

# Long-run real exchange rate changes and the properties of the variance of $k$ -differences\*

Sungwook Park, Masao Ogaki<sup>†</sup>

February 16, 2007

## Abstract

Engel (1999) computes the variance of  $k$ -differences for each time horizon using the method of Cochrane (1988) in order to measure the importance of the traded goods component in U.S. real exchange rate movements. The importance of traded goods should decrease as the horizon increases if the law of one price holds for traded goods in the long run. However, Engel finds that the variance of  $k$ -differences decreases only initially and then increases as  $k$  approaches the sample size. He interprets the increasing variance as evidence of an increase in the long-run importance of the traded goods component. By contrast, we show that the variance of  $k$ -differences tends to return to the initial value as  $k$  approaches the sample size whether the variable is stationary or unit root nonstationary. Our results imply that the increasing variances for  $k$ -values close to the sample size cannot be interpreted as evidence of an increase in the importance of the traded goods component in the long run. We find that our test results regarding the variance of  $k$ -differences are consistent with smaller importance of the traded goods component in the longer run.

Keywords: Real exchange rate, Variance ratio, Traded and nontraded goods  
JEL Classification: F31

---

\*For helpful comments, we thank Joseph Kaboski, Pok-sang Lam, J. Huston McCulloch, Mototsugu Shintani, Yixiao Sun, seminar participants at Bank of Japan, Keio University, Osaka University, and the Midwest Econometrics Group 2006 Meeting.

<sup>†</sup>Department of Economics, The Ohio State University, 410 Arps Hall, 1945 N. High Street, Columbus, OH 43210. *E-mail address*: park.622@osu.edu

# 1 Introduction

Since Balassa (1964) and Samuelson (1964), the disaggregation of the economy into internationally traded and nontraded sectors has been one of the main building blocks in many open economy models. In those two-good models, if the law of one price holds in the prices of traded goods, then the real exchange rate (RER) is determined by the movement of its nontraded goods component which consists of the relative prices of nontraded goods.

Since Isard (1977), however, empirical evidence has clearly shown that, in the short run, the law of one price does not hold even for available measures of traded goods. Thus, the Balassa-Samuelson view focusing on the role of nontraded goods had been thought to better fit the long run.

Contrary to this traditional view, Engel (1999) presented empirical results which can be interpreted to imply that almost all U.S. RER movements can be accounted for by movements of the traded good component at all time horizons. Engel (1999) himself refrains from reaching a decisive conclusion about the long-run time horizon and argues that his results are mainly about short and medium horizons because of the small number of observations. Nevertheless, some authors have taken Engel's (1999) results as evidence against the relevance of the traditional dichotomy of goods in modeling long run real exchange rate movements<sup>1</sup>. For example, Obstfeld (2001) writes;

This is a striking contradiction of the Harrod-Balassa-Samuelson theory. International divergences in the relative consumer price of "tradables" are so huge that the theoretical distinction between supposedly arbitrated tradables prices and completely sheltered nontradables prices offers little or no help in understanding U.S. real exchange rate movements, even at long

---

<sup>1</sup>Those who discuss Engel's (1999) empirical findings in the context of the long run movements of the real exchange rate include Alexius and Nilsson (2000), Chinn (2006), Obstfeld (2001), Sarno and Taylor (2002), Sarno and Valente (2006), Schnatz, Vijselaar, and Osbat (2004), and Taylor and Taylor (2004).

horizons.

In his paper, Engel measures the importance of the traded goods component in accounting for U.S. real exchange rate movements by adopting the variance of  $k$ -differences used in Cochrane (1988). In this paper, we challenge this widely accepted interpretation of Engel's results about the long run movement of the RER by analyzing properties of the limit distribution of the variance of  $k$ -differences when  $k$  is close to the sample size.

The variance of  $k$ -differences of a time series  $z_t$  is denoted as  $V_k(z)$  in this paper. As in Cochrane (1988),  $V_k(z)$  is defined as follows:

$$V_k(z) \equiv \frac{T}{(T-k)(T-k+1)k} \sum_{t=0}^{T-k} [z_{t+k} - z_t - k\overline{\Delta z}]^2, \quad (1)$$

$$\text{where } \overline{\Delta z} = \frac{1}{T}(z_T - z_0).$$

According to the definition in Equation (1), the variance of  $k$ -differences is a variance of  $k$ -period differences centered around the sample mean of the difference.

Cochrane (1988) shows that  $V_k(z)$  is asymptotically equivalent to the Bartlett kernel estimator of the long run variance of  $\Delta z_t$ . If the law of one price holds for traded goods in the long run, then the long run variance of the traded goods component of the RER is zero since it is stationary. Thus, based on Cochrane (1988), Engel (1999) expects that  $V_k(z)$  for the traded goods component will converge down to zero as  $k$  increases if the traditional Balassa-Samuelson view is true for the long run<sup>2</sup>.

However, Engel's empirical results show that  $V_k(z)$  for the traded goods component decreases at first but increases towards the end of time horizons, most prominently in the case of the US-Canada RER.<sup>3</sup> Engel interprets the rise in the later part of the

---

<sup>2</sup>In Engel (1999), the formula for  $V_k(z)$  is a little different from Cochrane's (1988). Engel does not divide it by  $k$ . Thus, Engel says, "One expects the variance of  $k$ -differences of  $x_t$  [the traded goods component] to converge [to a finite number] as  $k$  gets large."

<sup>3</sup>As we shall see in the next section, what Engel (1999) actually computes is the ratio of  $V_k$  of the traded goods component to that of the real exchange rate. However, the shape of the graph is mainly determined by the numerator. It is because  $V_k$  of the nontraded goods component is expected to remain

graph as an increase in the importance of the traded goods component in the long run movement of the RER<sup>4</sup>.

This paper shall show that  $V_k(z)$  for  $k \cong T$  tends to go back to the initial value on average as  $k$  gets closer to the sample size, whether the variable of interest is mean-reverting or not. As such, Engel's (1999) observation about the long run time horizons may come simply from this statistical property of the variance of  $k$ -differences and have little to do with the long run properties of the real exchange rate.

Our findings in this paper imply that the variances of  $k$ -differences in the middle range of  $k$ 's are more relevant to the long run than those at  $k$ 's close to both ends of the time span. Thus, the fall of the graph of the variance of  $k$ -differences in the middle range of Engel's results favors the smaller importance of the traded goods component in the longer run. However, Engel (1999) finds that the fall is not statistically meaningful from his test based on the variance of  $k$ -differences. After some adjustments in Engel's testing method, however, our results show that the fall of the graph is statistically significant for some countries, meaning that Engel's test results are not very robust. As such, arguing that the nontraded goods component plays the same minimal role for the long run movement of the US real exchange rate based on Engel's empirical results is less convincing.

The evidence in this paper is consistent with recent works. Kakkar and Ogaki (1999) run a cointegration regression of the real exchange rate on its nontraded goods component and find that the nontraded goods component can explain long run real exchange rate movements fairly well. Related evidence for the usefulness of the dichotomy of goods in understanding the real exchange movements is found in a line of studies on the half-life<sup>5</sup> of the real exchange rate. Crucini and Shintani (2002), Kim (2005), and

---

constant on average if it is random walk.

<sup>4</sup>p.513 in Engel (1999). Applying Engel's approach to bilateral Asian-Pacific real exchange rates, Parsley (2001) also finds the rise in the later part of the graph in the case of US-Hong Kong and interprets it as an increase in the variability of the traded goods component in that time horizon (p.9).

<sup>5</sup>Half life is the time it takes for half the effects of a given shock to dissipate.

Kim and Ogaki (2004) find that half-lives of the RER based on traded good prices are shorter than those of the RER based on nontraded good prices. Crucini, Telmer, and Zachariades (2005) also find that the law of one price holds better for traded goods than for nontraded goods in data for over 500 goods. Taylor and Taylor (2004) state that the Harrod-Balassa-Samuelson model of equilibrium real exchange rates is attracting renewed interest as a desirable modification [of PPP theory] after languishing for some years in relative obscurity.

The rest of the paper is organized as follows. Section 2 will review the existing literature on the asymptotic distribution of  $V_k(z)$  and provide the main theoretical result of this paper. Section 3 discusses the implication of this paper's result for Engel's findings and presents our test results based on the variance of  $k$ -differences. Section 4 concludes.

## 2 The statistical properties of the variance of $k$ -differences

### 2.1 Existing theories on the statistical properties of $V_k(z)$

Throughout the paper, suppose that the following Assumption 1 holds for a random variable,  $z_t$ .

