compare my results with theirs. Given δ^i and θ_t estimates, Eq. (7) can be represented as,

$$\hat{r}^i_t = \alpha^i + \rho \hat{r}^i_{t-1} + \sum_{j=1}^{k} \beta^i_j \Delta \hat{r}^i_{t-j} + \eta^i_t, \tag{8}$$

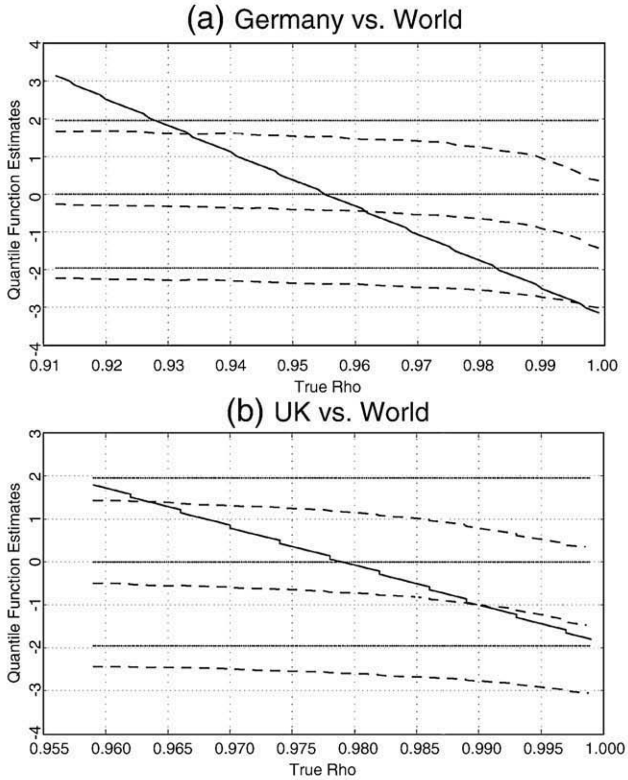
where $\hat{r}^i_t = r^i_t - \hat{\delta}^i \hat{\theta}_t$ denotes a common (global) shock detrended excess stock price of the country *i*. In other words, ρ can be interpreted as a persistence parameter to idiosyncratic shocks. Note that the present model nests a common time effect model as a special case of $\delta^i = 1$, $\forall i = 1, 2, \cdots, N$. Then, the SUR estimation of Eq. (8) can be implemented as described by Levin et al. (2002).

[8] Covariance (weakly) stationarity requires constant first and second moments for all *t*.

[9] In order to derive Eq. (3) in Balvers et al. (2000) from (1), one has to assume $\lambda^i = \lambda^r$ where *r* refers to the reference country. Otherwise, the unobserved term P^{*i}_{t+1} in their Eq. (1) cannot be cancelled out and remains in their estimation equation.

[10] They use Andrew's (1993) method to calculate the median-unbiased estimates and the corresponding confidence intervals, which requires Gaussian error terms.

[11] When θ_t is serially correlated, additional lagged factor terms should be added.



Notes: i) Sample periods are 1969:12-2007:9 ($T = 454$). ii) 2.5%, 50%, and 97.5% quantile function estimates (dashed lines) and the grid-*t* statistics (solid line) were obtained from 10,000 nonparametric bootstrap replications from the empirical distribution on 50 grid points in the neighborhood of the least squares point estimates (Hansen 1999).

Fig. 1. Median-unbiased estimator and the grid-*t* confidence interval.

Note also that one can assess the relative importance of common shocks to the variation of r_t^i by the following variance decomposition. That is, conditional on all coefficients and lagged values of r_t^is, the following holds.[12]

$$Var\left(r_t^i\right) = Var\left(\eta_t^i\right) + Var\left(\delta^i \theta_t\right), \tag{9}$$

owing to the orthogonality assumption of θ_t to idiosyncratic shock η^i_t.

3. The median-unbiased estimator and the grid bootstrap confidence interval

It is well-known that the least squares estimator for ρ^i in Eq. (6) is significantly downward-biased when the regression equation for autoregressive processes includes an intercept (α^i).[13] It should also be noted that the distribution of η_t is, in general, unknown and far from being a normal distribution. This implies that the conventional normal approximation for constructing confidence intervals performs very poorly, and inherits such downward bias. The extent of this so-called small-sample bias becomes even more severe as ρ^i approaches to unity.

When the distribution is unknown, researchers often rely on bootstrap procedures. For example, one may construct percentile bootstrap or bootstrap-*t* confidence intervals (Efron and Tibshirani 1993). However, these methods evaluate the empirical distribution of the parameter estimates only on the least squares point estimate. This implicitly assumes that the bootstrap quantile functions are constant around the point estimate, which is *not* correct for the autoregressive process.

Fig. 1 clearly shows such downward bias. Detailed explanations of the figure will be presented in what follows. The figure plots the theoretical values for the 2.5%, 50%, and 97.5% quantile functions of the *t*-statistics for ρ^i based on the normal approximation (dotted lines), which are −1.96, 0.00, and 1.96, respectively. The corresponding quantile function estimates (dashed lines),

[12] I thank an anonymous referee for pointing this out.

