Journal of International Money and Finance 30 (2011) 572–586

Contents lists available at ScienceDirect

Journal of International Money and Finance

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

A re-examination on dissecting the purchasing power parity puzzle

Jyh-Lin Wu [a,b,*], Chingnun Lee [a], Tzu-Wei Wang [c]

[a] Institute of Economics, National Sun Yat-Sen University, Kaohsiung, Taiwan
[b] Department of Economics, National Chung-Cheng University, Chia-Yi, Taiwan
[c] Department of Finance and Risk Management, Shu-Te University, Kaohsiung, Taiwan

JEL classification:
F31

Keywords:
Purchasing power parity puzzle
Real exchange rates
Local projections
Impulse-response function
Half-life

A B S T R A C T

The purpose of this paper is to examine the source of a real exchange-rate adjustment based on the impulse-response function constructed from local projections when the true data-generating process (DGP) is unknown. This work extends the local-projection method proposed by Jordà [2005. Estimation and inference of impulse responses by local projections. American Economic Review 95, 161–182] to allow for variables that are $I(1)$ and exhibit cointegration. Our paper shows that nominal exchange-rate adjustments dominate in the reversion toward PPP regardless of a nominal exchange-rate shock or a price shock. It is also shown that the half-life of real exchange rates is close to that of nominal exchange rates. Since these results are consistent with those of Cheung et al. [Cheung, Y.W., Lai, K.S., Bergman, M., 2004. Dissecting the PPP puzzle: the unconventional roles of nominal exchange rate and price adjustments. Journal of International Economics 64, 135–150], we therefore conclude that their main findings are robust to possible misspecifications in the true DGP.

1. Introduction

For decades, the purchasing power parity (PPP) hypothesis has been a major research field within empirical international finance. Existing empirical evidence based on unit-root tests provided mixed results for the long-run validity of PPP (Mark, 1990; Abuaf and Jorion, 1990; Taylor and Sarno, 1998;

* Corresponding author. Institute of Economics, National Sun Yat-Sen University, Kaohsiung, Taiwan.
E-mail addresses: ecdjlw@ccu.edu.tw, jlwu2@mail.nsysu.edu.tw (J.-L. Wu).

0261-5606/$ – see front matter
doi:10.1016/j.jimonfin.2011.01.004

O'Connell, 1998; Cheung and Lai, 1993; Michael et al., 1997). Although several articles appear to support the mean reversion of real exchange rates, their estimates of the half-life range from 3 to 5 years (Rogoff, 1996). These half-life estimates seem too long to be explained by sticky-price models which predict a half-life of 1–2 years. This is the PPP puzzle outlined by Rogoff (1996).

Since the seminal work of Rogoff (1996), the adjustment mechanism of PPP deviations has become an interesting issue in literature. Some articles examine the dynamic adjustment of real exchange rates, but the literature about the adjustment mechanism of nominal exchange rates and prices resulting from PPP deviations is limited. Two articles investigate the adjustment mechanism of PPP deviations. First, Engel and Morley (2001) use a state-space model to show that, while it takes years for nominal exchange rates to converge, it only takes a few months for prices to do so. Next, Cheung et al. (2004) use a linear vector error-correction model (VECM) to reach a similar conclusion. Cheung et al. (2004) also conclude that "about 60–90% of real exchange-rate convergence is driven by nominal exchange rates and the contribution of prices to PPP convergence is relatively minor". Their findings sharply contrast with conventional explanations based on nominal rigidities and therefore offer a new look at theoretical modeling of real exchange-rate dynamics.

Cheung et al. (2004) apply a VECM model to construct impulse-response functions (IRFs) given that nominal exchange rates and prices are $I(1)$ but cointegrated. The conventional VAR-VECM approach assumes that the true data-generating process (DGP) follows a vector autoregressive (VAR) process. A reduced rank VAR specification may also be misspecified in a set of cointegrated variables, since an error-correction model may be formed by a reduced rank vector autoregressive moving average (VARMA) specification (Engle and Granger, 1987).[1] Hence, the construction of IRFs based on the VAR-VECM model is not robust to misspecifications in the true DGP.

Jordà (2005) proposes a local-projection approach that uses multi-step direct forecasts as an alternative to a VARMA transformation.[2] Furthermore, he shows that IRF estimates derived from local projections are consistent and asymptotically normal. One advantage of Jordà's local projections is that they allow IRFs to be constructed without knowing the true DGP.[3] In other words, the IRFs constructed from local projections are robust to misspecifications in the true DGP.

Unfortunately, Jordà's local-projection method cannot be directly applied to cases where variables are $I(1)$, since the standard multi-step forecast and projection theorem requires the existence of the second moment of variables.[4] Lin and Tsay (1996) find, based on monthly financial and macroeconomic data of six major economies, that direct forecasting outperforms VECM-based forecasting in the presence of unknown unit roots and cointegration – even though unit roots and cointegration are ignored in direct forecasting. Theoretically, cointegration implies the existence of a long-run equilibrium among the variables, which could be exploited to improve the accuracy of long-term forecasts. The simulations of Lin and Tsay (1996) show the long-term forecasting benefit of modeling cointegrating relationships among variables, if they exist. Therefore, having the correct knowledge of cointegrating relationships among variables improves the accuracy of forecasts.

The purpose of our paper is to re-examine the adjustment mechanism of prices and nominal exchange rates based on IRFs constructed from local projections. This paper extends Jordà's local-projection method to obtain IRFs in a model where variables are integrated of order one, $I(1)$, but with

1 Let $\mathbf{y}_t$ be an $n \times 1$ random vector such that all components of $\mathbf{y}_t$ are $I(1)$ and that there exists a vector $\boldsymbol{\beta}$ such that $\boldsymbol{\beta}'\mathbf{y}_t \sim I(0)$. Engle and Granger (1987) show that there exists an error-correction representation with $\mathbf{z}_t = \boldsymbol{\beta}'\mathbf{y}_t$, an $r \times 1$ vector of stationary random variables: $\mathbf{A}^*(L)(1-L)\mathbf{y}_t = -\boldsymbol{\alpha}\boldsymbol{\beta}'\mathbf{y}_{t-1} + \mathbf{d}(L)\epsilon_t$. It is worth noting that the above vector error-correction model is not a VAR model unless $\mathbf{d}(L) = \mathbf{I}$.

2 Jordà (2005) establishes the equivalence of IRFs calculated by local projections and VAR, respectively, when the true DGP follows a VAR process.