**Assumption 1** For a random variable,  $z_t$ , assume that  $\Delta z_t = d + \psi(L)\varepsilon_t = d + \sum_{j=0}^{\infty} \psi_j \varepsilon_{t-j}$ , where  $\sum_{j=0}^{\infty} j \cdot |\psi_j| < \infty$  and  $\{\varepsilon_t\}$  is an *i.i.d.* sequence with mean zero, variance  $\sigma^2$ , and finite fourth moment<sup>6</sup>. Define

---

<sup>6</sup>We follow the assumption in Proposition 17.3 in Hamilton (1994) in order to apply the functional central limit theorem to unit root nonstationary processes with serial correlation.

$$\begin{aligned}\gamma_j &\equiv E[(\Delta z_{t+j} - d)(\Delta z_t - d)] = \sigma^2 \sum_{s=0}^{\infty} \psi_s \psi_{s+j} \quad \text{for } j = 0, 1, 2, \dots \\ \lambda &\equiv \sigma \sum_{j=0}^{\infty} \psi_j = \sigma \cdot \psi(1) \\ \Omega &\equiv \sum_{j=-\infty}^{\infty} \gamma_j = \lambda^2\end{aligned}$$

The variance of  $k$ -differences has been studied in the context of the variance ratio test for the random walk hypothesis. Earlier works focused on the case when  $k$  is much smaller than the sample size. As in Lo and MacKinlay(1999)<sup>7</sup>, the variance of  $k$ -differences can be expressed as a weighted average of sample autocovariances, only with a small difference in order of  $o_p(T^{-1/2})$ ;

$$\begin{aligned}V_k(z) &= \sum_{\tau=-k+1}^{k-1} \frac{k-|\tau|}{k} \widehat{\gamma}_\tau + o_p(T^{-1/2}), \\ \text{where } \widehat{\gamma}_\tau &\equiv \frac{1}{T} \sum_{t=1}^{T-|\tau|} (\Delta z_t - \overline{\Delta z})(\Delta z_{t+|\tau|} - \overline{\Delta z}).\end{aligned}\tag{2}$$

Hence, when  $k$  is relatively small and fixed, by the law of large numbers,

$$V_k(z) \rightarrow \sum_{\tau=-k+1}^{k-1} \frac{k-|\tau|}{k} \gamma_\tau \text{ as } T \rightarrow \infty.\tag{3}$$

Especially when  $k = 1$ ,

$$V_1(z) \rightarrow \gamma_0 \text{ as } T \rightarrow \infty.$$

For the variance ratio test, in which the test statistic is defined as follows:

$$VR_k(z) \equiv V_k(z)/V_1(z),\tag{4}$$

---

<sup>7</sup>p54, chapter 3.

it is possible to show the following asymptotic distribution of  $VR_k(z)$  :

$$\sqrt{T}(VR_k(z) - 1) \xrightarrow{D} N(0, \sigma_k^2), \quad (5)$$

where  $\sigma_k^2$  is a simple function of  $k^8$ .

The variance of  $k$ -differences is asymptotically equivalent to the Bartlett kernel estimator for the long-run variance of  $\Delta z_t$ , as pointed out in Cochrane (1988)<sup>9</sup>. Equation (2) illustrates the fact. The first term of the right hand side in equation (2) is the definition of the Bartlett kernel estimator with the lag length of  $k$ . Newey and West (1987) show that the Bartlett kernel estimator converges to the long-run variance of  $\Delta z_t$  as  $T \rightarrow \infty$  and  $k \rightarrow \infty$  at a much slower growth rate<sup>10</sup>,  $O(T^{1/4})$ . Thus, under certain conditions,

$$V_k(z) \rightarrow \Omega \quad \text{when } k/T \rightarrow 0 \text{ as } T \rightarrow \infty. \quad (6)$$

However, it turns out that the variance ratios do not converge to a point, are severely right skewed for relatively large  $k$  in a small sample, and are not asymptotically normally distributed as in equation (5). So it is not appropriate to apply conventional asymptotics to this case. Richardson and Stock (1989) study the limit distribution of  $V_k(z)$  when  $k/T \rightarrow b > 0$  under the null of a random walk, and Deo and Richardson (2003) extend Richardson and Stock's (1989) result to the process which contains both permanent and transitory components. They find that  $V_k(z)$  does not converge to a limit but to a nondegenerate limiting distribution, which is a functional of a Brownian motion as

---

<sup>8</sup>For instance, if  $\psi(L) = 1$  and  $\varepsilon_t$  is an iid normal random variable with variance  $\sigma^2$ , then  $\sigma_k^2 = \frac{2(2k-1)(k-1)}{3k}$ .

<sup>9</sup>Actually, what Cochrane (1988) shows is that the population variance of  $k$ -differences is exactly equal to the population counterpart of the Bartlett estimator of long run variance. After replacing the two population concepts with the sample counterparts, the equality becomes an asymptotic equivalence.

<sup>10</sup>Later, Andrews (1991) shows that the Bartlett kernel estimator can attain consistency with the bandwidth at growth rate  $o(T^{1/2})$ .

follows:

$$V_k(z) \Rightarrow \frac{\Omega}{(1-b)^2 b} \int_b^1 [W(r) - W(r-b) - bW(1)]^2 dr, \quad (7)$$

as  $k/T \rightarrow b$  and  $T \rightarrow \infty$ ,

where  $W(r)$  is a standard Brownian motion.

Unlike the case when  $k/T \rightarrow 0$ , the limit distribution of  $V_k(z)$  in this case is significantly different from that of the Bartlett kernel estimator. To see the difference, we can rewrite the expression for  $V_k(z)$  in equation (1) as follows<sup>11</sup>.

$$V_k(z) = \frac{T^2}{(T-k)(T-k+1)} \left( \hat{\Omega}_k - \frac{1}{Tk} \sum_{t=1}^{k-1} S_t^2 - \frac{1}{Tk} \sum_{t=T-k+1}^{T-1} S_t^2 \right) \quad (8)$$

where  $\hat{\Omega}_k$  is the Bartlett kernel estimator with the bandwidth of  $k$ ,

and  $S_t$  is the partial sum process,  $\sum_{i=1}^t u_i$ .

In equation (8), the main differences between the Bartlett kernel estimator of the long run variance and the variance of  $k$ -differences are the two partial sum processes in the parenthesis in equation (8). This difference indicates the fact that the variance of  $k$ -differences underweights observations around both endpoints as mentioned in Cochrane (1988).

To learn the difference between the variance of  $k$ -differences and the Bartlett kernel estimator for large  $k$ 's, we compute the mean of each term in equation (8). As in equation (9), Kiefer and Vogelsang (2005) provide the analytical form of the limit distribution of the Bartlett kernel estimator and its mean when  $k/T \rightarrow b > 0$ .<sup>12</sup>

<sup>11</sup>This expression is inspired by Cai and Shintani (2006) and Kiefer and Vogelsang (2002).

<sup>12</sup> $\widetilde{W}(r)$  is called a Brownian bridge. Davidson (1994, p.445) explains this as a Brownian motion tied down at both ends.



$$\widehat{\Omega}_k \Rightarrow Q(b) \equiv \frac{2\Omega}{b} \left[ \int_0^1 \widetilde{W}(r)^2 dr - \int_0^{1-b} \widetilde{W}(r+b)\widetilde{W}(r) dr \right], \quad (9)$$

where  $\widetilde{W}(r) \equiv W(r) - \lambda W(1)$

$$E(Q(b)) = \Omega \left( 1 - b + \frac{b^2}{3} \right). \quad (10)$$

From the functional central limit theorem and the continuous mapping theorem,

$$\frac{1}{Tk} \sum_{t=1}^{k-1} S_t^2 \Rightarrow \frac{\Omega}{b} \int_0^b \widetilde{W}(r)^2 dr, \quad (11)$$

$$\frac{1}{Tk} \sum_{t=T-k+1}^{T-1} S_t^2 \Rightarrow \frac{\Omega}{b} \int_{1-b}^1 \widetilde{W}(r)^2 dr, \quad \text{as } k/T \rightarrow b \text{ and } T \rightarrow \infty. \quad (12)$$

The expectations of the limit distributions for the two partial sum processes are as follows.

$$E \left( \frac{\Omega}{b} \int_0^b \widetilde{W}(r)^2 dr + \frac{\Omega}{b} \int_{1-b}^1 \widetilde{W}(r)^2 dr \right) = \Omega \left( b - \frac{2}{3}b^2 \right)^{13}. \quad (13)$$

Equation (11) shows that the mean of the Bartlett kernel estimator when  $b > 0$  is proportional to, but different from the long run variance,  $\Omega$ . Equation (13) shows that the variance of  $k$ -differences before the small sample correction may be even further away from long run variance on average than the Bartlett kernel estimator. The small correction term adjusts the mean of the variance of  $k$ -differences to the level of the long run variance<sup>14</sup>: That is,

$$\frac{T^2}{(T-k)(T-k+1)} \rightarrow \frac{1}{(1-b)^2} \quad \text{as } k/T \rightarrow b \text{ and } T \rightarrow \infty.$$

<sup>13</sup>The following formula on p.445 in Davidson (1994) is used to compute the expectation:  $E(\widetilde{W}(t)\widetilde{W}(s)) = \min(t, s) - ts$ .

<sup>14</sup>It can be shown that the three limit distributions in equations (9), (11), and (12) are consistent with Deo and Richardson's (2003) limit distribution as in equation (7).

Meanwhile,

$$\begin{aligned} E(Q(b)) - E\left(\frac{\Omega}{b} \int_0^b \widetilde{W}(r)^2 dr + \frac{\Omega}{b} \int_{1-b}^1 \widetilde{W}(r)^2 dr\right) &= \Omega\left(1 - b + \frac{b^2}{3}\right) - \Omega\left(b - \frac{2}{3}b^2\right) \\ &= \Omega(1 - b)^2. \end{aligned}$$

To sum up, when  $k/T \rightarrow b > 0$ , the limit distribution of variance of  $k$ -differences is significantly different from that of the Bartlett kernel estimator. On the other hand, both the Bartlett kernel estimator and the variance of  $k$ -differences converge to a limit distribution which is the multiple of the long run variance and a nuisance-parameter-free distribution. The nuisance-parameter-free distributions depend only on the value of  $b$  and are invariant to the distribution of each variable.<sup>15</sup> Since the variance of  $k$ -differences is proportional to the long run variance when  $k/T \rightarrow b > 0$ , the ratio of the variance of  $k$ -differences at large  $k$ 's may contain some information about the ratio of long run variances even though the variance of  $k$ -differences is no longer consistent.