[13] The least squares regression of an autoregressive process with a constant is equivalent to running regressions for demeaned observations. Then, the regression error is correlated with the independent variable, because the current and future values of the dependent variable are embedded in the sample mean. It can be shown that this creates a downward bias (see, among others, Kendall, 1954 for univariate estimations and Nickell, 1981 for panel estimations).

however, are far from being constant, and are below their normal approximation values. Such downward bias becomes more severe as the true value of ρ^i approaches to unity.

These issues have been resolved by Andrews (1993) for the first-order autoregressive process. By evaluating empirical distributions of parameter estimates over a grid of values in the relevant parameter space for ρ^i, he shows how to obtain the exactly median-unbiased estimates as well as exact confidence intervals. Andrews and Chen (1994) extended this work for the higher order autoregressive process such as Eq. (6), and showed how to obtain approximately median-unbiased estimates and confidence intervals.

Balvers et al. (2000) applied Andrews' (1993) method for a univariate AR(1) process to the panel context and corrected their estimates for small-sample bias. It should be noted that they corrected for the bias using Andrews' (1993) method rather than using Andrews and Chen's (1994) method, though their estimation model is a higher order autoregressive process. This is problematic because median-bias in their AR(p) model becomes *random* in the presence of nuisance parameters such as βs. Put it differently, bias should be measured for realized nuisance parameter values conditional on the true value for ρ^i. This implies their bias-corrections based on an AR(1) representation would be inaccurate.

It should be also noted that these methods require Gaussianity assumption for η_t^i, which may not be appropriate for asset prices such as stock indices. Hansen (1999) proposes a nonparametric grid bootstrap method that significantly improves over the conventional methods. His method does not require any distributional assumption on error terms, but provides excellent first-order coverage even when ρ^i is unity.

The median-unbiased estimate $\tilde{\rho}^i$ and its corresponding confidence interval can be obtained as follows. Following Hansen (1999), I define the grid-t statistic at each of M grid points $\rho_j^i \in [\rho_1^i, \rho_2^i, \cdots \rho_M^i]$ around the neighborhood of the least square point estimate $\hat{\rho}^i$.

$$t_N\left(\rho_j^i\right) = \frac{\hat{\rho}^i - \rho_j^i}{se\left(\hat{\rho}^i\right)}, \tag{10}$$

where $se(\hat{\rho}^i)$ denotes the least squares standard error of $\hat{\rho}^i$.

Then, for each grid point ρ_j^i, I generate pseudo samples for each country nonparametrically. Implementing least square estimations for B bootstrap iterations at each of M grid points of ρ_j^i, we obtain the (α quantile) grid-t bootstrap quantile function estimates, $\hat{q}^*_{N,\alpha}(\rho_j^i) = \hat{q}^*_{N,\alpha}(\rho_j^i, \varphi(\rho_j^i))$, where φ denotes nuisance parameters such as βs that are functions of ρ_j^i. Note that each function is evaluated at each grid point ρ_j^i rather than at the point estimate.[14] Then, I smooth the estimated quantile functions using kernel regression to obtain smoothed estimate $\tilde{q}^*_{N,\alpha}(\rho_j^i)$ for each grid point ρ_j^i.[15]

Finally, the median-unbiased estimator is,

$$\tilde{\rho}^i = \rho_j^i \in R, \text{ s.t. } t_N\left(\rho_j^i\right) = \tilde{q}^*_{N,50\%}\left(\rho_j^i\right), \tag{11}$$

and the corresponding bias-corrected 95% grid-t confidence interval $[\tilde{\rho}_L^i, \tilde{\rho}_U^i]$ is as follows.[16]

$$\tilde{\rho}_L^i = \rho_j^i \in R \text{ s.t. } t_N\left(\rho_j^i\right) = \tilde{q}^*_{N,97.5\%}\left(\rho_j^i\right) \text{ and } \tilde{\rho}_U^i = \rho_j^i \in R \text{ s.t. } t_N\left(\rho_j^i\right) = \tilde{q}^*_{N,2.5\%}\left(\rho_j^i\right). \tag{12}$$

Corresponding half-life estimates are obtained by the conventional method using the formula $\ln(1/2)/\ln(\rho^i)$.[17] For the half-life in years, one needs to divide the number by the frequency count of the data in one year (e.g., 12 with monthly frequency).

Bias-correction for panel-SUR estimates can be similarly implemented by matching bootstrap quantile function estimates for SUR estimate ρ with the corresponding grid-t statistic.

4. Data and summary statistics

I utilize two data sets. For univariate estimations of speeds of convergence, I employ monthly frequency stock indices from Morgan Stanley Capital International (MSCI) for stock market indices for 18 developed countries and the MSCI World index. For panel-SUR estimations, annual frequency data is used to compare my results with those of Balvers et al. (2000). The observations span from December of 1969 to September of 2007. The observations are end-of-period value-weighted index prices in US dollar terms that include gross dividends.[18]

I provide preliminary summary statistics of national stock index deviations from the reference index in Table 1. Following Balvers et al. (2000), I choose the World index or the US index as the reference index. The mean values of the index deviations

[14] When they are evaluated at the point estimate, the estimator reduces to the Efron and Tibshirani's (1993) bootstrap-t estimator.

[15] Following Hansen (1999), I used the Epanechinikov kernel $K(u) = 3(1-u^2)/4I(|u| \leq 1)$, where $I(\cdot)$ is an indicator function. The bandwidth parameter was chosen by least squares leave-one-out cross-validation.