3 When variables under investigation are $I(0)$, a standard VAR approach is applied to construct an IRF. The conventional approach may not be appropriate if one is interested in the true dynamics of underlying process. This is because VAR specifications are typical parsimonious and likely misspecified and hence a VAR(k) model is not likely to be the exact representation of the process under study (Chang and Sakata, 2007). In addition, an unappealing feature of conventional IRF estimates is that these estimates are constructed by extrapolating the parsimonious VAR model which will typically impose smooth decay in the estimated IRF and therefore excludes potentially high-order dynamics of underlying process.

4 To be specific, it requires that variables are in a Hilbert space, $L^2(\Omega, \mathcal{F}, P)$ (Weiss, 1991; Brockwell and Davis, 1991, Chapter 2; Ing, 2003).

cointegration among them. The key idea of our extension is to project the differenced variables on lag differenced variables and cointegrating residuals which are constructed based on a theoretical implication. The IRFs of level variables are then constructed from the accumulation of all related projecting coefficients that were used.

This work finds that nominal exchange-rate adjustments dominate the reversion to PPP regardless of shocks, and that the half-life of real exchange rates is close to that of nominal exchange rates. These results are consistent with those of Cheung et al. (2004). Given this agreement, we hypothesize that the interesting findings of Cheung et al. (2004) are robust to possible misspecifications in the true DGP.

The organization of the paper is given as follows. Section 2 extends Jordà's local-projection method to the case where variables are cointegrated $I(1)$ variables. Section 3 shows that projection errors follow a moving average process when the true DGP is a cointegrated VAR process – a finding that can be used to improve estimation efficiency. Section 4 estimates the extent to which real exchange-rate evolution can be explained by either nominal exchange rates or prices through our modified local-projection method. The 90% confidence interval of the relative contribution of nominal exchange-rate and price adjustments in the reversion toward PPP is also constructed. In addition, this work similarly constructs the half lives of nominal and real exchange rates and prices, along with each half-life's 90% confidence interval. Finally, concluding remarks are given in Section 5.

2. Local-projection-based IRFs when variables are a cointegrated *I*(1) process

An impulse-response function measures the time profile of the effect of shocks on the (expected) future value of variables in a dynamic system. According to Koop et al. (1996), the generalized impulse-response function of $\mathbf{y}_t$ at horizon h is defined as follows:

$$IR(t,h,\delta,\mathbf{\Omega}_{t-1}) = E(\mathbf{y}_{t+h}|\mathbf{v}_t = \delta,\mathbf{\Omega}_{t-1}) - E(\mathbf{y}_{t+h}|\mathbf{v}_t = \mathbf{0},\mathbf{\Omega}_{t-1}), \quad h = 1,2,\ldots, \tag{1}$$

where δ is a $n\times 1$ vector indicating the shocks, $\mathbf{0}$ is a $n\times 1$ vector of zeroes, $\mathbf{v}_t$ is a $n\times 1$ vector of additive random shocks, $\mathbf{\Omega}_{t-1}$ denotes the information set including values of variables up to $t-1$, and $E(\cdot|\cdot)$ denotes the best mean square predictor. If both $\mathbf{y}_t$ and the elements of $\mathbf{\Omega}_{t-1}$ are in Hilbert space $L^2(\Omega,\mathcal{F},P)$, then the impulse-response function defined in equation (1) based on linear projection can be easily obtained by projecting $\mathbf{y}_{t+h}$ onto a linear function of elements of $\mathbf{\Omega}_{t-1}$. This is the idea of Jordà's (2005) local-projection method.

But Jordà's (2005) method cannot be so simply applied when the variables under investigation are $I(1)$ (that is $\mathbf{y}_t \notin L^2$) and have a cointegration relationship. To solve this problem, we now define the evolution of the first difference of $\mathbf{y}_t$ as follows. Consider the Hilbert space $L^2(\Omega,\mathcal{F},P)$ defined on an abstract probability space and let $\Delta\mathbf{y}_t = \mathbf{y}_t - \mathbf{y}_{t-1} = (\Delta Y_{t,1},\Delta Y_{t,2},\ldots,\Delta Y_{t,n})'$, $\beta'\mathbf{y}_t = (Z_{t,1},Z_{t,2},\ldots,Z_{t,r})'$, $t=1,2,\ldots,T$ be a sequence of n and r $(r<n)$ dimensional stochastic vector, respectively. Both $\Delta\mathbf{y}_t$ and $\beta'\mathbf{y}_t$ are in $L^2(\Omega,\mathcal{F},P)$ and are adapted to the information set $\{\mathbf{\Omega}_t, t=1,2,\ldots,T\}$, where $\mathbf{\Omega}_t = \sigma(\Delta\mathbf{y}_1,\ldots,\Delta\mathbf{y}_t,\beta'\mathbf{y}_t)$ is the sigma-field on which $\Delta\mathbf{y}_t$ is measurable. That is, $\mathbf{y}_t$ is a $n\times 1$ $I(1)$ random vector with cointegrating matrix β such that $\beta'\mathbf{y}_t$ is a $r\times 1$ $I(0)$ random vector. In such a case, modeling a cointegrating relationship among variables is helpful because cointegrating residuals provide useful information for improving forecasts (Engle and Yoo, 1987; Reinsel and Ahn, 1988). In this respect, this paper considers projecting $\Delta\mathbf{y}_{t+h}$ onto the linear space generated by $(\beta'\mathbf{y}_{t-1}, \Delta\mathbf{y}_{t-1}, \Delta\mathbf{y}_{t-2}, \ldots, \Delta\mathbf{y}_{t-k+1})'$. Specifically, this work defines:[5]

$$\mathbf{P}(\Delta\mathbf{y}_{t+h}|\beta'\mathbf{y}_{t-1},\Delta\mathbf{y}_{t-1},\Delta\mathbf{y}_{t-2},\ldots,\Delta\mathbf{y}_{t-k+1}) = (P_{h_k}\Delta Y_{t+h,1},P_{h_k}\Delta Y_{t+h,2},\ldots,P_{h_k}\Delta Y_{t+h,n})', \tag{2}$$

where

$$P_{h_k}\Delta Y_{t+h,i} = c_i + \sum_{j=1}^{r} a_{ij}Z_{t-1,j} + \sum_{m=1}^{n} b_{im}^{(1)}\Delta Y_{t-1,m} + \cdots + \sum_{m=1}^{n} b_{im}^{(k-1)}\Delta Y_{t-k+1,m}, \quad \forall i = 1,\ldots,n.$$

By the projection theorem, the set of coefficients c_i, a_{ij} and $b_{im}^{(1)}$, …, $b_{im}^{(k-1)}$ all exist and satisfy the following two equations for all $i=1, 2, \ldots, n$:

[5] Vector projection is just the collection of scalar projection in a vector (Brockwell and Davis, 1991, p. 421).