Engel's (1999) inference about the relative importance of the traded goods component in the RER movement relies on the statistical properties of the variance of  $k$ -differences at  $b = 0$  even when  $k$  is relatively big. The fact that the variance of  $k$ -differences has different limit distributions depending on the value of  $b$  raises doubt about Engel's inference. However, there seems to be some hope for Engel's argument about large  $k$ 's because of the proportionality of the limit distribution of the variance of  $k$ -differences to the long run variance at  $b > 0$ . Even so, it should also be noted that there exists a noticeable difference between the limit distribution of the Bartlett kernel estimator and that of the variance of  $k$ -differences at  $b > 0$ <sup>16</sup>.

---

<sup>15</sup>Using this property, Kiefer and Vogelsang (2005) are able to construct a test with this inconsistent Bartlett kernel estimator. They call their approach "fixed- $b$  asymptotics" and the conventional approach "small- $b$  asymptotics" Sun, Phillips, and Jin (2006) show another way to utilize "fixed- $b$  asymptotics".

<sup>16</sup>Later in this paper, we shall see how different the variance of  $k$ -differences can be from the Bartlett kernel estimator for a given  $b$  in Figure 2.

## 2.2 Statistical properties of $V_k(z)$ when $k$ is close to the sample size

The limit distribution of the Bartlett kernel estimator in Kiefer and Voglesang (2005) is applicable for  $k/T \rightarrow b$  in the interval of  $(0,1]$ , including the case when  $b = 1$ . On the other hand, the limit distribution of the variance of  $k$ -differences in equation (7) is applicable only for  $b$  in  $(0,1)$ . In equation (7), the limit distribution is not defined when  $b = 1$  since both numerator and denominator in the limit distribution become zero in this case<sup>17</sup>. So we cannot say that the variance of  $k$ -differences is proportional to the long run variance when  $b = 1$ . In other words, while there is continuity in the limit distribution of the Bartlett kernel estimator at  $b = 1$ , such continuity does not exist for the limit distribution of the variance of  $k$ -differences. Thus, at this point, the difference between the Bartlett kernel estimator and the variance of  $k$ -differences is so huge that the two are not even close to each other.

Unlike the previous cases when  $b < 1$ , only a small of number of observations are used to compute the variance of  $k$ -differences at  $b = 1$  regardless of the sample size. For example, when  $k = T - 1$ , there are only two observations available for any given sample size. As a result, conventional asymptotic theory is not applicable. Due to this restriction, we characterize the statistical properties of the variance of  $k$ -differences with the mean of its limit distribution instead of the analytical expression for the limit distribution itself.

The exact analytical solution for the mean of the limit distribution can be computed for the case when  $k = T - 1$ , the largest possible value of  $k$ . It turns out that there exists a symmetric relationship between the two extreme cases when  $k = 1$  and when  $k = T - 1$ . For the case when  $k < T - 1$ , the symmetry is not exact but approximate. The following proposition establishes a statistical property of the variance of  $k$ -differences when  $k = T - 1$ , the largest possible  $k$ .

---

<sup>17</sup>In particular, the fact that the whole term in the parenthesis in the equality (8) is zero when  $k = T$  is consistent with Kiefer and Vogelsang's (2002) proof.

**Proposition 1** Under Assumption 1, the limit of the mean of  $V_k(z)$  when  $k$  is the largest possible, i.e.  $T - 1$ , is equal to the variance of the change, which is equal to the limit of  $V_1(z)$ . That is,

$$\lim_{T \rightarrow \infty} E(V_{T-1}(z)) = \gamma_0 = \lim_{T \rightarrow \infty} V_1(z). \quad (14)$$

*Proof of Proposition 1.* First, without loss of generality, let's assume that the drift term,  $d$ , in Assumption 1 is zero<sup>18</sup>.

Next, let's transform equation (1) into the following:

$$\begin{aligned} V_k(z) &\equiv \frac{T}{(T-k)(T-k+1)k} \sum_{t=0}^{T-k} [z_{t+k} - z_t - k\overline{\Delta z}]^2 \\ &= \frac{T}{(T-k)(T-k+1)k} \sum_{j=0}^{T-k} \left[ \sum_{t=1}^k \{ \Delta z_{t+j} - \overline{\Delta z} \} \right]^2 \\ &= \frac{T}{(T-k)(T-k+1)k} \sum_{j=0}^{T-k} \left[ \sum_{t=1}^k u_{t+j} \right]^2, \text{ where } u_t \equiv \Delta z_t - \overline{\Delta z} \end{aligned} \quad (15)$$

To deal more easily with the case when  $k$  is close to the sample size, let  $m \equiv T - k$ .

Then

$$\begin{aligned} V_k(z) &= V_{T-m}(z) \\ &= \frac{T}{m(m+1)(T-m)} \sum_{j=0}^m \left[ \sum_{t=1}^{T-m} u_{t+j} \right]^2 \\ &= \frac{T}{m(m+1)(T-m)} \sum_{j=0}^m \left[ \sum_{i=1}^j u_i + \sum_{s=T-m+j+1}^T u_s \right]^2 \end{aligned} \quad (16)$$

---

<sup>18</sup>When  $d \neq 0$ , all the following steps in the proof hold true for  $\Delta \tilde{z}_t \equiv \Delta z_t - d$  after  $\Delta \tilde{z}_t$  replaces  $\Delta z_t$ .

The last equality holds because  $\sum_{t=1}^T u_t = \sum_{t=1}^T (\Delta z_t - \overline{\Delta z}) = 0$  :

$$\begin{aligned}
0 &= \sum_{t=1}^T u_t = \sum_{i=1}^j u_i + \sum_{t=1}^{T-m} u_{t+j} + \sum_{s=T-m+j+1}^T u_s \\
&\Rightarrow - \sum_{t=1}^{T-m} u_{t+j} = \sum_{i=1}^j u_i + \sum_{s=T-m+j+1}^T u_s \\
&\Rightarrow \left[ \sum_{t=1}^{T-m} u_{t+j} \right]^2 = \left[ \sum_{i=1}^j u_i + \sum_{s=T-m+j+1}^T u_s \right]^2.
\end{aligned} \tag{17}$$

Especially when  $m = 1$  or  $k = T - 1$ , from equation (16),

$$V_{T-1}(z) = \frac{T}{(1+1)(T-1)} \sum_{j=0}^1 \left[ \sum_{i=1}^j u_i + \sum_{s=T+j}^T u_s \right]^2 = \frac{T}{(T-1)} \frac{1}{2} [u_T^2 + u_1^2]. \tag{18}$$

The first term on the right hand side of equation (18) is

$$\begin{aligned}
\frac{T}{T-1} u_1^2 &= \frac{T}{T-1} \left( \Delta z_1 - \frac{1}{T} \sum_{s=1}^T \Delta z_s \right)^2 \\
&= \frac{T}{T-1} \left[ (\Delta z_1)^2 - \frac{2}{T} \sum_{s=1}^T \Delta z_1 \Delta z_s + \frac{1}{T^2} \left( \sum_{s=1}^T \Delta z_s \right)^2 \right].
\end{aligned} \tag{19}$$

By taking the unconditional expectation of equation (19),

$$\begin{aligned}
E \left( \frac{T}{T-1} u_1^2 \right) &= \frac{T}{T-1} \left( \gamma_0 - \frac{2}{T} \sum_{j=0}^{T-1} \gamma_j + \frac{1}{T} \sum_{j=-T+1}^{T-1} \frac{T-j}{T} \gamma_j \right) \\
&= \frac{T}{T-1} \left( \frac{T-1}{T} \gamma_0 - \frac{1}{T} \sum_{j=-T+1}^{T-1} \gamma_j + \frac{1}{T} \sum_{j=-T+1}^{T-1} \frac{T-j}{T} \gamma_j \right) \\
&= \gamma_0 - \frac{1}{T-1} \sum_{j=-T+1}^{T-1} \gamma_j + \frac{1}{T-1} \sum_{j=-T+1}^{T-1} \frac{T-j}{T} \gamma_j \\
&\rightarrow \gamma_0 - \frac{1}{T-1} \Omega + \frac{1}{T-1} \Omega \\
&\rightarrow \gamma_0 \text{ as } T \rightarrow \infty.
\end{aligned} \tag{20}$$

Finally, from equations (18) and (20),

$$\begin{aligned} E(V_{T-1}(z)) &= \frac{1}{2} \left[ E\left(\frac{T}{T-1}u_T^2\right) + E\left(\frac{T}{T-1}u_1^2\right) \right] \\ &\rightarrow \frac{1}{2}(\gamma_0 + \gamma_0) = \gamma_0 \quad \text{as } T \rightarrow \infty. \end{aligned} \quad (21)$$

In equation (18),  $V_{T-1}(z)$  is a function of  $u_t^2$  but not a function of any  $u_{t+j}u_t$ ,  $j \neq 0$ .  $V_1(z)$  is also a function of  $u_t^2$  as follows:

$$V_1(z) = \frac{T}{(T-1)T} \sum_{j=0}^{T-1} \left[ \frac{1}{T} \sum_{t=1}^T u_{t+j} \right]^2 = \frac{T}{(T-1)T} \frac{1}{T} \sum_{t=1}^T u_t^2. \quad (22)$$

By the law of large numbers,

$$V_1(z) \rightarrow \gamma_0 \quad \text{as } T \rightarrow \infty. \quad \blacksquare \quad (23)$$

Proposition 1 shows that the final value of  $E(V_k(z))$  goes back to its initial value as  $k$  varies from 1 to  $T-1$ . The proposition indicates that  $E(V_k(z))$  for the largest  $k$  has little to do with the long run movement of the variable since its limit is  $\gamma_0$ , which represents the shortest run movement of the variable. So, if we treat it as an estimator of the long run variance,  $V_{T-1}(z)$  has a severe bias.