[16] Hansen (1999) did not explicitly describe how to obtain the median-unbiased estimator using his grid-bootstrap method. However, it is straightforward to use the 50% quantile function estimates to get the median-unbiased estimator in line of Andrews (1993). Steinsson (2008) and Kim and Ogaki (2009) also employ similar approaches.

[17] This method implicitly assumes that the price deviation monotonically decays. In order to calculate half-life allowing for non-monotonic convergence, one can use the impulse-response functions. However, the difference between these two methods is often negligible.

[18] When one uses an average data set such as the IFS share price index, the ρ^i estimate is subject to an upward time aggregation bias (see Taylor, 2001). We avoid such a bias using end-of-period data.

Table 1
Summary statistics of national stock index deviations from the reference index.

Country	World reference index				US reference index			
	Mean	S.D.	JBT	JBT_A	Mean	S.D.	JBT	JBT_A
Australia	−0.571	0.254	1.172	0.282	−0.476	0.320	51.65**	4.834**
Austria	0.132	0.412	15.74**	1.193	0.227	0.512	31.08**	2.658
Belgium	0.587	0.357	12.17**	1.533	0.682	0.352	24.80**	2.285
Canada	−0.115	0.305	29.88**	2.688	−0.020	0.366	18.04**	1.994
Denmark	0.549	0.291	9.559**	0.887	0.644	0.289	68.92**	5.644**
France	0.072	0.217	19.22**	1.750	0.167	0.229	14.23**	1.512
Germany	0.040	0.167	8.573**	0.955	0.135	0.241	10.67**	1.420
Hong Kong	1.437	0.474	15.59**	1.468	1.532	0.469	31.22**	4.496*
Italy	−1.017	0.359	116.3**	6.944**	−0.923	0.419	32.77**	3.102*
Japan	0.804	0.490	20.49**	1.578	0.899	0.664	17.39**	1.334
Netherlands	0.580	0.415	33.01**	2.712	0.675	0.377	43.28**	4.137*
Norway	0.353	0.307	14.32**	2.650	0.447	0.395	6.523**	0.152
Singapore	0.691	0.442	13.15**	1.687	0.786	0.550	18.33**	1.666
Spain	−0.420	0.466	11.88**	0.556	−0.325	0.440	25.91**	1.424
Sweden	0.639	0.485	19.90**	1.781	0.734	0.444	18.62**	1.818
Switzerland	0.260	0.273	27.83**	2.588	0.354	0.215	4.075	1.061
UK	0.178	0.199	129.8**	16.57**	0.273	0.230	42.37**	5.159**
US	−0.095	0.185	10.97**	1.756				

Notes: i) Sample periods are 1969:12–2007:9 (T = 454) and 1969–2006 (T = 38 and N_{US} = 17 and N_{World} = 18) for JBT and JBT_A, respectively ii) JBT and JBT_A refer to the normality test statistics by Jarque and Bera (1980) for the monthly and annual data, respectively. The null hypothesis is that each stock index deviation is normally distributed. The critical values were taken from Deb and Sefton (1996) to deal with a size distortion problem using an asymptotic chi-square distribution with two degrees of freedom. iii) * and ** refer to the rejections of the null hypothesis of normality at the 10% and 5% significance level, respectively.

relative to the World index range from −1.017 for Italy to 1.437 for Hong Kong, and the standard deviations vary from 0.167 for Germany to 0.490 for Japan. The statistics for the stock index deviations relative to the US index exhibit similar patterns.

I also check for normality of error terms with the conventional Jarque–Bera test. It is well-known that the test suffers from a serious size distortion problem when an asymptotic $\chi^2(2)$ distribution is used. So I use the critical values from Deb and Sefton (1996) to deal with the oversize problem. Interestingly, I found strong evidence against the normality assumption from monthly relative stock prices. The tests accept the null hypothesis of normality at 10% significance level only for 3 countries with the US index, and for 6 countries with the World index. With the annual frequency data, I obtained much more favorable results for Gaussianity.[19] Since I obtained mixed results, I plotted empirical distributions for annual data using the Gaussian kernel and report them in Fig. 2. With a few exceptions, most empirical distributions seem far from being normal. Based on this finding, I employ Hansen's (1999) nonparametric bias-correction method that is robust to the distributional assumptions.

As an additional preliminary test, I implement a panel unit root test proposed by Im et al. (2003). Their test procedure is much less restrictive than the Levin et al. (2002) type panel test, since Im et al.'s (2003) test does not require any homogeneity assumption on the convergence (and thus persistence) parameter. That is, ρ^i in Eq. (6) need *not* be identical for all countries.

As in Balvers et al. (2000), I also obtain strong evidence that favors mean reversion (see Table 2). It is very important to recognize that this does *not* mean that all stock indices are stationary, because most panel unit root tests that include the Levin et al. (2002) type test and Im et al.'s (2003) test reject the null of a unit root when even a single index is stationary.[20]

More importantly, these tests employ a cross-section independence assumption. As Phillips and Sul (2003) point out, panel unit root tests with no cross-section dependence consideration are over-sized.[21] Put it differently, those tests tend to reject the null of unit root too often. Recognizing this, I implement a meta-z test proposed by Choi (2001) and Phillips and Sul (2003) for my panel data after controlling for cross-section dependence.[22] It is interesting to see that the null of unit root cannot be rejected irrespective of the choice of reference indices. Therefore, controlling for cross-section dependence produces much weaker evidence for mean reversion of stock index deviations. In other words, I found much weaker evidence for the usefulness of the contrarian investment strategy.