$$E(\Delta Y_{t+h,i} - P_{h_k}\Delta Y_{t+h,i})Z_{t-1,j} = 0, \quad \forall j = 1,\ldots,r,$$

$$E(\Delta Y_{t+h,i} - P_{h_k}\Delta Y_{t+h,i})\Delta Y_{t-l,m} = 0, \quad \forall l = 1,\ldots,k-1; \;\; \forall m = 1,\ldots,n.$$

We next denote the forecast error, $\mathbf{u}_{t+h}^{(h)}$, by:

$$\mathbf{u}_{t+h}^{(h)} = \Delta\mathbf{y}_{t+h} - \mathbf{P}(\Delta\mathbf{y}_{t+h}|\beta'\mathbf{y}_{t-1}, \Delta\mathbf{y}_{t-1}, \Delta\mathbf{y}_{t-2}, \ldots, \Delta\mathbf{y}_{t-k+1}). \tag{3}$$

Equation (2) can now be substituted into equation (3) to obtain:

$$\Delta\mathbf{y}_{t+h} = \mathbf{c}^{(h)} + \mathbf{A}^{(h+1)}\beta'\mathbf{y}_{t-1} + \mathbf{B}_1^{(h+1)}\Delta\mathbf{y}_{t-1} + \cdots + \mathbf{B}_{k-1}^{(h+1)}\Delta\mathbf{y}_{t-k+1} + \mathbf{u}_{t+h}^{(h)}, \quad h = 1,\ldots,s, \tag{4}$$

where

$$\mathbf{c}^{(h)} = [c_i]_{n\times 1}, \;\; \mathbf{A}^{(h+1)} = [a_{ij}]_{n\times r}, \;\; \mathbf{B}_l^{(h+1)} = \left[b_{im}^{(l)}\right]_{n\times n}, \;\; l = 1,\ldots,k-1.$$

It is worth noting that the forecast error, $\mathbf{u}_{t+h}^{(h)}$, is not necessarily a white noise process.

According to equation (1), the IRF of $\Delta\mathbf{y}_t$ from the local linear projection ($\mathrm{IR}_\Delta^{\mathrm{LP}}$) based on equation (4) is:[6]

$$IR_\Delta^{LP}(t,h,\delta,\mathbf{\Omega}_{t-1}) = \left(\mathbf{A}^{(h)}\beta' + \mathbf{B}_1^{(h)}\right)\delta, \quad h = 1,2,\ldots s. \tag{5}$$

And, since $\mathbf{y}_{t+h} = \Delta\mathbf{y}_{t+h} + \Delta\mathbf{y}_{t+h-1} + \cdots + \Delta\mathbf{y}_{t+1} + \mathbf{y}_t$, then, by the principle of linearity inherent to the linear projection function, the IRF of $\mathbf{y}_t$ ($\mathrm{IR}^{\mathrm{LP}}$) is

$$IR^{LP}(t,h,\delta,\mathbf{\Omega}_{t-1}) = (\mathbf{I} + \mathbf{E}_h\beta' + \mathbf{G}_h)\delta = \mathbf{\Phi}_h\delta, \quad h = 1,2,\ldots,s,$$

where $\mathbf{E}_h = \sum_{j=1}^{h}\mathbf{A}^{(j)}$ and $\mathbf{G}_h = \sum_{j=1}^{h}\mathbf{B}_1^{(j)}$ are the matrices of cumulative effects.[7] The impulse response of the ith variable to a shock of the jth variable at horizon h corresponds to the (i,j) elements of the $n \times n$ matrix, $\mathbf{\Phi}_h$. Following Jordà (2005), the consistent (but possibly not the most-efficient) estimator of $\mathrm{IR}^{\mathrm{LP}}(t,h,\delta,\mathbf{\Omega}_{t-1})$, $\widehat{\mathrm{IR}}^{\mathrm{LP}}(t,h,\delta,\mathbf{\Omega}_{t-1})$, for a given δ, is:

$$\widehat{IR}^{LP}(t,h,\delta,\mathbf{\Omega}_{t-1}) = \left(\mathbf{I} + \widehat{\mathbf{E}}_h\beta' + \widehat{\mathbf{G}}_h\right)\delta, \quad h = 1,2,\ldots,s,$$

where $\widehat{\mathbf{E}}_h = \sum_{j=1}^{h}\widehat{\mathbf{A}}^{(j)}$, $\widehat{\mathbf{G}}_h = \sum_{j=1}^{h}\widehat{\mathbf{B}}_1^{(j)}$, and $\widehat{\mathbf{A}}^{(j)}$ and $\widehat{\mathbf{B}}_1^{(j)}$, $h = 1,2,\ldots s$, are estimates obtained by regressing equation (4) for each horizon h by the least-square method.

One way to derive the statistical reliability of the above impulse-response estimates is to apply an asymptotic normal approximation. To construct the asymptotic confidence interval of the impulse-response function of $\mathbf{y}_t$ from a linear projection, under a given δ, we need the covariance matrices of $\widehat{\mathbf{A}}^{(h)}$ and $\widehat{\mathbf{A}}^{(g)}$, $\widehat{\mathbf{B}}_1^{(h)}$ and $\widehat{\mathbf{B}}_1^{(g)}$, and $\widehat{\mathbf{A}}^{(h)}$ and $\widehat{\mathbf{B}}_1^{(m)}$ for $h, g, m = 1,\ldots s$. Jordà (2005, 2009) adopts a heteroscedastic and autocorrelated (HAC) consistent estimator to estimate the asymptotic variance–covariance matrix of $\widehat{\mathrm{IR}}_\Delta^{\mathrm{LP}}$ by stacking s local-projection equations in (4).[8,9] The above mentioned estimator is given as follows:

[6] Because of the assumption that errors are additive and that $\mathbf{\Omega}_{t-1}$ is already known by time t, we can see that both $\Delta\mathbf{y}_t$ and $\mathbf{y}_t$ will be equally affected by a shock at time t.

[7] Using the fact that $\mathbf{y}_t - E(\mathbf{y}_t|\mathbf{\Omega}_{t-1}) = \Delta\mathbf{y}_t - E(\mathbf{y}_t|\mathbf{\Omega}_{t-1})$, the unnormalized forecasted mean-squared error (MSE) would be $\mathrm{MSE}(E(\mathbf{y}_{t+h}|\mathbf{\Omega}_{t-1})) = E[(\mathbf{u}_{t+h}^{(h)} + \mathbf{u}_{t+h-1}^{(h-1)} + \cdots + \mathbf{u}_{t+1}^{(1)} + \mathbf{u}_t^{(0)})(\mathbf{u}_{t+h}^{(h)} + \mathbf{u}_{t+h-1}^{(h-1)} + \cdots + \mathbf{u}_{t+1}^{(1)} + \mathbf{u}_t^{(0)})']$. This result is useful for computing the forecast-error variance decomposition.