Equation (18) shows that  $V_{T-1}(z)$  can be expressed only by  $u_t^2$ . No higher order sample autocovariance terms,  $u_t u_{t+\tau}$  ( $\tau \neq 0$ ), appear in equation (18). It indicates that  $E(V_{T-1}(z))$  is associated only with the shortest run movement of the variable. In the proof of the proposition, equation (17) is the key to derive equation (18). Equation (17) holds because the mean of the change is unknown and estimated by  $\overline{\Delta z}$ . So estimated unknown drift is a source of bias. The Bartlett kernel estimator also has such bias due to the estimated unknown drift term. Equation (10) shows the bias from the long run variance<sup>19</sup>. The bias grows bigger as  $\tau$  increases. The Bartlett kernel dampens the bias

<sup>19</sup>The Bartlett kernel estimator is a weighted sum of  $\widehat{\gamma}_\tau$ . The estimate of autocovariance,  $\widehat{\gamma}_\tau$ , is biased

by assigning smaller weights to higher order sample autocovariances, but the variance of  $k$ -differences reverses the effect by underweighting observations near both endpoints.<sup>20</sup>

$V_k(z)$  is continuous with respect to  $k$ . Proposition 1 implies that, when  $k$  is close to the sample size, there is a central tendency for  $V_k(z)$  to go back toward the initial value of  $V_k(z)$  no matter what DGP  $z_t$  follows. To get an idea about the quasi-symmetry of  $E(V_k(z))$  near both ends of the time horizon, let's find a similar expression to equations (21) and (22) for  $V_k(z)$  when  $k = 2$  and  $T - 2$ . From equation (9), when  $k = 2$ ,

$$\begin{aligned} V_2(z) &= \frac{T}{(T-1)} \frac{1}{(T-1)} \sum_{j=0}^{T-k} \left[ \sum_{t=1}^k u_{t+j} \right]^2 \\ &= \frac{T}{(T-1)} \frac{1}{(T-1)} \sum_{j=1}^{T-1} [u_j + u_{j+1}]^2 \end{aligned} \quad (24)$$

while, by equation (16) when  $k = T - 2$ ,

$$\begin{aligned} V_{T-2}(z) &= \frac{T}{(T-1)} \frac{1}{3} \sum_{j=0}^{T-k} \left[ \sum_{t=1}^k u_{t+j} \right]^2 \\ &= \frac{T}{(T-1)} \frac{1}{3} \{ [u_1 + u_2]^2 + [u_{T-1} + u_T]^2 + [u_1 + u_T]^2 \}. \end{aligned} \quad (25)$$

Hence,  $V_2(z)$  is associated with zero and the first order sample autocovariance term—namely,  $u_t^2$  and  $u_t u_{t+1}$ . On the other hand,  $V_{T-2}(z)$  is affected by  $u_t^2$ ,  $u_t u_{t+1}$  and  $u_1 u_T$ . Under the summability condition in Assumption 1,  $\sum_{j=0}^{\infty} j \cdot |\psi_j| < \infty$ , the high order autocovariance term,  $u_1 u_T$ , should be negligible on average. It hints that, for fixed and small  $m$ ,  $V_{T-m}(z)$  is mainly associated with the  $(m-1)^{th}$  or lower order autocovariance terms as is  $V_m(z)$ .

Engel (1999) infers the importance of the traded goods component in the long run

---

when the mean is unknown. See Theorem 6.2.2 in Fuller (1996) and Percival (1993) for more details.

<sup>20</sup>Campbell and Mankiw (1987) already warned that one must be careful not to misinterpret the behavior of  $V_k(z)$  as  $k$  increases to the point where it approaches  $T$  when the sample mean is used. However, their formula goes to zero instead of  $\gamma_0$  because it does not have the small sample correction term in equation (1).

movement of the US RER based on the asymptotics in equation (6). According to equation (6), for small and fixed  $k'$ s, the larger  $k$  is, the longer run movements of  $z_t$   $V_k(z)$  represents. Contrary to what equation (6) indicates, however, when  $k$  is close to the sample size,  $V_k(z)$  seems to get associated with lower order autocovariances as  $k$  increases up to the sample size.

### 2.3 Simulation results for the distribution of $V_k(z)$

Table 1 recapitulates our discussion so far on the limit of  $V_k(z)$  or the limit of its mean over various time horizons. Under Assumption 1,  $V_1(z)$  converges to  $\gamma_0$ . As  $k$  increases,  $V_k(z)$  converges to the long run variance of  $\Delta z_t$  when  $k/T \rightarrow 0$ . However, if  $k$  is big enough compared with the sample size resulting in  $k/T \rightarrow b > 0$ , then  $V_k(z)$  converges to a limit distribution and not to a number. In this case, we can show that, from equation (7), the mean of the limit distribution is the long run variance. Finally, as  $k$  gets close to the sample size, the mean of the limit distribution of  $V_k(z)$  goes back to its initial value,  $\gamma_0$ .

If  $z_t$  is a random walk, the limit of  $V_k(z)$  or the limit of its mean continues to be  $\gamma_0$  irrespective of  $k$ . This follows from  $\gamma_\tau = 0$  for each  $\tau \neq 0$ , implying that the long run variance of  $\Delta z_t$  is equal to the variance of the change,  $\gamma_0$ . If  $z_t$  is stationary, on the other hand, although the graph of the mean of  $V_k(z)$  starts from the same point ( $V_1(z) \rightarrow \gamma_0$ ), it will go down toward zero as  $k$  grows. When the variable is stationary, its long run variance of the first difference is zero. Thus, the limit distribution in this case degenerates to zero. Thus, even in case of  $k/T \rightarrow \delta > 0$ ,  $V_k(z)$  converges to zero. Later, as  $k$  grows close to the sample size, the mean of  $V_k(z)$  goes back toward the initial value,  $\gamma_0$ .



Table 1. The limit of  $V_k(z)$  or the limit of the mean of its distribution

	$k = 1$	$k$ is small and fixed	$k/T \rightarrow 0$	$k/T \rightarrow b > 0$	$k = T - 1$
general case	$\gamma_0$	$\sum_{\tau=-k+1}^{k-1} \frac{k -  \tau }{k} \gamma_\tau$	$\Omega$	$\Omega$	$\gamma_0$
random walk	$\gamma_0$	$\gamma_0$	$\gamma_0$	$\gamma_0$	$\gamma_0$
stationary	$\gamma_0$	$\sum_{\tau=-k+1}^{k-1} \frac{k -  \tau }{k} \gamma_\tau$	0	0	$\gamma_0$

\*  $\gamma_\tau$  is the  $\tau$ -th order autocovariance of  $\Delta z_t$ , and  $\Omega$  is the long run variance of  $\Delta z_t$

By means of a Monte Carlo simulation, we get the mean and 90% confidence intervals of  $V_k$  for each  $k = 1, 2, \dots, T - 1$  from 5,000 simulated series of pure random walks and a stationary AR(1) as in Figure 1.<sup>21</sup> In the graph, bold lines are the means of  $V_k(z)$  in the simulation while normal lines represent 90% two-sided confidence intervals. The solid lines are for the stationary AR(1) process, and the dotted lines are for the random walk process.

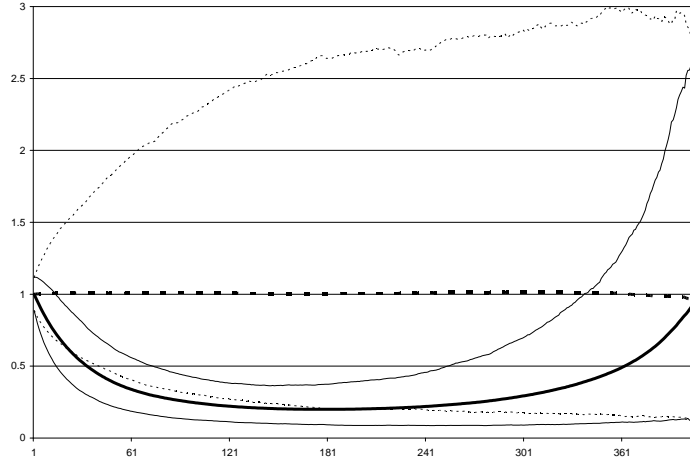
Figure 1 illustrates our findings in Proposition 1. The graph for the mean of  $V_k(z)$  for each DGP starts at its variance of the change and ends at the same value. Note especially that the mean of  $V_k(z)$  for random walk does not change much as we see in Table 1. On the other hand, the mean of  $V_k(z)$  for the stationary AR(1) shows as a U-shaped graph.

