Economic Modelling 36 (2014) 161–171

Contents lists available at ScienceDirect

Economic Modelling

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

Does Purchasing Power Parity hold? New evidence from wild-bootstrapped nonlinear unit root tests in the presence of heteroskedasticity

Jen-Je Su [a], Adrian (Wai-Kong) Cheung [b,*], Eduardo Roca [a]

[a] Department of Accounting, Finance and Economics, Griffith University, Nathan, Queensland 4111, Australia
[b] Department of Finance and Banking, School of Economics and Finance, Curtin University, Bentley, Western Australia 6102, Australia

ARTICLE INFO

Article history:
Accepted 17 September 2013

JEL classification:
C10
C12
C22
C50

Keywords:
(Nonlinear) unit root test
Heteroskedasticity
Wild bootstrapping
Purchasing power parity

ABSTRACT

In spite of the extensive research which has already been undertaken, the issue as to whether Purchasing Power Parity (PPP) empirically holds, continues to be strongly debated. Existing studies have been criticized for their reliance on unit root tests which are deemed to suffer from certain weaknesses such as the size distortion bias arising from heteroskedasticity. In this paper, we provide new evidence on PPP based on a new methodology that overcomes this problem. We use the widely accepted KSS (Kapetanios et al., 2003) non-linear unit root tests which we, however, wild bootstrapped. Through Monte Carlo simulation, we demonstrate that the wild-bootstrapped KSS is robust to heteroskedasticity-induced size distortion problem. We apply this method to test PPP across 61 countries over the period 1994 to 2012 — a period characterized by a number of crises such as the Asian Financial Crisis, Russian Crisis, dotcom crisis, Global Financial Crises, among others, and therefore, intense heteroskedasticity. Our results provide strong evidence against PPP. This paper contributes to both the international financial economics and econometrics literatures.

1. Introduction

Purchasing Power Parity (PPP), as first articulated by Cassel (1918, 1922), is one of the key assumptions in open macroeconomics and international finance models; however, after thirty years of intensive research, the empirical validity of PPP remains far from conclusive. PPP is a simple idea based on the law of one price which postulates that identical goods should sell at the same price in different countries and that the exchange rates between currencies will allow this to happen. In the absolute version of the PPP, the exchange rate would simply be the ratio of the price levels between countries. On the other hand, in the relative version of the PPP, the change in exchange rates offsets the differential in the relative change in prices between countries which implies that the real exchange rate (RER) will be stationary (see Sarno and Taylor, 2002; Taylor, 2006; Taylor and Taylor, 2004 for a detailed discussion of PPP). Hence, the PPP hypothesis is typically examined by testing if RERs are stationary using unit root tests. Earlier studies, which were often based on the Dickey–Fuller (DF) tests, found little evidence to support PPP (e.g. Adler and Lehmann (1983)). The failure has been attributed to the low power of the DF tests and the literature has consequently moved on to enhance the testing power using various approaches (see, for example, Rogoff (1996) and Taylor and Taylor (2004)).

One of the new developments in the PPP empirics is the use of nonlinear unit root tests. Due to the existence of transactions costs, according to Michael et al. (1997) and Taylor (2001), RERs may revert to the long-run equilibrium only when they are sufficiently distant from the long-run equilibrium; in other words, RERs are globally stationary but adjusted in a nonlinear (threshold-like) fashion. The DF-type tests, which are developed in a linear (autoregressive) context, may exhibit low power against such nonlinear stationarity. To address the problem, a range of nonlinear unit-root tests have been suggested — see, for example, Enders and Granger (1998), Kapetanios et al. (2003) (hereafter, KSS), Bec et al. (2004), and Sollis (2009). Among these, KSS (2003) is probably the most widely recognized and applied. KSS (2003) proposed a unit-root test using an auxiliary regression model that approximates the exponential smooth transition autoregressive (ESTAR) process by Taylor series. Using the KSS test, stronger supporting evidence of PPP has been found — see, among others, KSS (2003), Liew et al. (2004), Bahmani-Oskooee et al. (2007), Pesaran et al. (2009), and Zhou and Kutan (2011).

Unit root tests may suffer non-trivial size distortion in the presence of *conditional* heteroskedasticity (e.g. generalized autoregressive heteroskedasticity (GARCH)) as well as *unconditional* time-varying variance (non-stationary volatility). As shown in Valkanov (2005), with strong GARCH effect (i.e. when the GARCH process is nearly integrated and the volatility parameter is relatively large) often observed in economic/financial time series, convergence of the finite-sample DF

* Corresponding author.