[8] The lag order (k) in equation (4) is assumed to be the same at each horizon h for notational simplicity.

[9] The block on the diagonal is the same as the HAC variance–covariance estimates obtained by regressing equation (4) for each horizon h with a least-square method. The above derivation is not reported here but is available upon request from authors.

$$\widehat{\mathrm{Var}}\left(\mathrm{vec}\left(\widehat{\mathbf{\Psi}}\right)\right) = \left[(\mathbf{I}\otimes\mathbf{X})'(\mathbf{I}\otimes\mathbf{X})\right]^{-1}(\mathbf{I}\otimes\mathbf{X})'\widehat{\mathbf{\Sigma}}(\mathbf{I}\otimes\mathbf{X})\left[(\mathbf{I}\otimes\mathbf{X})'(\mathbf{I}\otimes\mathbf{X})\right]^{-1}, \tag{6}$$

where (ignoring the constant terms):

$$\mathbf{\Psi} = \begin{bmatrix} \mathbf{A}'^{(1)} & \mathbf{A}'^{(2)} & \dots & \mathbf{A}'^{(s)} \\ \mathbf{B}_1'^{(1)} & \mathbf{B}_1'^{(2)} & \dots & \mathbf{B}_1'^{(s)} \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ \mathbf{B}_{k-1}'^{(1)} & \mathbf{B}_{k-1}'^{(2)} & \dots & \mathbf{B}_{k-1}'^{(s)} \end{bmatrix}, \quad \mathbf{X}_t = \left[(\boldsymbol{\beta}'\mathbf{y}_t)', (\Delta\mathbf{y}_{t-1})', \dots, (\Delta\mathbf{y}_{t-k+1})'\right],$$

and $\mathbf{\Sigma} = \mathrm{Var}(\mathrm{vec}(\mathbf{u}'_{t+1}, \dots, \mathbf{u}'_{t+s}))$. The construction of the asymptotic variance–covariance matrix of $\widehat{\mathrm{IR}}^{\mathrm{LP}}$ is straightforward after estimating the matrix of $\widehat{\mathrm{Var}}(\mathrm{vec}(\widehat{\mathbf{\Psi}}))$. For example, let $\mathbf{\Psi}^* = \begin{bmatrix} \mathbf{A}^{(1)}\boldsymbol{\beta}' + \mathbf{B}_1^{(1)} \\ \vdots \\ \mathbf{A}^{(s)}\boldsymbol{\beta}' + \mathbf{B}_1^{(s)} \end{bmatrix}$, then the standard deviation of the ith variable to a shock of the jth variable at horizon h is available from the summation of the square root of the appropriate diagonal and non-diagonal entries of the covariance matrix, $\widehat{\mathrm{Var}}(\mathrm{vec}(\widehat{\mathbf{\Psi}}^*))$.

Note that, although the above variance–covariance matrix only reflects the uncertainty associated with the slope parameter estimates, this is not a fundamental limitation. The tentative assumption, in this paper, that δ is constant, was only made for illustrative purposes to facilitate the illustration of how to construct the asymptotic variance of impulse-response functions. Consequently, a component that incorporates the estimation uncertainty associated with the estimate of δ has been ignored (see Kilian and Kim, 2009). Also note that our assumption of a constant δ is innocuous to our empirical results, because our confidence intervals (discussed in Section 4 below) are simulated based on a bootstrap that re-constructs $\widehat{\delta}^*$ for each replication.

3. Asymptotic equivalence to VAR-VECM when DGP is a cointegrated VAR process

Jordà (2005) shows that the linear projection approach will be asymptotically equivalent to the VARMA approach if the data-generating process is a stationary VAR process. In this section, this paper extends Jordà's result by showing that there is an asymptotic equivalence between the linear projection approach and the VAR-VECM approach under the case that the data-generating process is a non-stationary but cointegrated VAR process.

Consider a n-dimensional linear vector autoregressive data-generating process of order k:

$$\mathbf{y}_t = \mathbf{\Phi} + \mathbf{\Pi}_1\mathbf{y}_{t-1} + \mathbf{\Pi}_2\mathbf{y}_{t-2} + \cdots + \mathbf{\Pi}_k\mathbf{y}_{t-k} + \epsilon_t, \quad t = 1, \dots, T, \tag{7}$$

where $\mathbf{\Phi}$ is a vector of constants and $\mathbf{\Pi}_i$ is a matrix of coefficients for $i = 1,\dots,k$. The error term, ϵ_t, is assumed to be identically, independently and normally distributed with a zero mean and a constant variance. The initial values, $\mathbf{y}_0$, $\mathbf{y}_{-1}$, …, $\mathbf{y}_{-k+1}$, are taken as given. It is well known that equation (7) can be re-written in a vector error-correction form:

$$\Delta\mathbf{y}_t = \mathbf{\Pi}\mathbf{y}_{t-1} + \sum_{i=1}^{k-1}\mathbf{\Gamma}_i\Delta\mathbf{y}_{t-i} + \mathbf{\Phi} + \epsilon_t, \quad t = 1, \dots, T, \tag{8}$$

where $\mathbf{\Pi} = -\mathbf{I} + \sum_{i=1}^{k}\mathbf{\Pi}_i = \boldsymbol{\alpha}\boldsymbol{\beta}'$ with $\mathrm{rank}(\mathbf{\Pi}) = r < n$ and $\mathbf{\Gamma}_i = -\sum_{j=i+1}^{k}\mathbf{\Pi}_j$. This paper assumes that all roots of the determinant of $\mathbf{A}(z) \equiv \mathbf{I} - \mathbf{\Pi}_1 z - \mathbf{\Pi}_2 z^2 - \cdots - \mathbf{\Pi}_k z^k = 0$ are either outside the unit circle or are equal to one. Such an assumption ensures that $\mathbf{y}_t$ is integrated of order one and cointegrated.[10]

The current research rewrites equation (8) to a linear first-order system:

[10] Please refer to Johansen (1991) and Hansen (2005) for detailed discussion.