Then, next observation from Figure 1 is that the mean of  $V_k(z)$  for the stationary AR(1) has the closest value to its long run variance, zero, in the middle range of time horizons. The minimum of the mean of  $V_k(z)$  over different  $k$ 's is 0.2 at  $k = 179$ , a little less than half of the sample size. Hence,  $V_k(z)$  in the middle range of time horizons seems more relevant to the long run movement of the variable than that at the time horizons close to the sample size.

<sup>21</sup>In the simulation, the number of observation is 408 ( $t = 0, 1, 2, \dots, 407$ ) as in the first data set in Engel (1999). The AR(1) coefficient for the stationary process is set to be .94387 which implies that the half life is one year in a monthly data. The variance of the difference in each process is set to be one. The error terms in each series are assumed to be normal.

Figure 1

The distribution of the variance of  $k$ -differences



\* The solid line is for the  $AR(1)$ , and the dotted line is for the pure random walk.

★ Bold lines are the means, and normal lines are the 90% confidence intervals.

Another observation is that the mean of  $V_k(z)$  for the stationary  $AR(1)$  process even in the middle range of time horizons is clearly above its long run variance, i.e. zero. As an estimate of the long run variance,  $V_k(z)$  has a severe upward bias when the variable is stationary  $AR(1)$  even at the most relevant time horizons. Since  $V_k(z)$  in equation (1) is defined as a sum of squared terms, the value of  $V_k(z)$  in a finite sample should be always positive. Thus, the issue here is how close the value of  $V_k(z)$  is to the true long run variance. The simulation result shows that the 34 year time-span in Engel (1999) even with fairly short half life is not enough to get an estimate close to the true long run variance.

In terms of the width of the confidence intervals, the confidence interval for the stationary  $AR(1)$  process is much narrower than that for the random walk in the middle range of time horizons. Intuitively, the narrower confidence interval of the stationary  $AR(1)$  may be related to the fact that the limit distribution of  $V_k(x)$  does not converges to a nondegenerate distribution but goes to a number even when  $k/T \rightarrow b \geq 0$ .

The next observation for the confidence intervals is that the two distributions become indiscernible as  $k$  gets close to the sample size. Hence, the test based on  $V_k(z)$  for  $k$  close to the sample size will suffer from very low power with the null hypothesis of a random walk against the alternative hypothesis of a stationary AR(1).

We can compare this with simulation results for the Bartlett kernel estimator. As is apparent in equation (10), the mean of the Bartlett kernel estimator is getting smaller as  $b$  increases when the variable follows a random walk while the mean of the variance of  $k$ -differences remains constant because of the small sample correction term. To compare the two statistics, we divide the Bartlett kernel estimator by the terms in parentheses in equation (10). After the adjustment, we find no difference between the mean of the Bartlett kernel estimator and that of the variance of  $k$ -differences when the variable follows a pure random walk. On the other hand, the two are very different for large  $k$ 's in the case of a stationary AR(1).

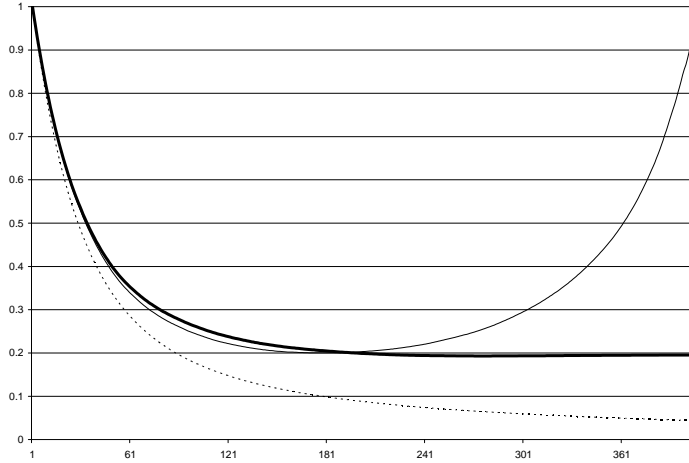
Figure 2 shows the mean of the simulation results for both the Bartlett kernel estimator after the adjustment and that for the variance of  $k$ -differences in the case of a stationary AR(1). In the figure, the bold solid line is the mean of the Bartlett kernel estimator, and the normal solid line is for the variance of  $k$ -differences. The dotted line is the mean of the population counterpart of the variance of  $k$ -differences<sup>22</sup>. Both the Bartlett kernel estimator and the variance of  $k$ -differences are above the population counterpart on average. There is not much difference between the Bartlett kernel estimator and the variance of  $k$ -differences for the first half of the time horizons. However, for the second half, the two statistics are very different. The Bartlett kernel estimator does not change much in this region while the variance of  $k$ -differences goes back to the initial level.

---

<sup>22</sup>If  $x_t = \rho x_{t-1} + \varepsilon_t$  with  $0 < \rho < 1$ ,  $\varepsilon_t \sim iid(0, \sigma_\varepsilon^2)$ ,  
 $E(V_k(x)) = \frac{2(1-\rho^k)}{k(1-\rho^2)}\sigma_\varepsilon^2 = \frac{(1-\rho^k)}{k(1-\rho)}\sigma_{\Delta x}^2$ .

Figure 2

Mean of  $V_k$  and mean of Bartlett kernel estimator for a stationary AR(1)



Another model for our simulation is an integrated AR(1) which is considered as a possible DGP of log stock price in Lo and MacKinlay (1988). An integrated AR(1) process with a positive AR coefficient is more persistent than a pure random walk. The integrated AR(1) model in Lo and MacKinlay is

$$\Delta z_t = \kappa \cdot \Delta z_{t-1} + \epsilon_t, \quad \text{where } \epsilon \sim i.i.d.N(0, \sigma_\epsilon^2) \text{ and } |\kappa| < 1. \quad (26)$$

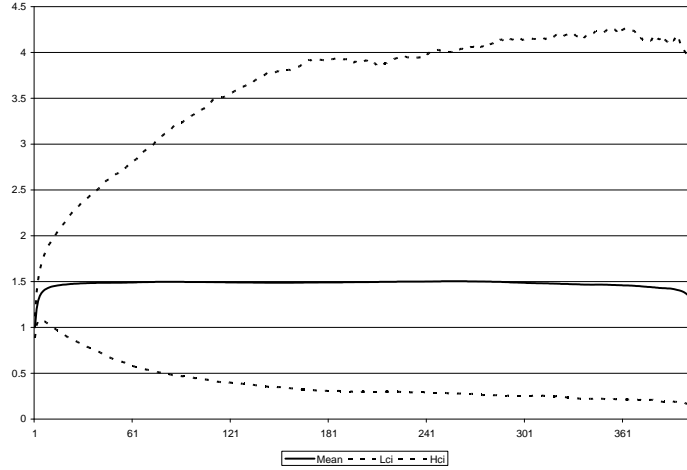
In the simulation,  $\kappa = .2$ ,  $\sigma_{\Delta z}^2 = 1$ , and the sample size is set to be 408 for comparison with the previous simulation. Figure 3 represents the simulation result.

Figure 3 also illustrates the result in Proposition 1. For  $k$  close to the sample size, the mean goes back to the variance of the change as  $k$  increases<sup>23</sup>. Unlike in our first simulation results, the mean of the variance of  $k$ -differences soon reaches the level of its long run variance and stays around this level throughout the middle time horizons.

<sup>23</sup>In this case, the long run variance of  $\Delta z_t$ ,  $\Omega$  is  $\sigma_\epsilon^2/(1 - \kappa)^2$ , whereas  $\gamma_0 = \sigma_\epsilon^2/(1 - \kappa^2)$ . Thus,  $\Omega/\gamma_0 = 1.5$ . In this example, for comparison with Figure 1,  $\gamma_0$  is set to be one. Then  $\Omega = 1.5$ .

Figure 3

The distribution of the variance of  $k$ -differences (ARIMA(1,1,0) Model)



\* Solid lines are the means, and dotted lines are the 90% confidence intervals.

So far, the error terms in the DGP are assumed to be normally distributed. We also performed the same simulation above assuming that the error terms follow t-distribution with 3 degrees of freedom. In this case, the mean of the graph is the same as in Figures 1 and 2. On the other hand, the confidence intervals for the short run time horizons are wider than those for the normal distribution case. However, as  $k$  increases, the confidence intervals converge to those in Figures 1 and 2.

In conclusion, first, our simulation results in Figures 1 and 2 show that the mean of  $V_k(z)$  does go back to  $\gamma_0$  as Proposition 1 states, irrespective of the DGP of  $z_t$  under Assumption 1. Second, in terms of its mean,  $V_k(z)$  reaches the closest point to the long run variance not in the end but in the middle of the time horizons. Third, while the Bartlett kernel estimator and  $V_k(z)$  are very close to each other in the first half of the time horizon, the two are quite different in the second half. Finally, there are two differences between the case of a stationary AR(1) and the case of an integrated AR(1). First, when  $k$  is around the middle of the sample size, the mean of  $V_k(z)$  for an integrated AR(1) is very close to the long run variance while that for a stationary

AR(1) has a severe upward bias. Second, the slope of the graph of the mean of  $V_k(z)$  for a stationary AR(1) is much less steep at both ends of the graph than that for an integrated AR(1).