5. Empirical results

Assuming ρ^i is common for all countries, Balvers et al. (2000) implemented a Seemingly Unrelated Regression type unit root test[23] for stock index deviations, and found strong evidence in favor of mean reversion. Our panel unit root test results are consistent with theirs only when a cross-section independence assumption is employed. Allowing for cross-section dependence, however, the evidence of mean reversion becomes much weaker. Given mixed evidence, I shift the focus to a more practical issue and measure how quickly individual stock price deviations converge to their equilibrium values.

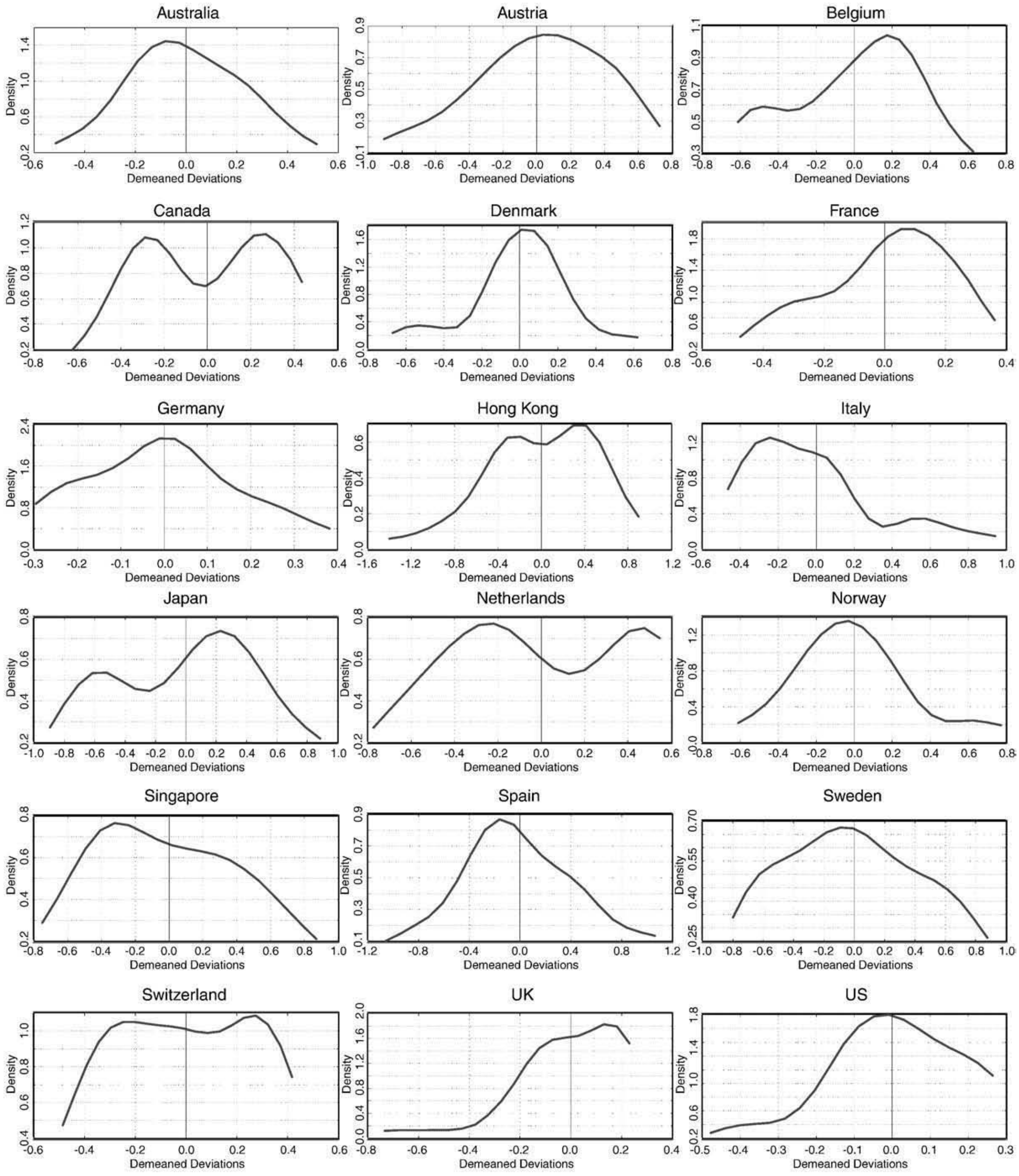
[19] I thank an anonymous referee who pointed out this possibility.

[20] The other extreme is the panel unit root test by Taylor and Sarno (1998), which accepts the unit root hypothesis when at least one series is nonstationary.

[21] O'Connell (1998) also reports that the limiting distribution of panel unit root test statistics is no longer correct when disturbances are correlated. Controlling for cross-section dependence, he finds no evidence for purchasing power parity.

[22] I acknowledge the use of Sul's code available at http://homes.eco.auckland.ac.nz/dsul013/gauss/mf.htm.

[23] Their test procedures are based on the formal test by the earlier version of the Levin et al. (2002) test. However, they use empirical critical values with no asymptotic theory.