$$\mathbf{W}_t = \mathbf{F}\mathbf{W}_{t-1} + \mathbf{v}_t, \tag{9}$$

with:

$$\mathbf{W}_t = \begin{bmatrix} \mathbf{y}_t \\ \Delta\mathbf{y}_t \\ \Delta\mathbf{y}_{t-1} \\ \vdots \\ \Delta\mathbf{y}_{t-k+2} \end{bmatrix}; \ \mathbf{F} = \begin{bmatrix} \alpha\beta' + \mathbf{I} & \mathbf{\Gamma}_1 & \mathbf{\Gamma}_2 & \cdots & \mathbf{\Gamma}_{k-2} & \mathbf{\Gamma}_{k-1} \\ \alpha\beta' & \mathbf{\Gamma}_1 & \mathbf{\Gamma}_2 & \cdots & \mathbf{\Gamma}_{k-2} & \mathbf{\Gamma}_{k-1} \\ \mathbf{0} & \mathbf{I} & \mathbf{0} & \cdots & \mathbf{0} & \mathbf{0} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \cdots & \mathbf{I} & \mathbf{0} \end{bmatrix}; \ \mathbf{v}_t = \begin{bmatrix} \mathbf{\Phi} + \epsilon_t \\ \mathbf{\Phi} + \epsilon_t \\ \mathbf{0} \\ \vdots \\ \mathbf{0} \end{bmatrix}.$$

Equation (9) is the companion form of the VECM, which allows us to construct h-ahead forecasts as follows:

$$\mathbf{W}_{t+h} = \mathbf{v}_{t+h} + \mathbf{F}\mathbf{v}_{t+h-1} + \cdots + \mathbf{F}^h\mathbf{v}_t + \mathbf{F}^{h+1}\mathbf{W}_{t-1}.$$

The second equation of (9) states:

$$\begin{aligned}\Delta\mathbf{y}_{t+h} = {} & (\mathbf{\Phi} + \epsilon_{t+h}) + (\mathbf{F}^1_{2,1} + \mathbf{F}^1_{2,2})(\mathbf{\Phi} + \epsilon_{t+h-1}) + \cdots + (\mathbf{F}^h_{2,1} + \mathbf{F}^h_{2,2})(\mathbf{\Phi} + \epsilon_t) + \mathbf{F}^{h+1}_{2,1}\mathbf{y}_{t-1} \\ & + \mathbf{F}^{h+1}_{2,2}\Delta\mathbf{y}_{t-1} + \mathbf{F}^{h+1}_{2,3}\Delta\mathbf{y}_{t-2} + \cdots + \mathbf{F}^{h+1}_{2,k}\Delta\mathbf{y}_{t-k+1},\end{aligned} \tag{10}$$

where ${\mathbf{F}_{i,j}}^h$ is the (i,j)th $n \times n$ block of the matrix $\mathbf{F}^h$ (i.e., $\mathbf{F}$ raised to the power of h). The Appendix of this paper shows that ${\mathbf{F}_{2,j}}^{h+1} \rightarrow \mathbf{0}$ as $h \rightarrow \infty$, $\forall j = 1,2,\ldots,k$. Hence, the infinite vector-moving-average representation of $\Delta\mathbf{y}_t$ can be derived as follows:

$$\Delta\mathbf{y}_t = \gamma + \epsilon_t + (\mathbf{F}^1_{2,1} + \mathbf{F}^1_{2,2})\epsilon_{t-1} + \cdots + (\mathbf{F}^h_{2,1} + \mathbf{F}^h_{2,2})\epsilon_{t-h} + \cdots. \tag{11}$$

The impulse-response function of $\Delta\mathbf{y}_t$ from equation (11) is therefore given by:

$$IR^{MA}_{\Delta}(t, h, \delta, \mathbf{\Omega}_{t-1}) = (\mathbf{F}^h_{2,1} + \mathbf{F}^h_{2,2})\delta, \quad h = 1, 2, \ldots, s. \tag{12}$$

Equation (10) can be used to establish the relationship between the impulse-response function calculated by local projections and that by the conventional VAR-VECM method. To justify the above statement, this paper rewrites equation (10) as follows:

$$\begin{aligned}\Delta\mathbf{y}_{t+h} = {} & [\mathbf{I} + (\mathbf{F}^1_{2,1} + \mathbf{F}^1_{2,2}) + \cdots + (\mathbf{F}^h_{2,1} + \mathbf{F}^h_{2,2})]\mathbf{\Phi} + \mathbf{F}^{h+1}_{2,1}\mathbf{y}_{t-1} + \mathbf{F}^{h+1}_{2,2}\Delta\mathbf{y}_{t-1} + \mathbf{F}^{h+1}_{2,3}\Delta\mathbf{y}_{t-2} + \cdots \\ & + \mathbf{F}^{h+1}_{2,k}\Delta\mathbf{y}_{t-k+1} + [\epsilon_{t+h} + (\mathbf{F}^1_{2,1} + \mathbf{F}^1_{2,2})\epsilon_{t+h-1} + \cdots + (\mathbf{F}^h_{2,1} + \mathbf{F}^h_{2,2})\epsilon_t].\end{aligned} \tag{13}$$

Comparing equations (4) and (13), this paper obtains the following equalities:

$$\begin{aligned}\mathbf{c}^{(h)} &= [\mathbf{I} + (\mathbf{F}^1_{2,1} + \mathbf{F}^1_{2,2}) + \cdots + (\mathbf{F}^h_{2,1} + \mathbf{F}^h_{2,2})]\mathbf{\Phi}, \\ \mathbf{A}^{(h+1)}\beta' &= \mathbf{F}^{h+1}_{2,1}, \\ \mathbf{B}^{(h+1)}_1 &= \mathbf{F}^{h+1}_{2,2}, \\ \mathbf{u}^{(h)}_{t+h} &= \epsilon_{t+h} + (\mathbf{F}^1_{2,1} + \mathbf{F}^1_{2,2})\epsilon_{t+h-1} + \cdots + (\mathbf{F}^h_{2,1} + \mathbf{F}^h_{2,2})\epsilon_t.\end{aligned} \tag{14}$$

Equation (14) establishes the equivalence of impulse-response functions constructed from local projections and from the VAR-VECM approach, when the DGP follows a cointegrated VAR process as in equation (7). It is worth noting that the error term from local projections, $\mathbf{u}^{(h)}_{t+h}$, is a moving average of forecast errors in equation (7) from time t to $t+h$. This special structure of the error term from local projections allows GLS estimation of the system block by block (Jordà, 2005).

4. Empirical investigation

Monthly data for consumer price indices and exchange rates (foreign currencies per US dollar) over the period from April 1973 through December 1998 have been downloaded from the IMF's international financial statistics. Four industrial countries are considered in our empirical investigation: France, Germany, Italy and the United States. The data source, sample period and countries in the

Table 1
Results from unit-root tests.

	DF-GLS		
	FRN/US	GER/US	ITA/US
s_t	−1.172	−0.389	0.384
p_t	−0.675	0.114	0.190
q_t	−2.139**	−1.945*	−2.133**

Notes: FRN, GER and ITA denote France, Germany and Italy, respectively. "*" and "**" indicate significance at the 10% and 5% level, respectively. DF-GLS is the unit-root statistic provided by Elliott et al. (1996). The lag order of the model is selected based on the MAIC rule.

sample are the same as those in Cheung et al. (2004). Real exchange rates are defined as the difference between the nominal exchange rates and the relative price, i.e. $q_t = s_t - (p_t{}^* - p_t^{us})$, where s_t and q_t are the nominal and real exchange rates, respectively, and where p_t^{us} and p_t^* are the price level in the US and a foreign country, respectively.