### 3 Implication of Proposition 1 for Engel's ratio of $V_k(z)$

#### 3.1 Ratio of Variance of $k$ -differences in Engel (1999)

As in Engel (1999), we define the real exchange rate,  $q_t$ , as

$$q_t \equiv s_t + p_t^* - p_t. \quad (27)$$

where  $s_t$  is the log of the nominal exchange rate,  $p_t$  is the log of the general price index of the home country, and  $p_t^*$  is the log of the general price index of the foreign country.

Engel (1999) regards the log of the general price index as a weighted average of traded- and nontraded-goods prices:

$$p_t = (1 - \alpha)p_t^T + \alpha p_t^N, \quad (28)$$

$$p_t^* = (1 - \beta)p_t^{T*} + \beta p_t^{N*} \quad (29)$$

Superscripts  $T$  and  $N$  indicate traded and nontraded goods each. An asterisk represents the foreign country.  $\alpha$  and  $\beta$  are the shares of nontraded goods in each country's price index.

The RER can be decomposed by

$$q_t = x_t + y_t, \quad (30)$$

where

$$x_t \equiv s_t + p_t^{T*} - p_t^T, \quad (31)$$

$$y_t \equiv \beta(p_t^{N*} - p_t^{T*}) - \alpha(p_t^N - p_t^T). \quad (32)$$

$x_t$ , the traded goods component, is the relative price of traded goods between the two countries while  $y_t$ , the nontraded goods component, is a weighted difference of the relative prices of nontraded goods in each country.

Engel (1999) measures the importance of the traded goods component in explaining US RER movements with the ratio of the variance of  $k$ -differences of  $x_t$  over that of  $q_t$ ,  $RV_k$ :<sup>24</sup>

$$RV_k = \frac{Var(x_t - x_{t-k})}{Var(q_t - q_{t-k})} = \frac{V_k(x)}{V_k(x) + V_k(y)} \quad (33)$$

*assuming  $x$  and  $y$  are uncorrelated*

### 3.2 An illustration with highly tradable goods

In order to illustrate the implication of Proposition 1 for Engel's method, we first apply his method to data involving highly tradable goods for which the law of one price is likely to hold. Our purpose in this exercise is to show that the ratio of the variances of  $k$ -differences for stationary  $x_t$  is likely to have a U-shaped graph.

Burstein, Neves and Rebelo (2003) point out that distribution costs are so large for consumer goods that the law of one price may not hold at the retail price level. For this reason, Burstein, Eichenbaum and Rebelo (2005 and 2006) use the prices of pure-traded goods at the dock<sup>25</sup>. Following Burstein, Eichenbaum and Rebelo (2006), we measure the prices of traded goods using a geometric average of import and export prices and compute  $x_t$  in equation (31) with those prices. Then we construct data for  $y_t$  as the

---

<sup>24</sup>Engel (1999) mainly uses the ratio of the mean-squared errors (MSE), the sum of the squared drift and variance of  $k$ -differences, in order to measure the movement comprehensively. However, he states that the results based on the variance of  $k$ -differences are not very different from the results based on MSE's for US RER. His inference in the paper is based on the properties of the variance of  $k$ -differences. For simplicity we only consider the ratio of the variance of  $k$ -differences in this paper.

<sup>25</sup>Betts and Kehoe (2006) also show that the choice of the price series significantly affects the statistical measure of the relative importance of the traded goods component in the real exchange rate movement.

difference between the RER and  $x_t$ .<sup>26</sup> For the RER, the CPI general indexes for both countries are used.

The data are collected from the IFS CD ROM. The sample period is 1973:01-2002:12. Among the ten bilateral real exchange rates with the US in Burstein, Eichenbaum and Rebelo (2006), graphs for the US-Italy RER are presented in Figure 4 since our unit root test and stationarity test results consistently indicate that its traded goods component is likely to be stationary.

Figure 4 is a graph for  $V_k(x)$  of the US-Italy RER, and Figure 5 is for  $RV_k$ . As far as the long run movement of RER is concerned,  $V_k$  can be interpreted as an estimator of the long-run variance following the asymptotics in equation (6). According to the traditional Balassa-Samuelson theory, the numerator of  $RV_k$  should converge to zero since the long run variance of the traded good component, which is stationary, is zero. On the other hand, the denominator of  $RV_k$  must have a positive value because the nontraded good component is unit-root nonstationary. As a whole, therefore,  $RV_k$  should converge to zero as  $k$  increases at an appropriate rate as the sample size increases. In other words, the importance of the traded goods component should be small in the long run.

Given this result, it is tempting to interpret the rise of  $V_k$  and  $RV_k$  for large  $k$  in Figures 4 and 5 as evidence against the Balassa-Samuelson theory in the long-run. However, Proposition 1 shows that, for large  $k$ , the asymptotics in equation (6) are not applicable and that  $V_k$  has a tendency to go back to the initial level as  $k$  gets closer to the sample size irrespective of whether the variable is stationary or difference stationary. Due to the statistical properties of its numerator and denominator,  $RV_k$  also has a tendency to go back to the neighborhood of the initial level. Therefore, the rise of  $V_k$  and  $RV_k$  for large  $k$  is more likely due to the property of  $V_k$  specified in Proposition 1 than to the properties of US RER long run movements.

---

<sup>26</sup>So the decomposition of the RER is based on equation (3) in Engel (1999) rather than equation (1) in Engel's paper. In other words,  $y_t = (p_t^* - p_t^{T*}) - (p_t - p_t^T)$ .



Figure 4

Variiances of  $k$ -differences for  $x_t$  of US-Italy pure traded goods

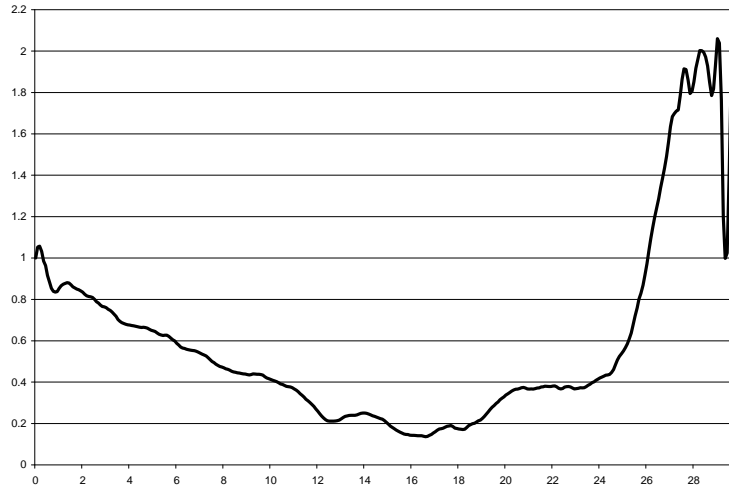
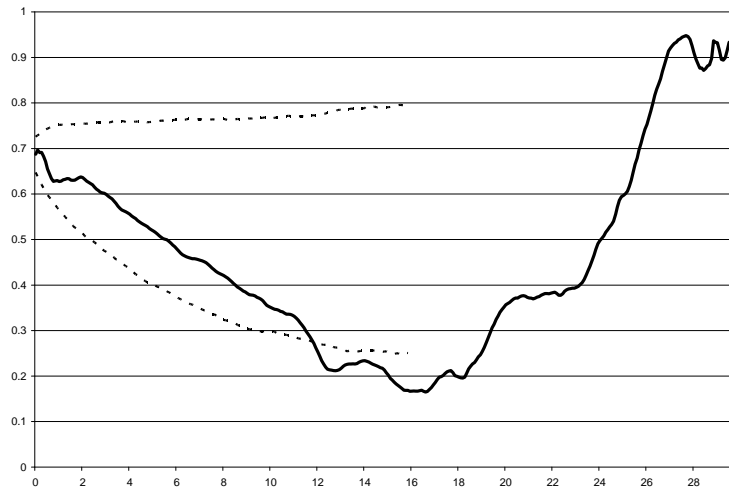


Figure 5

Ratio of  $V_k$  for US-Italy pure traded goods



If we focus on the fall of  $V_k$  and  $RV_k$  for the first half of the graph in these figures, the results with Engel's method are consistent with those with unit root tests in Table 2. The importance of the traded goods component in explaining the movement of the real exchange rate at each time horizon becomes smaller in the longer run. The graph

in Figure 5 is clearly in favor of stationarity of  $x_t$  since the solid line,  $RV_k$ , is, in the longer periods, under the lower dotted line which is the critical value of the null that  $x_t$  follows a random walk<sup>27</sup>. Thus, it is important not to interpret the rise of the ratio in the second half of Figure 5 as evidence against the Balassa-Samuelson theory in the long-run.

### 3.3 Reexamination of Engel's Empirical Results

We now reexamine Engel's results in light of our findings in this paper. Solid lines in Figure 6 plot the graphs of  $RV_k$  for the US RER computed from Engel's (1999) first data set in his paper<sup>28</sup>. For short time horizons, the ratios are all over 90%. For the middle range of time horizons, the ratios go down except for the US-Italy RER, although the magnitude of change varies from country to country. And finally, for long time horizons, the ratios move back to higher levels. The most prominent case is for the US-Canada RER.