Notes: i) Sample periods are 1969-2006 ($T = 38$). ii) The Gaussian kernel with optimal bandwidth selection method was used. Using alternative kernels such as the Epanechnikov kernel produces similar shapes.

Fig. 2. Kernel estimation of the density function for the international stock index deviations from the World index.

Unlike panel-SUR models as in Balvers et al. (2000) or in the present paper, our univariate estimation model does not require any homogeneity assumption so that we are able to measure speeds of convergence for each stock price. In order to preserve the cross-sectional correlation, however, we extract the entire vector of residuals $\{\tilde{\eta}_t^1, \tilde{\eta}_t^2, \cdots, \tilde{\eta}_t^N\}$ in generating pseudo population samples for each country.

Table 2
Panel unit root test results.

	Cross-section dependence	World reference index	US reference index
$\bar{t}_{IPS}$	No	−2.065**	−2.384**
z_{CPS}	Yes	41.88	30.11

Notes: i) The number of lags (k) was chosen by the general-to-specific rule (Hall, 1994) with maximum 2-year lags. ii) Sample periods are 1969–2006 ($T=38$ and $N_{US}=17$ and $N_{World}=18$). iii) $\bar{t}_{IPS}$ denotes the Im et al.'s (2003) test statistic and z_{CPS} is the z-test statistics proposed by Choi (2001) and Phillips and Sul (2003). iv) Critical values were obtained by interpolating the values from Mark (2001) and from Sul's website. v) ** refers to the rejection of the null hypothesis of common unit root at the 5% significance level. vi) The p values for z_{CPS} were 56.3 and 16.6 for the world and the US reference index, respectively. vii) The factor loading coefficient estimates are available from the author upon request.

Table 3
Bias-corrected estimates and grid bootstrap confidence intervals: national stock index deviations from the World index.

Country	$\tilde{\rho}$	CI	HL	CI
Australia	0.982	[0.957,1.000]	3.134	[1.314,∞)
Austria	0.990	[0.975,1.000]	5.476	[2.263,∞)
Belgium	1.000	[0.982,1.000]	∞	[3.165,∞)
Canada	1.000	[0.985,1.000]	∞	[3.804,∞)
Denmark	1.000	[0.973,1.000]	∞	[2.110,∞)
France	0.981	[0.958,1.000]	2.952	[1.340,∞)
Germany	0.962	[0.933,0.996]	1.477	[0.836,3.54]
Hong Kong	0.983	[0.963,1.000]	3.448	[1.511,∞)
Italy	0.972	[0.953,0.993]	2.028	[1.199,7.921]
Japan	1.000	[0.990,1.000]	∞	[5.753,∞)
Netherlands	1.000	[0.995,1.000]	∞	[11.52,∞)
Norway	0.976	[0.955,1.000]	2.387	[1.250,∞)
Singapore	0.985	[0.968,1.000]	3.728	[1.792,∞)
Spain	0.997	[0.983,1.000]	19.57	[3.431,∞)
Sweden	1.000	[0.992,1.000]	∞	[7.492,∞)
Switzerland	1.000	[0.981,1.000]	∞	[3.053,∞)
UK	0.990	[0.964,1.000]	5.809	[1.555,∞)
US	0.993	[0.980,1.000]	8.472	[2.888,∞)
Median	0.992	[0.974,1.000]	7.141	[2.187,∞)

Notes: i) The number of lags (k) was chosen by the general-to-specific rule (Hall, 1994) with maximum 12 month lags. ii) Sample periods are 1969:12–2007:9 ($T=454$). iii) For each stock index deviation, the bias-corrected 95% confidence interval as well as the bias adjusted estimate was obtained by 10,000 nonparametric bootstrap replications from the empirical distribution on 50 grid points in the neighborhood of the least squares point estimates (Hansen, 1999). iv) With the 90% bias-corrected confidence intervals, the corresponding half-life confidence intervals were finite for one additional country, Norway ([2.39,24.7]).

Without small-sample bias-corrections, we obtained compact 95% confidence intervals for ρ^i for each country irrespective of the choice of reference indices.[24] The corresponding median values of half-life point estimates are 4.250 and 2.892 years when the World index and the US index serve as a reference index, respectively. The 95% confidence intervals were also compact in their median values, and ranged from about 1 to 8 years, which is roughly consistent with the work of Balvers et al. (2000). These results seem to clearly support the value of the contrarian strategy.[25]

This finding in favor of the contrarian strategy is completely upset when we correct for the small-sample bias using the grid bootstrap method (See Tables 3 and 4). For each country, we implemented 10,000 bootstrap simulations on 50 fine grid points totaling 500,000 nonparametric simulations. One notable finding is that some of the point estimates become unity, which implies that stock price deviations may not be corrected in finite time. In particular, the bias-corrected point estimates $\tilde{\rho}^i$ for Netherlands and Sweden were always unity irrespective of the reference indices.

The most surprising finding is that for most countries, the 95% confidence intervals include unity, which implies that it may take infinite time for the price indices to halfway adjust to their long-run equilibrium values. This conclusion remains valid even with the 90% confidence intervals, as the confidence intervals become finite for only one additional country when the World index is the reference index. Put differently, I find very strong statistical evidence *against* the contrarian strategy, contrary to previous results.