The purpose of this section is to apply the method of local projections, discussed in Section 2, to investigate the source of real exchange-rate adjustment toward PPP. This work first examines the unit-root hypothesis of nominal exchange rates, prices ($p_t \equiv p_t{}^* - p_t^{us}$) and real exchange rates based on the DF-GLS statistic provided by Elliott et al. (1996). The lag order of the model is selected based on the modified AIC rule. Results from Table 1 point out that the unit-root hypothesis of nominal exchange rates and prices is not rejected at conventional levels. This paper then applies Johansen's (2002) Bartlett-corrected trace test to determine the existence of a long-run relationship between nominal exchange rates and prices.[11] The lag order of the VAR system for France, Germany and Italy is 5, 3 and 3, respectively, which is consistent with what was found in Cheung et al. (2004). Results from Table 2 indicate that the hypothesis of no cointegration is rejected but the hypothesis of the existence of at least one cointegrating vector is not rejected at conventional levels for all countries. The above results support the existence of a long-run relationship between prices and nominal exchange rates. This work then examines whether the cointegrating vector of nominal exchange rates and relative prices is not significantly different from (1, −1). Results from Table 2 point out that the restriction of the cointegrating vector being (1, −1) is rejected, at conventional levels, for all countries.

It is worth noting that if nominal exchange rates are cointegrated with relative prices and the cointegrating vector is (1, −1) then real exchange rates must be stationary. Froot and Rogoff (1995) suggest that the most direct way to examine the long-run PPP specification is to perform unit-root tests on real exchange rates. Results displayed in the third row of Table 1 indicate that the unit-root hypothesis of real exchange rates is rejected at conventional levels. This indicates that nominal exchange rates are cointegrated with prices and the cointegrating vector is not significantly different from (1, −1). Accordingly, this paper supports that the long-run PPP condition hold and hence replaces the cointegrating residual of the model by real exchange rates.

To examine the sources of real exchange-rate adjustments, IRFs based on local projections, described in Section 2, are first constructed. To obtain the largest sample size possible in a given data set, the current research applies moment estimators to estimate the coefficients of the local-projection equation, rather than OLS estimators as proposed by Jordà (2005).[12] The method provided by Pesaran and Shin (1998) is applied to scale the hypothesized vector of shocks, δ^*, and hence our results are invariant to the ordering of variables in the projection vector.

[11] Johansen (2002) points out that the limiting distribution of the conventional trace statistic is often a poor approximation to the finite sample distribution, and hence proposes a Bartlett-corrected trace test to improve the finite sample performance of the conventional trace test.

[12] The projection coefficients can be calculated from a linear transformation of the first and second moments of variables (Hamilton, 1994, p. 86). In projecting a variable h-period ahead conditional on its $k-1$ most recent observations, the OLS estimators are constructed based on the first $h+(k-1)$ observations and therefore only use the sample size of $T-h-(k-1)$ to estimate all theoretical moments in a data set of T observations (Hamilton, 1994, p. 75). This paper adopts moment estimators obtained by equating theoretical and sample moments and therefore will have different sample size on each individual moment estimator. OLS and moment estimators have the same limit law (Brockwell and Davis, 1991, p. 240 and 262; and Fuller, 1996, p. 313).

Table 2
Results from cointegration tests.

Hypothesis	Statistics	FRN [5]	GER [3]	ITA [3]
A. Cointegration tests on s_t and p_t				
$r=0$	BTr	41.01**	33.90**	46.78**
$r=1$	BTr	4.17	2.43	3.33
B. Testing the hypothesis of the cointegrating vector being (1, −1)				
CV = (1, −1)	LR	25.02**	10.62**	14.91**

Note: FRN, GER and ITA denote France, Germany and Italy, respectively. $r=0$ and $r=1$ indicate the hypothesis of existing at least 0 and 1 cointegrating vector, respectively. CV = (1, −1) indicates the hypothesis of the cointegrating vector being (1, −1). BTr is the Barlett corrected trace statistic provided by Johansen (2002). LR is the likelihood ratio statistic of testing the cointegrating vector being (1, −1), which has a χ^2 distribution. A number in a bracket is the lag order of the VAR system. "**" indicates significance at the 5% level.

Following the suggestion of Kilian and Kim (2009), this paper computes the confidence interval of estimates by a block bootstrap method (Künsch, 1989, and Liu and Singh, 1992).[13,14] To preserve the correlation across variables in data, the overlapping blocks of l consecutive $(k+1)$ tuples of $(\Delta\mathbf{y}_{t+h}, \boldsymbol{\beta}'\mathbf{y}_{t-1}, \Delta\mathbf{y}_{t-1}, \ldots, \Delta\mathbf{y}_{t-k+1})$ are drawn with replacement. The bootstrapped sample is used to construct $\sum_{j=1}^{h}\widehat{\mathbf{A}}^{*(j)}$ and $\sum_{j=1}^{h}\widehat{\mathbf{B}}_1^{*(j)}$ (and hence $\widehat{\mathbf{E}}_h^*$ and $\widehat{\mathbf{G}}_h^*$) in equation (5).[15] For each bootstrap replication, this paper constructs $\widehat{\boldsymbol{\delta}}^*$ based on Pesaran and Shin (1998). Kilian and Kim (2009) indicate that it is not clear, in literature, how to estimate the variance of the bootstrapped local-projection impulse-response estimators for constructing the studentized t^* statistic. Instead of constructing a symmetric percentile-t confidence interval, this paper constructs a nominal $(1-\alpha)$% confidence interval by the percentile method which is defined as: $P(\widehat{\theta}^{\mathrm{LP}*}_{\alpha/2} \le \theta \le \widehat{\theta}^{\mathrm{LP}*}_{1-\alpha/2}) = 1-\alpha$, where $\widehat{\theta}^{\mathrm{LP}*}_{\alpha/2}$ and $\widehat{\theta}^{\mathrm{LP}*}_{1-\alpha/2}$ are the $\alpha/2$ and $1-\alpha/2$ quantiles of the distribution of $\widehat{\theta}^{\mathrm{LP}*}$. The block size is set to eight for all horizons, since this will produce the smallest local-projection confidence intervals amongst all of the different possible block sizes between 3 and 15.[16]