0264-9993/$ – see front matter © 2013 Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.econmod.2013.09.029

distribution to the asymptotic distribution appears to be very slow. As a result, the usual DF test tends to be oversized in the presence of strong GARCH effect (see also Kim and Schmidt (1993) and Su (2011)). Cook (2006) also found that the size distortion due to GARCH can be even more severe when the (nonlinear) KSS test is considered. In addition, many economic and financial variables (including foreign exchange rates) are characterized by the existence of permanent volatility breaks (see, for example, Loretan and Phillips (1994) and Busetti and Taylor (2003)). In theory, permanent shifts in volatility can greatly affect linear unit root inference not only in finite samples but also asymptotically (Hamori and Tokihisa (1997), Kim et al. (2002) and Cavaliere (2004)). However, the impact of volatility shifts to nonlinear unit root tests remains largely unexplored.

We propose to use the wild bootstrap method to overcome the aforementioned size problems. The idea behind this method is to replicate in resampled data the pattern of heteroskedasticity in the original data. As shown in Cavaliere and Taylor (2008, 2009) the wild bootstrap inference is robust to both conditional and unconditional heteroskedasticity and is able to achieve the (infeasible) size-corrected power of the usual (*linear*) unit root tests. Interestingly, even though the statistical performance of wild-bootstrapping has not yet been examined in the literature for the *nonlinear* unit root tests, there are two empirical works that have already incorporated the wild bootstrap method to the KSS test (Arghyrou and Gregoriou, 2007, 2008). Both papers found apparently conflicting results between the standard and bootstrapping inferences. On one hand, Arghyrou and Gregoriou (2007) found that the standard and bootstrap approaches arrive in the same conclusion when the DF test is considered: none of the 7 bilateral RERs (against the US dollar) they examined is stationary. On the other hand, Arghyrou and Gregoriou (2008) reported that non-stationarity is rejected in 6 out of the 7 RERs when the standard KSS inference is applied but, strikingly, only 1 rejection with the associated wild bootstrap inference.[1] Given the findings of Arghyrou and Gregoriou (2007, 2008), one might doubt if the results that are biased toward PPP from the KSS test are nothing but a product of size distortion owing to heteroskedasticity. Once heteroskedasticity is appropriately accounted for, can KSS test still produce more support for PPP than DF?

In this paper, we aim to shed lights on these issues. First, we show that the KSS test is more size-distorted than the DF test in the presence of non-constant variances. Second, we show that the wild bootstrap method that works well with the DF test also works satisfactorily with the KSS test. Third, we apply the DF and KSS tests to 61 real effective exchange rates (REERs) with data from the Bank of International Settlement (BIS). We find that wild bootstrap implementation of the KSS test produces much less rejection of non-stationarity than standard implementation but the KSS test still rejects more often than the DF test, suggesting that the real exchange rates are non-stationary in most of these countries. This finding also implies that the PPP does not hold and arbitrage opportunities exist. Possible economic explanations include transaction costs (Dumas, 1992; Obtstfeld and Taylor, 1997; Sercu, Uppal and Van Hulle, 1995), limits to arbitrage (Zussman, 2003), heterogeneous agents (Reitz and Taylor, 2008), presence of target zones (Krugman, 1991), central bank interventions (Dominguez, 1998; Lee, 2011).

Our paper contributes to both the international financial economics and the econometrics literatures. In terms of its contributions to the international financial economics literature, our paper provides fresh evidence on PPP — an issue that is still highly debated notwithstanding the large body of research on this topic which has built up since the early 1970s. It has been suggested that in order to move this debate forward constructively, there is a need for empirical studies which incorporate non-linearities, as first pointed out by Rogoff (1996) and use data sets with more extensive coverage (Sarno and Taylor, 2002; Taylor, 2006; Taylor and Taylor, 2004). Our study fills these important gaps in the literature. As mentioned, our paper examines an extensive number of countries — 61 in total, and is based on a more updated data set over a period of 18 years (1994–2012) that is characterized by the occurrence of a number of financial crises and therefore, of high heteroskedasticity.