Although, Engel refrains from reaching a decisive conclusion because of the small number of independent observations for large  $k$ 's, he interprets the rise of the graph for the US-Canada RER in longer time horizons as an increase in the importance of the traded goods component (p.513), implying that the traditional Balassa-Samuelson theory does not work even in the long run.

However, previous discussions in this paper show that  $V_k$ 's for  $k$ 's near the sample size have little to do with the long run movement of the variables. Thus, the rise of the graph for the US-Canada  $RV_k$ 's at large  $k$ 's may not be interpreted as an increase in the importance of the traded goods component for long run time horizons.

Instead, our simulation results indicate that  $V_k$ 's in the middle range of time horizons

---

<sup>27</sup>How we construct the critical value will be explained in detail in the next subsection.

<sup>28</sup>The data are monthly from January 1962 to December 1995 for Canada, France, Germany, Italy, Japan, and the United States. Thus it has 408 observations (so  $T = 407$ ). CPI's for goods are used for traded goods prices, and CPI's for services are used for nontraded goods prices. See Appendix A of Engel (1999) for more details.

are more relevant to the long run movement of the variable while  $V_k'$ s at both ends of the time horizon are associated with the short run movement. If  $x_t$  is AR(1) and  $y_t$  is a random walk, the graph for  $RV_k'$ s is likely to be U-shaped on average while the graph should be close to a flat line if both  $x_t$  and  $y_t$  are random walks. The graphs in Figure 6 show a U-shape except for US-Italy so that  $RV_k'$ s in the middle range have smaller value than those at both ends. It may imply that the traded good component becomes less important in the longer run in accounting for the movement of the US RER.

Although the graphs look U-shaped,  $RV_k'$ s in the middle range may not be statistically different from those at both ends. With  $RV_k'$ s computed from the data, Engel (1999) tries to test his null hypothesis that the law of one price for the traded goods does not hold. Since there is no standardized asymptotic distribution of  $RV_k$ , Engel uses a parametric bootstrap method to compute the confidence intervals of  $RV_k$ . Under his null, he supposes that both  $x_t$  and  $y_t$  are random walks with drift. Engel, then, shows that the  $RV_k$  at every time horizon in the data is within the two-sided 95% confidence interval of  $RV_k$ . As such, Engel does not find any evidence for a less important role for the traded goods component in the longer run movement of RER. Engel compares his results with Kakkar and Ogaki's (1999) which are in favor of an important role for the nontraded goods component in the long run. He attributes the difference to the low power of the tests to distinguish between unit roots and stationarity in relatively short time spans.

We believe that Engel's confidence interval is not tight enough and needs some adjustments. Those adjustments lead to a different conclusion from that in Engel (1999). We find that Engel's empirical results are not very robust. Our adjustments to Engel's testing method include the following:

First, we perform a one-sided test, as opposed to the two-sided test in Engel (1999). Our main interest in this paper is the long run movement of the RER. In the long run,  $V_k$  is the estimator of long run variance. If the law of one price for traded goods does

not hold in the long run, then  $x_t$  is nonstationary and the long run variance of  $\Delta x_t$  will have a positive value. On the other hand, if the law of one price holds in the long run, then  $x_t$  is stationary and the long run variance of  $\Delta x_t$  is zero. Therefore,  $RV_k$  under the null that both  $x_t$  and  $y_t$  are random walks should be statistically larger than that under the alternative hypothesis. Thus, lower dotted line in Figure 6 is the critical value under the null that both  $x_t$  and  $y_t$  are random walk. That is, if  $RV_k$  from the data is lower than the confidence intervals, then the test rejects the null.

Second, we report the confidence interval only up to half of the sample size while Engel (1999) reports up to the largest possible  $k$ . Since Cochrane (1988), it has been known that the variance of  $k$ -differences for large  $k$  is not reliable. Admitting the inaccuracy of the statistics in his paper, Engel reports it for the entire time horizon probably because he believes that a small piece of information about the long run is better than no information. However, since our findings indicate that  $RV_k$  for large  $k$  has little to do with the long run, there is not much gain from reporting inaccurate test results for large  $k$ 's.

Third, we do not allow drift either in  $x_t$  or in  $y_t$ . While the literature on the Balassa-Samuelson theory has given some models which allow drift in  $y_t$ , it is difficult to find a model which explains why  $x_t$  may have drift. Engel does not provide a theoretical explanation for it either.

Fourth, with the same testing method, it is easy to flip the null. We can compute the confidence interval under the null that  $x_t$  is stationary and that  $y_t$  is still a random walk. When  $RV_k$  from the data is above the confidence interval, the test rejects the null. An additional problem in this case is how to specify the data generating process of  $x_t$ . Apparently, there are many different kinds of stationary processes. We report the simplest case in which  $x_t$  is  $AR(1)$  in this paper.

The dotted lines in Figure 6 are critical values for one-sided tests with 5% size after the changes we make. The lower dotted line is the critical value under the null that  $x_t$  is

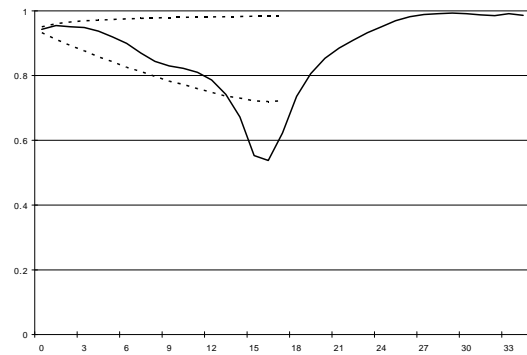
a random walk without drift while the upper dotted line is the critical value under the null that  $x_t$  is stationary  $AR(1)$ . Overall, the graphs at many time horizons are within the two dotted lines, indicating the low power problem pointed out by Engel (1999). However, unlike the results in Engel (1999),  $RV'_k$ 's for long time horizons are below the lower dotted line, rejecting the null that both  $x_t$  and  $y_t$  are random walk for Canada, France, and Germany. In the case of Italy and Japan, on the other hand,  $RV'_k$ 's for short time horizons are above the upper dotted line, rejecting the null that  $x_t$  is  $AR(1)$  and  $y_t$  is a random walk.

To sum up, the test based on  $RV'_k$ 's computed from the data and bootstrap critical values does not necessarily support Engel's null hypothesis but provides some evidence for smaller importance of the traded goods component in accounting for longer run RER movement although the evidence is not as clear as the result for highly tradable goods in the previous subsection.

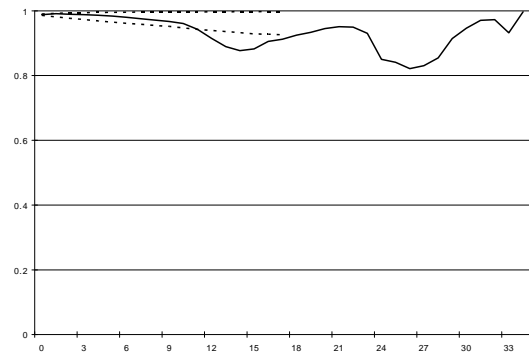
Figure 6

Ratio of the variances of  $k$ -differences for the US RER(1962:01-1995:12)

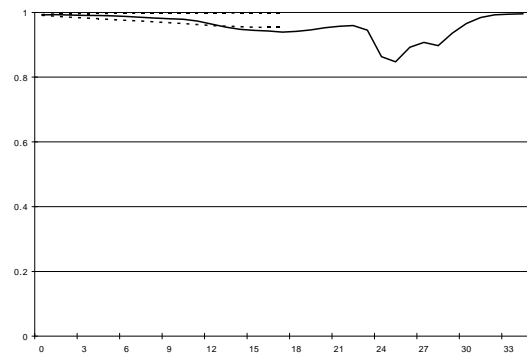
a) Canada



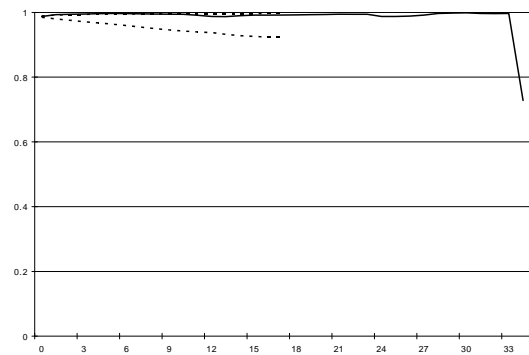
b) France



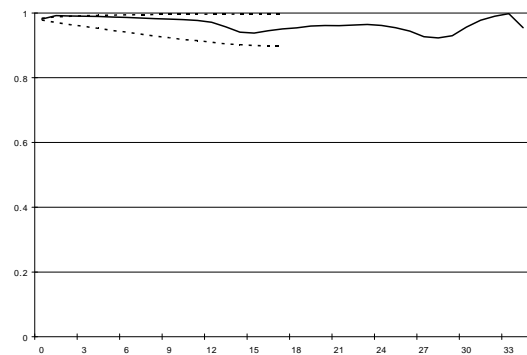
c) Germany



d) Italy



e) Japan



## 4 Conclusion

According to the traditional Balassa-Samuelson view, the traded goods component of the real exchange rate is stationary while the nontraded goods component has a unit root. The long run variance of the traded goods component is zero while the real exchange rate itself has a positive long run variance because of the unit root in the nontraded goods component.