This paper decomposes real exchange-rate dynamics by constructing the impulse-response function of nominal exchange rates, prices and real exchange rates under both a price and nominal exchange-rate innovations. Since real exchange rates are defined as the difference between nominal exchange rates and prices, the relationship between the IRFs of nominal and real exchange rates and prices is given as follows:

$$IR_{q_{it}} = IR_{s_{it}} - IR_{p_{it}}, \tag{15}$$

where the subscript i indicates a nominal exchange-rate shock $(i=1)$ or a price shock $(i=2)$. Taking the first difference of equation (15), this work measures the relative contribution of nominal exchange rates and prices as follows:

$$\xi_{pi} = 1 - \xi_{si},$$

where $\xi_{si} \equiv \Delta\mathrm{IR}_{s_{it}}/\Delta\mathrm{IR}_{q_{it}}$, $\xi_{pi} \equiv -\Delta\mathrm{IR}_{p_{it}}/\Delta\mathrm{IR}_{q_{it}}$, and where ξ_{si} and ξ_{pi} each represent the proportion of the real exchange-rate evolution explained by, respectively, the nominal exchange rate and the price.

[13] Instead of generating the confidence interval of impulse-response estimates, this paper constructs the confidence interval of the relative contribution of nominal exchange rates and relative prices and that of half lives. These confidence intervals are nonlinear functions of the impulse-response coefficient estimates, and hence the delta expansion of the asymptotic distribution of the impulse-response estimator is complicate. Because of this, the current research adopts a block bootstrap to construct the confidence interval of the relative contribution of the nominal exchange rate and price and the confidence interval of half lives.

[14] Although our estimators are based on sample moments, Romano and Thombs (1996) show that block bootstrap resampling methods can be used to approximate the distribution of sample moment estimators under weak assumptions.

[15] One may apply a bias-corrected method to estimate parameters and then construct the confidence interval of estimates accordingly, but Kilian and Kim (2009) point out that the improvement to the bias-corrected method is limited.

[16] This paper also constructs a local-projection bootstrapped confidence interval by varying the block size with horizons, but results show that a fixed block size of eight produces the smallest confidence interval.

Since the sum of relative contribution to nominal exchange rates and prices is equal to one ($\xi_{pi}+\xi_{si}=1$), this work only reports the impact of a shock on nominal exchange rates, to save space. If ξ_{si} is greater than 0.5 but less than 1.0, then the current research claims that the nominal exchange-rate adjustment dominates in the reversion toward PPP after facing the *i*th shock. This indicates that, facing a positive shock, the price and nominal exchange-rate adjustments work in opposite directions. The price adjusts to correct the PPP deviation, but the nominal exchange rate moves further away from equilibrium driving the real exchange rate away from its long-run equilibrium level. If ξ_{si} is greater than 1, then ξ_{pi} must be negative since $\xi_{si}+\xi_{pi}=1$. This indicates that when facing a positive shock, nominal exchange-rate and price adjustments work in the same direction, driving the real exchange rate away from its long-run equilibrium level. Furthermore, the nominal exchange-rate adjustment dominates the price adjustment in driving real exchange rates.

Table 3 reports the estimated proportion of real exchange-rate evolution explained by the nominal exchange rate ($\widehat{\xi}_{si}$), the 90% confidence interval of the proportion, and the standard deviation of the estimated proportion under both a nominal exchange rate and a price shock. The above standard deviation and confidence interval are constructed based on 1000 block bootstrap replications with the block size being set to eight for all horizons. Results from the third and eighth columns of Table 3 indicate that the nominal exchange-rate adjustment generally dominates the reversion toward PPP regardless of shocks, since $\widehat{\xi}_{si}>0.5$ for all but two cases.[17] Given a nominal exchange-rate shock, however, results from the fourth column of Table 3 fail to reject the hypothesis of $\xi_{s1}=0$ (percentile confidence interval contains zero) in one, five and nine out of nine horizons for France, Germany and Italy, respectively. Moreover, ξ_{s2} is insignificant, regardless of horizons, for Italy. The standard deviation of $\widehat{\xi}_{s1}$ (SE_1) under different horizons ranges from 2.4 to 19.7 for France, 3.1 to 49.3 for Germany and 3.7 to 68.1 for Italy. Given a price shock, the standard deviation of $\widehat{\xi}_{s2}$ (SE_2) under different horizons varies from 1.5 to 148.7 for France, 1.6 to 31 for Germany and 3.6 to 25.8 for Italy. The standard deviations of $\widehat{\xi}_{s1}$ and $\widehat{\xi}_{s2}$ under different horizons are much larger than those founded in Cheung et al. (2004).

Due to the second power term in estimating a variance from sample average, the value of the standard error can be arbitrarily large when there are extreme values in the sample (Kim and White, 2004). Therefore, the above large standard deviations and wide 90% confidence intervals could be due to the existence of few extreme values of $\widehat{\xi}_{si}$.[18] After detailed inspection of the histogram of the 1000 bootstrap estimates of ξ_{si}, this paper finds few extreme estimates of ξ_{si}.[19] Hence, the current research removes these extreme estimates and then re-constructs the standard deviation of $\widehat{\xi}_{si}$ and the 90% confidence interval of ξ_{si}. Given a nominal exchange-rate shock, results from the sixth column of Table 3 indicate that ξ_{s1} is significantly different from zero at all horizons regardless of countries.[20] Results from the seventh column of Table 3 indicate that the standard deviation of $\widehat{\xi}_{s1}$ (SE_1*) under different horizons varies around 0.27, 0.31 and 0.50 for France, Germany and Italy, respectively. This paper only removes about one-tenth of $\widehat{\xi}_{s1}$ (according to the criterion in footnote 19) regardless of horizons and countries but the reduction in the standard deviation of $\widehat{\xi}_{s1}$ is substantial as indicated by the fifth and seventh columns of Table 3. Similar results are obtained under a price shock. Although the standard errors of $\widehat{\xi}_{si}$ decreases after removing the extreme values, they are still larger than those found in Cheung et al. (2004). Kilian and Kim (2009) point out that the estimated variance of local-projection-based estimates is larger than that of VAR-based estimates, and hence our results are consistent with theirs.

Based on a VAR-VECM framework, Cheung et al. (2004) point out that real exchange-rate convergence is driven by nominal exchange rates and the contribution of prices to PPP convergence is relatively minor. Our results from Table 3 are generally consistent with theirs, since ξ_{si} is significant (after

[17] The estimated proportions of real exchange-rate evolution explained by the nominal exchange rate from a price shock, ξ_{s2}, is −2.84 for FRN at the horizon of 10. This is due to the sharp fluctuation of the impulse-response function for nominal and real exchange rates indicated by Fig. 1.