I also find very weak evidence in support of the homogeneity assumption ($\rho^i=\rho$, $\forall i$,), a key assumption of Balvers et al. (2000). For example, the price deviation for Italy relative to the World index exhibits strong mean reversion with a half-life ranging around 1 to 7 years, while for all other countries except Germany one cannot exclude the possibility of an infinite half-life for mean reversion. This result is important, since it suggests that the results of Balvers et al. (2000) might inherit the drawback of the Levin et al. (2002) type panel/SUR unit root test. It is well-known that such tests may reject the joint null hypothesis (H_0: $\rho^1=\cdots=\rho^N=\rho=1$) even when a single series is stationary ($\rho^i<1$, $\rho^j=1$, $\forall j\neq i$). It is thus possible that test results might be contaminated by the inclusion of outliers such as Italy and Germany when the World index serves as the reference index.

[24] We choose the number of lags (k) by the general-to-specific rule (Hall, 1994), as recommended by Ng and Perron (2001). We constructed the 95% confidence intervals nonparametrically by utilizing percentile estimates of ρ^i from 10,000 residual-based bootstrap simulations on the point estimates (Efron and Tibshirani, 1993).

[25] All results are available upon request.

Table 4
Bias-corrected estimates and grid bootstrap confidence intervals: national stock index deviations from the US index.

Country	$\tilde{\rho}$	CI	HL	CI
Australia	0.987	[0.966,1.000]	4.436	[1.679,∞)
Austria	0.993	[0.979,1.000]	8.073	[2.699,∞)
Belgium	0.988	[0.972,1.000]	4.762	[2.013,∞)
Canada	0.999	[0.986,1.000]	66.43	[4.109,∞)
Denmark	0.982	[0.962,1.000]	3.234	[1.491,∞)
France	0.979	[0.955,1.000]	2.760	[1.250,∞)
Germany	0.982	[0.958,1.000]	3.165	[1.331,∞)
Hong Kong	0.978	[0.956,1.000]	2.565	[1.282,∞)
Italy	0.980	[0.962,1.000]	2.853	[1.477,∞)
Japan	0.997	[0.985,1.000]	16.81	[3.886,∞)
Netherlands	1.000	[0.986,1.000]	∞	[3.975,∞)
Norway	0.985	[0.968,1.000]	3.941	[1.768,∞)
Singapore	0.989	[0.975,1.000]	5.159	[2.297,∞)
Spain	0.991	[0.976,1.000]	6.413	[2.361,∞)
Sweden	1.000	[0.984,1.000]	∞	[3.646,∞)
Switzerland	0.984	[0.961,1.000]	3.487	[1.460,∞)
UK	0.980	[0.957,1.000]	2.862	[1.307,∞)
Median	0.987	[0.968,1.000]	4.436	[1.768,∞)

Notes: i) The number of lags (k) was chosen by the general-to-specific rule (Hall, 1994) with maximum 12 month lags. ii) Sample periods are 1969:12–2007:9 ($T=454$). iii) For each stock index deviation, the bias-corrected 95% confidence interval as well as the bias adjusted estimate was obtained by 10,000 nonparametric bootstrap replications from the empirical distribution on 50 grid points in the neighborhood of the least squares point estimates (Hansen, 1999). iv) The 90% bias-corrected confidence intervals were still infinite for all countries.

I turn to our panel-SUR estimation results with annual frequency data. To directly compare our results with those of Balvers et al. (2000), we impose the homogeneity assumption on the speed of convergence.

I first implement a variance-decomposition analysis to see whether a cross-section independence assumption is appropriate to our aggregate macroeconomic data. For this purpose, I employ an approach of Imbs et al. (2005) assuming that θ_t is approximated by the cross-sectional mean of $\bar{r}_t = N^{-1}\sum_{i=1}^{N} r_t^i$. Mark and Sul (2008) show that this simple measure is highly correlated with other alternative popular measures of common factors. Regressing each r_t^i on θ_t and a constant, we obtain the factor loading coefficient δ^i. Under the assumption of orthogonality of θ_t to idiosyncratic shock η^i, Eq. (9) can be used to obtain measures of relative importance of common factors to each relative stock price. As we can see in Table 5, the role of common factor θ_t in explaining the variation of each deviation seems far from being negligible. This finding supports the cross-section consideration in estimating the speeds of convergence in our panel-SUR model.

Table 6 reports our panel-SUR estimation results. I first estimate the common persistence parameter ρ with cross-section dependence consideration. The 95% confidence bands are obtained by 10,000 nonparametric bootstrap simulations at the point

Table 5
Variance decomposition of r_t into a common factor and idiosyncratic components.