Cochrane (1988) shows that the variance of  $k$ -differences ( $V_k$ ) is asymptotically equivalent to the Bartlett kernel estimator of long run variance. Engel (1999) computes  $V_k$  for the real exchange rate and for its traded goods component. Based on Cochrane (1988), Engel expected that the ratio of  $V_k$  of the traded goods component to  $V_k$  of the real exchange rate would converge to zero as  $k$  increases if the traditional theory were true.

In contrast to the traditional view, Engel finds that the ratios decrease at first but increase at the end of the time horizons, most prominently in case of the US-Canada RER. Engel interprets this as an increase in the importance of the traded goods component at longer run time horizons. Based on the empirical results, Engel concludes that the behavior of the traded goods component is indistinguishable from the behavior of a random walk.

This paper, however, shows that the mean of the variance of  $k$ -differences for the largest  $k$ ,  $V_{T-1}$ , converges to the limit of the variance of the first difference,  $V_1$ . Therefore, if  $V_k$  falls as  $k$  increases,  $V_k$  tends to rise as  $k$  approaches  $T - 1$  irrespective of whether the variable of interest is stationary or unit root nonstationary. This means that the rise of the graph at  $k$  close to the sample size in Engel (1999) cannot be interpreted as evidence for unit root nonstationarity of the traded goods component in the real exchange rate.

While  $V_k$  for  $k$  close to the sample size does not reflect the long run properties of the variable, the simulation results in the paper show that  $V_k$  will get closer to the long

run variance as  $k$  increases from one to time horizons in the middle range. The ratio of  $V_k$  in Engel (1999) decreases in the first half of time horizons, which indicates that the nontraded goods component plays a more important role in the longer run.

Engel (1999) show that the  $RV'_k$ s from the data are all within the confidence intervals he constructs under the null that there is no change in the importance of the nontraded goods component over different time horizons. On the contrary, after some adjustments of the testing method, our test results provide some evidence consistent with a more important role for the nontraded goods component at longer time horizons for some countries.

Cochrane (1988) pointed out that the variance of  $k$ -differences for large  $k$  is less reliable. He explains that the degrees of freedom of  $V_k$  are roughly equal to the number of nonoverlapping long runs, which is less than two when  $k$  is more than half of the sample size. Considering the inaccuracy due to low degrees of freedom, Cochrane (1988) reports his results at time horizons only up to one fourth of the sample size. Lo and MacKinlay (1988) also report their simulation results at time horizons up to half of the sample size. It is exceptional in the literature to report up to the longest time horizon as in Engel (1999), and Engel admits that his longer run horizon numbers are less reliable, probably based on Cochrane's (1988) argument.

What is new in this paper, though, is that  $V_k$  for  $k$  close to the sample size not only has a big variance due to its low degrees of freedom but also has little to do with the long run movement of the variable.



## References

- [1] Alexius, Annika, Nilsson, Jonny, 2000, Real exchange rates and fundamentals: evidence from 15 OECD countries. *Open Economies Review* 11, 383-397.
- [2] Andrews, Donald W.K., 1991. Heteroskedasticity and autocorrelation consistent covariance matrix estimation. *Econometrica* 59, 817-858.
- [3] Balassa, Bela, 1964. The purchasing power parity doctrine: a reappraisal. *Journal of Political Economy* 72, 584-596.
- [4] Betts, Caroline M., Kehoe, Timothy J., 2006, U.S. real exchange rate fluctuations and relative price fluctuations. *Journal of Monetary Economics* 53, 1297-1326.
- [5] Burstein, Ariel, Eichenbaum, Martin, Rebelo, Sergio, 2005. Large devaluations and the real exchange rate. *Journal of Political Economy* 113, 742-784.
- [6] Burstein, Ariel, Eichenbaum, Martin, Rebelo, Sergio, 2006. The importance of nontradable goods' prices in cyclical real exchange rate fluctuations. *Japan and the World Economy* 18, 247-253.
- [7] Burstein, Ariel, Neves, João, Rebelo, Sergio, 2003. Distribution costs and real exchange rate dynamics during exchange-rate-based stabilizations. *Journal of Monetary Economics* 50, 1189-1214.
- [8] Cai, Ye, Shintani, Mototsugu, 2006, On the alternative long-run variance ratio test for a unit root. *Econometric Theory* 22, 347-372.
- [9] Campbell, John Y., Mankiw, Gregory, 1987. Are output fluctuations transitory? *The Quarterly Journal of Economics* 102, 857-880.
- [10] Chinn, Menzie D., 2006, A primer on real effective exchange rates: determinants, overvaluation, trade flows and competitive devaluation. *Open Economies Review* 17, 115-143.

- [11] Cochrane, J. H., 1988. How big is the random walk in GNP? *Journal of Political Economy* 96, 893-920.
- [12] Crucini, Mario J., Shintani, Mototsugu, 2002. Persistence in law-of-one-price deviations: evidence from micro-data. Working Papers 02-W22R, Department of Economics, Vanderbilt University.
- [13] Crucini, Mario J., Telmer, Chris I, Zachariadis, Marios, 2005. Understanding European real exchange rates. *American Economic Review* 95, 724-738.
- [14] Deo, Rohit S., Richardson, Matthew, 2003, On the asymptotic power of the variance ratio test. *Econometric Theory* 19, 231-239.
- [15] Davidson, James, 1994, *Stochastic limit theory*. Oxford University Press.
- [16] Engel, Charles, 1999, Accounting for U.S. real exchange rate changes. *Journal of Political Economy* 107, 507-38.
- [17] Fuller, Wayne A., 1996, *Introduction to statistical time series*, 2nd edition. John Wiley & Sons.
- [18] Isard, Peter, 1977, How far can we push the 'law of one price'? *American Economic Review* 67, 942-948
- [19] Hamilton, James D., 1994, *Time Series Analysis*. Princeton University Press.
- [20] Kakkar, Vikas, Ogaki, Masao, 1999, Real exchange rates and nontradables: a relative price approach. *Journal of Empirical Finance* 6, 193-215.
- [21] Kiefer, Nicholas M., Vogelsang, Timothy J., 2002, Heteroskedasticity-autocorrelation robust standard errors using the Bartlett kernel without truncation. *Econometrica* 70, 2093-2095.

- [22] Kiefer, Nicholas M., Vogelsang, Timothy J., 2005, A new asymptotic theory for heteroskedasticity-autocorrelation robust tests. *Econometric Theory* 21, 1130-1164.
- [23] Kim, Jaebeom, 2005, Convergence rates to purchasing power parity for traded and nontraded goods: a structural error-correction model approach. *Journal of Business and Economic Statistics* 23, 76-86
- [24] Kim, Jaebeom, Ogaki, Masao, 2004, Purchasing power parity for traded and non-traded goods: a structural error correction model approach. *Monetary and Economic Studies* 22, 1-25.
- [25] Lo, Andrew W., MacKinlay, A. Craig, 1988, The size and power of the variance ratio test in finite samples: a Monte Carlo investigation. NBER Technical Working Papers: 66.
- [26] Lo, Andrew W., MacKinlay, A. Craig, 1999, *A non-random walk down Wall Street*, Princeton University Press.
- [27] Newey, W. K., West, K. D., 1987, A simple, positive semi-definite, heteroskedasticity and autocorrelation consistent covariance matrix. *Econometrica* 55, 703-08.
- [28] Obstfeld, Maurice, 2001, International macroeconomics: beyond the Mundell-Fleming model. *IMF Staff Papers*: 47, Special issue.
- [29] Parsley, David C., 2001, Accounting for real exchange rate changes in East Asia. Hong Kong Institute for Monetary Research Working Paper No.6/2001.
- [30] Percival, Donald B., 1993, Three curious properties of the sample variance and autocovariance for stationary processes with unknown mean. *The American Statistician* 47, 274-276.
- [31] Richardson, Matthew, Stock, James H., 1989, Drawing inferences from statistics based on multi-year asset returns. *Journal of Financial Economics* 25, 323-348.

- [32] Rogoff, Kenneth, 1996, The Purchasing Power Parity Puzzle. *Journal of Economic Literature* 34, 647-668.
- [33] Samuelson, Paul A., 1964, Theoretical notes on trade problems. *Review of Economics and Statistics* 46, 145-154.
- [34] Sarno, Lucio, Taylor, Mark P., 2002, Purchasing power parity and the real exchange rate. *IMF Staff Papers*: 49, No 1.
- [35] Sarno, Lucio, Valente, Giorgio, 2006, Deviations from purchasing power parity under different exchange rate regimes: do they revert and, if so, how? *Journal of Banking & Finance* 30, 3147-3169.
- [36] Schnatz, Bernd, Vajselaar, Focco, Osbat, Chiara, 2004, Productivity and the euro-dollar exchange rate. *Review of World Economics* 140, 1-30.
- [37] Sun, Yixiao, Phillips, Peter C. B., Jin, Sainan, 2006, Optimal bandwidth selection in heteroskedasticity-autocorrelation robust testing. *Yale University, Cowles Foundation Discussion Papers*: 1545.
- [38] Taylor, Alan M., Taylor, Mark P., 2004, The purchasing power parity debate. *Journal of Economic Perspectives* 18, 135-158.