In relation to our paper's contribution to the econometrics literature, through Monte Carlo simulation, we demonstrate that wild bootstrapping the KSS test eliminates the size distortion problem induced by heteroskedasticity. This has not been done previously in the literature. Arghyrou and Gregoriou (2008) incorporated wild bootstapping into KSS test but they did not conduct any simulation to prove that the KSS test works well with wild bootstrapping. The statistical properties of wild bootstrapping have been examined in the literature but this is only in relation to linear unit root tests such as the DF tests (see, for example, Cavaliere and Taylor (2008, 2009)).

The rest of the paper proceeds as follows. Section 2 discusses the Monte Carlo simulation setups and presents results from the simulations. Section 3 reports the empirical results and Section 4 concludes. A brief review of the unit root tests and wild bootstrapping procedure can be found in Appendix A.

2. Monte Carlo simulation setup and results

2.1. Setup for size issue and results

To investigate the size properties of the DF and KSS tests. The data-generating process is a drift-less integrated process $y_t = y_{t-1} + \varepsilon_t$, $t = 1,\ldots,T$, with heteroskedastic errors: $\varepsilon_t = \omega_t \sigma_t$ where σ_t is iid N(0,1) and the volatility parameter ω_t is specified as the following models.

1. No break: $\omega_t = 1$.
2. GARCH: $\omega_t = \phi_0 + \phi_1 \varepsilon_{t-1}^2 + \phi_2 \omega_{t-1}$.
3. (Exponential (Near-)Integrated) Stochastic volatility (SV): $\omega_t = \omega_0 \exp\left(b\,\xi_t/\sqrt{T}\right)$ where $\xi_t = (1 - c/T)\xi_{t-1} + k_t$, with $k_t \sim$ iid $N(0,1)$.
4. Single break (SB) in volatility: $\omega_t = \omega_0 + (\omega_1 - \omega_0) I\left(\frac{t}{T} \geq \tau\right)$, $\tau \in (0,1)$.
5. Double break (DB) in volatility: $\omega_t = \omega_0 + (\omega_1 - \omega_0) I\left(\tau_1 \leq \frac{t}{T} \leq \tau_2\right)$, $\tau_1, \tau_2 \in (0,1)$.
6. Trending Volatility (TV): $\omega_t = \omega_0 + (\omega_1 - \omega_0)\left(\frac{t-1}{T-1}\right)$.

Model 2 is the standard GARCH(1,1) model the parameters of which are set as follows: $(\phi_1, \phi_2) = (0.29, 0.7), (0.2, 0.7), (0.19, 0.8), (0.1, 0.8)$ and $\phi_0 = 1 - \phi_1 - \phi_2$. The SV process of Model 3 is generated with $b = 2$ and $c = 0$, 10. Models 4–6 refer to those non-stationary volatility cases considered in Cavaliere and Taylor (2008, 2009). Note that the variance of ε_t is $\mathrm{var}(\varepsilon_t) = \omega_t^2$. Model 4 corresponds to a single abrupt variance shift from ω_0^2 to ω_1^2 occurring at $t = \tau T$. Model 5 is with double shifts: the first shift at $t = \tau_1 T$ (from ω_0^2 to ω_1^2) and the second shift at $t = \tau_2 T$ (from ω_1^2 to ω_0^2). Model 6 generates smooth (trending) breaks over the whole sample period $t = 1$ to T. For these three models, we let $\omega_0 = 1$ and set $\delta\ (= \omega_0^2/\omega_1^2)$ equal to 1/2, 2, 1/5, 5, respectively. We note that $\delta < 1$ corresponds to a downward shift while $\delta > 1$ upward shift. For Model 4 (SB): we consider $\tau = 0.2$ (early break), 0.8 (late break). For Model 5 (DB), we set $\tau_2 = (1 - \tau_1)$ and $\tau_1 = 0.2, 0.8$.

All simulations are based on 20,000 replications with $T = 100$, 250 and done by GAUSS. For the wild bootstrapping, we use the warp-speed Monte Carlo method of Giacomini et al. (2013) using a single bootstrap re-resample (i.e. $M = 1$ in Step 3 of the bootstrap procedure described in Appendix A) and the bootstrap critical values are obtained based on the 20,000 Monte Carlo replications. Rejection frequencies are calculated at the 5% nominal significance level with the DF and KSS tests using the standard and bootstrap inferences, respectively. For simplicity, in both tests we set the augmentation equal to zero.[2] We report

[1] We note that Arghyrou and Gregoriou (2008) do not deal with GARCH or time-varying volatility; instead, the wild bootstrap method is used to account for non-normality.

[2] We have also considered augmented tests and, as expected, the results are very similar.