Country	World reference index				US reference index			
	$Var(r_t)$	$Var(\delta^i\theta_t)$	$\frac{Var(\delta^i\theta_t)}{Var(r_t)}$	δ^i	$Var(r_t)$	$Var(\delta^i\theta_t)$	$\frac{Var(\delta^i\theta_t)}{Var(r_t)}$	δ^i
Australia	0.067	0.002	0.027	0.335	0.104	0.033	0.316	0.808
Austria	0.180	0.034	0.191	1.463	0.281	0.174	0.620	1.857
Belgium	0.136	0.072	0.528	2.118	0.136	0.052	0.380	1.011
Canada	0.096	0.000	0.001	−0.086	0.137	0.026	0.186	0.711
Denmark	0.080	0.048	0.595	1.728	0.087	0.046	0.528	0.952
France	0.049	0.024	0.478	1.213	0.058	0.033	0.564	0.807
Germany	0.033	0.006	0.175	0.603	0.072	0.042	0.583	0.913
Hong Kong	0.274	0.111	0.406	2.639	0.272	0.068	0.251	1.161
Italy	0.142	0.005	0.036	−0.567	0.183	0.015	0.083	0.548
Japan	0.259	0.003	0.013	−0.457	0.476	0.297	0.624	2.427
Netherlands	0.186	0.051	0.274	1.785	0.163	0.022	0.134	0.656
Norway	0.103	0.028	0.268	1.311	0.163	0.096	0.589	1.379
Singapore	0.201	0.002	0.008	0.321	0.305	0.114	0.374	1.503
Spain	0.238	0.022	0.092	1.172	0.210	0.008	0.039	0.400
Sweden	0.242	0.073	0.299	2.129	0.210	0.023	0.110	0.675
Switzerland	0.077	0.036	0.469	1.500	0.052	0.015	0.294	0.553
UK	0.048	0.005	0.099	0.545	0.062	0.021	0.329	0.638
US	0.037	0.001	0.026	0.247				

Notes: i) Sample periods are 1969–2006 ($T=38$ and $N_{US}=17$ and $N_{World}=18$). ii) Cross-sectional average serves as a proxy variable for a common factor θ_t. iii) δ^i denotes the factor loading coefficient.

Table 6
Panel-SUR estimation with cross-section dependence consideration and grid bootstrap estimation results.

Panel-SUR estimates				
Reference	ρ	CI	HL	CI
World	0.833	[0.759,0.842]	3.784	[2.514,4.037]
		[0.767,0.836]		[2.612,3.879]
US	0.881	[0.791,0.875]	5.466	[2.958,5.199]
		[0.799,0.869]		[3.088,4.951]
Panel-SUR median-unbiased estimates				
Reference	$\tilde{\rho}$	CI	HL	CI
World	0.876	[0.822,0.933]	5.232	[3.526,10.06]
		[0.830,0.925]		[3.720,8.846]
US	0.948	[0.892,0.995]	13.07	[6.083,145.0]
		[0.901,0.989]		[6.640,63.39]

Notes: i) Sample periods are 1969–2006 ($T=38$ and $N_{\mathrm{US}}=17$ and $N_{\mathrm{World}}=18$). ii) The number of lags (k) was set at 1 by the BIC with a minimum lag 1 and maximum 2-year lags. iii) For each stock index deviation, the median bias-corrected 95% confidence interval as well as the bias adjusted estimate was obtained by 10,000 nonparametric bootstrap replications from the empirical distribution on 50 grid points in the neighborhood of the panel-SUR point estimates. iv) Confidence intervals written in small font denote the 90% confidence bands.

estimates. The implied half-life estimates are 3.784 and 5.466 years when the World index and the US index serve as a reference index, respectively. The 95% confidence bands range from around 3 to 5 years for each case. One interesting finding is that the panel-SUR estimates are severely downward-biased for ρ so that the 90% confidence band of the half-life with the US as a reference index does not include the corresponding point estimate. This possibility was raised by Andrews (1993). We, thus, correct for the bias by grid bootstrap method. Bias-corrected half-life estimates are 5.232 and 13.07 years for the World and the US as a reference index, respectively. Unlike the univariate estimations, the 95% confidence bands are finite even though the bands are still quite wide especially for the deviations from the US index. Even with 90% significance level, the confidence bands still far from being compact extending to over 60 years. Our results, therefore, cast doubt on the usefulness of the contrarian investment strategy.

6. Concluding remarks

This paper statistically evaluates the usefulness of the contrarian investment strategy in the context of national stock markets. When stock price indices share a common stochastic component, stock prices can deviate from the fundamental value trend only temporarily, and such deviations should die out eventually. Employing a panel approach, Balvers et al. (2000) provided strong empirical evidence of mean reversion and thus of the usefulness of the contrarian strategy.

My approach is different from theirs in two respects. First, I estimate the small-sample bias conditional on the realizations of nuisance parameters, because the bias in higher order autoregressive processes becomes random. Balvers et al. (2000) assume that the size of bias is constant even though they also employ a higher order autoregressive process. This results in inaccurate bias-correction. Second, we correct for small-sample bias nonparametrically with cross-section dependence consideration. Our variance-decomposition analysis provides ample evidence of non-negligible role of common shocks relative to idiosyncratic shocks in explaining the variation of stock price deviations from the fundamental values.

Correcting for small-sample bias, I obtain much longer (even infinite) half-life estimates than those found by Balvers et al. (2000). From univariate estimations with high frequency data, the upper bounds of the half-life estimates were not even finite for most countries, which imply that it may take an infinitely long time for *winner–loser reversals* to occur. Panel-SUR estimates were also quite long and confidence bands were far from being compact especially for the deviations from the US index. The contrarian investment strategy may generate excess returns. But how long should we wait for these returns?

References

Andrews, Donald W.K., 1993. Exactly median-unbiased estimation of first order autoregressive/unit root models. Econometrica 61, 139–165.

Andrews, Donald W.K., Chen, Hong Yuan, 1994. Approximately median-unbiased estimation of autoregressive models. Journal of Business and Economic Statistics 12, 187–204.

Bai, Jushan, Ng, Serena, 2004. A panic attack on unit roots and cointegration. Econometrica 72, 1127–1177.

Balvers, Ronald, Wu, Yangru, Gilliland, Erik, 2000. Mean reversion across national stock markets and parametric contrarian investment strategy. Journal of Finance 55, 745–772.

Bhojraj, Sanjeev, Swaminathan, Bhaskaran, 2006. Macromomentum: returns predictability in international equity indices. Journal of Business 79, 429–451.

Choi, In, 2001. Unit root tests for panel data. Journal of International Money and Finance 20, 249–272.

Cochrane, John H., 2008. The dog that did not bark: a defense of return predictability. Review of Financial Studies 21, 1533–1575.

Deb, Partha, Sefton, Martin, 1996. The distribution of a Lagrange multiplier test of normality. Economics Letter 51, 123–130.

DeBondt, Werner, Thaler, Richard, 1985. Does the stock market overreact? Journal of Finance 40, 793–805.

Efron, Bradley, Tibshirani, Robert J., 1993. An Introduction to the Bootstrap. Chapman and Hall/CRC, London, UK.

Fama, Eugene, French, Kenneth, 1988. Permanent and temporary components of stock prices. Journal of Political Economy 96, 246–273.

Hall, Alastair, 1994. Testing for a unit root in time series with pretest data-based model selection. Journal of Business and Economic Statistics 12, 461–470.

Hamilton, James D., 1994. Time Series Analysis. Princeton University Press, Princeton, New Jersey.

Hansen, Bruce E., 1999. The grid bootstrap and the autoregressive model. Review of Economics and Statistics 81, 94–607.

Im, Kyung So, Pesaran, M. Hashem, Shin, Yongcheol, 2003. Testing for unit roots in heterogeneous panels. Journal of Econometrics 115, 53–74.

Imbs, Jean, Mumtaz, Haroon, Ravn, Morten O., Rey, Helene, 2005. Ppp strikes back: aggregation and the real exchange rates. Quarterly Journal of Economics 120, 1–43.
Jarque, Carlos M., Bera, Anil K., 1980. Efficient tests for normality, homoscedasticity and serial independence of regression residuals. Economics Letters 6, 255–259.
Kasa, Kenneth, 1992. Common stochastic trends in international stock markets. Journal of Monetary Economics 29, 95–124.
Kendall, M.G., 1954. Note on bias in the estimation of autocorrelation. Biometrika 41, 403–404.
Kim, Hyeongwoo, Ogaki, Masao, 2009. Purchasing power parity and the taylor rule. Ohio State University Department of Economics Working Paper No. 09-03.
Kim, Myung Jig, Nelson, Charles R., Startz, Richard, 1991. Mean reversion in stock prices? A reappraisal of the empirical evidence. Review of Economic Studies 58, 515–528.
Levin, Andrew, Lin, Chien-Fu, Chu, Chia-Shang James, 2002. Unit root tests in panel data: asymptotic and finite-sample properties. Journal of Econometrics 108, 1–24.
Mark, Nelson C., 2001. International Macroeconomics and Finance: Theory and Econometric Methods. Blackwell Publishers, Oxford, UK.
Mark, Nelson C., Sul, Donggyu, 2008. Ppp Strikes Out: The Effect of Common Factor Shocks on the Real Exchange Rate. manuscript.
McQueen, Grant, 1992. Long-horizon mean-reverting stock prices revisited. Journal of Financial and Quantitative Analysis 27, 1–18.
Moon, Hyungsik Roger, Perron, Benoit, 2004. Testing for a unit root in panels with dynamic factors. Journal of Econometrics 122, 81–126.
Ng, Serena, Perron, Pierre, 2001. Lag length selection and the construction of unit root tests with good size and power. Econometrica 69, 1519–1554.
Nickell, Stephen, 1981. Biases in dynamic models with fixed effects. Econometrica 49, 1417–1426.
O'Connell, Paul G.J., 1998. The overvaluation of purchasing power parity. Journal of International Economics 44, 1–19.
Pesaran, M. Hashem, 2007. A simple panel unit root test in the presence of cross-section dependence. Journal of Applied Econometrics 22, 265–312.
Phillips, Peter C.B., Sul, Donggyu, 2003. Dynamic panel estimation and homogeneity testing under cross section dependence. Econometrics Journal 6, 217–260.
Poterba, James, Summers, Lawrence, 1988. Mean reversion in stock prices: evidence and implications. Journal of Financial Economics 22, 27–59.
Richards, Anthony J., 1995. Comovements in national stock market returns: evidence of predictability, but not cointegration. Journal of Monetary Economics 36, 631–654.
Richardson, Matthew, 1993. Temporary components of stock prices: a skeptic's view. Journal of Business and Economics Statistics 11, 199–207.
Richardson, Matthew, Stock, James, 1989. Drawing inferences from statistics based on multi-year asset returns. Journal of Financial Economics 25, 323–348.
Steinsson, Jón, 2008. The dynamic behavior of the real exchange rate in sticky-price models. American Economic Review 98, 519–533.
Taylor, Alan M., 2001. Potential pitfalls for the purchasing-power-parity puzzle? Sampling and specification biases in mean-reversion tests of the law of one price. Econometrica 69, 473–498.
Taylor, Mark P., Sarno, Lucio, 1998. The behavior of real exchange rates during the post-Bretton Woods period. Journal of International Economics 46, 281–312.