[18] The problem of existing extreme bootstrap estimates is not serious in Kilian and Kim (2009) but it is worth noting that their variables are $I(0)$ but ours are $I(1)$.

[19] Based on the box plot used in descriptive statistics, an extreme value of $\widehat{\xi}_{si}$ is detected if it is less than Q_1-3IQ or larger than Q_2+3IQ, where Q_1 is the lower quartile and Q_2 is the upper quartile and $IQ=(Q_2-Q_1)$ is the interquartile.

[20] Few of our estimates lie outside their confidence intervals since this paper constructs percentile intervals rather than centered percentile used by some authors such as Kazimi and Brownstone (1999).

Table 3
Relative contributions of nominal exchange rate and price adjustments to PPP reversion based on local projections.

	h	A nominal exchange-rate innovation					A price innovation				
		$\widehat{\xi}_{s1}$	CI_1	SE_1	CI_1*	SE_1*	$\widehat{\xi}_{s2}$	CI_2	SE_2	CI_2*	SE_2*
FRN/US	10	1.11	(−0.04, 1.91)	6.72	**(0.53, 1.49)**	0.28	−2.84	**(0.17, 1.78)**	2.08	**(0.45, 1.61)**	0.33
	30	0.89	**(0.06, 2.04)**	4.38	**(0.49, 1.45)**	0.28	2.01	**(0.22, 1.81)**	5.57	**(0.58, 1.48)**	0.26
	50	0.16	**(0.01, 1.83)**	2.72	**(0.54, 1.48)**	0.27	0.76	**(0.03, 1.83)**	3.48	**(0.46, 1.43)**	0.28
	70	0.64	**(0.17, 2.06)**	7.63	**(0.47, 1.45)**	0.28	0.95	**(0.29, 1.71)**	1.56	**(0.51, 1.42)**	0.26
	90	1.04	**(0.16, 1.73)**	9.36	**(0.52, 1.40)**	0.26	0.57	**(0.13, 2.01)**	12.69	**(0.53, 1.46)**	0.28
	130	1.11	**(0.13, 2.04)**	3.38	**(0.56, 1.44)**	0.26	0.97	**(0.04, 1.85)**	148.70	**(0.55, 1.41)**	0.25
	170	1.49	**(0.07, 1.81)**	2.48	**(0.50, 1.43)**	0.27	1.15	**(0.15, 1.82)**	1.90	**(0.44, 1.53)**	0.30
	200	1.33	**(0.18, 1.77)**	19.75	**(0.49, 1.49)**	0.28	0.54	**(0.11, 2.04)**	2.54	**(0.49, 1.45)**	0.33
	240	1.17	**(0.13, 1.83)**	4.18	**(0.37, 1.39)**	0.28	1.02	**(0.01, 1.81)**	1.49	**(0.49, 1.42)**	0.26
GER/US	10	1.07	(−0.11, 1.79)	3.94	**(0.41, 1.47)**	0.30	0.89	**(0.17, 1.78)**	4.37	**(0.42, 1.50)**	0.31
	30	2.43	(−0.07, 1.89)	8.91	**(0.37, 1.50)**	0.32	0.79	**(0.22, 1.81)**	6.72	**(0.45, 1.53)**	0.30
	50	1.45	**(0.01, 1.92)**	2.67	**(0.39, 1.50)**	0.31	1.07	**(0.03, 1.83)**	5.36	**(0.49, 1.46)**	0.28
	70	2.80	(−0.02, 1.89)	3.85	**(0.42, 1.62)**	0.33	1.16	**(0.29, 1.71)**	2.98	**(0.45, 1.41)**	0.28
	90	1.06	**(0.02, 1.82)**	3.11	**(0.42, 1.40)**	0.29	1.48	**(0.13, 2.01)**	1.62	**(0.44, 1.43)**	0.29
	130	1.01	(−0.20, 2.00)	15.62	**(0.39, 1.45)**	0.29	0.71	**(0.04, 1.85)**	30.97	**(0.41, 1.49)**	0.31
	170	0.76	(−0.37, 1.67)	9.97	**(0.42, 1.44)**	0.30	0.96	**(0.15, 1.82)**	2.28	**(0.39, 1.45)**	0.31
	200	0.97	**(0.09, 1.86)**	49.36	**(0.49, 1.46)**	0.30	1.20	**(0.11, 2.31)**	12.26	**(0.45, 1.47)**	0.30
	240	0.68	**(0.05, 1.90)**	14.55	**(0.49, 1.47)**	0.32	0.96	**(0.01, 1.81)**	2.71	**(0.50, 1.46)**	0.29
ITA/US	10	1.22	(−0.89, 2.79)	4.23	**(0.25, 1.88)**	0.48	3.81	(−0.56, 2.67)	3.61	**(0.24, 1.89)**	0.49
	30	0.99	(−0.67, 2.66)	50.31	**(0.07, 1.89)**	0.54	1.79	(−0.88, 2.71)	5.60	**(0.08, 1.79)**	0.49
	50	1.50	(−0.62, 2.20)	6.15	**(0.08, 1.82)**	0.51	0.68	(−0.92, 2.26)	21.26	**(0.20, 1.92)**	0.50
	70	0.92	(−0.42, 2.83)	68.11	**(0.14, 1.91)**	0.52	0.60	(−0.51, 2.23)	11.41	**(0.12, 1.81)**	0.49
	90	1.04	(−0.58, 2.92)	21.44	**(0.19, 1.79)**	0.48	1.03	(−0.50, 2.42)	25.79	**(0.10, 1.80)**	0.49
	130	1.00	(−0.68, 2.32)	8.05	**(0.16, 1.88)**	0.51	1.17	(−0.89, 2.84)	23.69	**(0.03, 1.91)**	0.54
	170	1.18	(−0.85, 3.13)	8.25	**(0.13, 1.79)**	0.50	0.84	(−0.44, 2.53)	6.66	**(0.21, 1.79)**	0.48
	200	0.92	(−0.38, 2.29)	3.85	**(0.24, 1.89)**	0.47	0.72	(−0.54, 2.55)	5.03	**(0.19, 1.94)**	0.49
	240	0.86	(−0.35, 2.67)	3.70	**(0.18, 1.94)**	0.51	1.37	(−0.52, 2.46)	5.88	**(0.13, 1.82)**	0.49

Notes: FRN, GRE and ITA denote France, Germany and Italy, respectively. The column $\widehat{\xi}_{si}$ indicates the estimated proportion of real exchange-rate evolution explained by the nominal rates under the ith shock. The columns SE_i and CI_i provide the standard error of $\widehat{\xi}_{si}$ and the 90% percentile confidential of ξ_{si} for $i = 1,2$. The standard error and confidence interval are constructed based on 1000 block bootstrap replications with the block size being set to eight. The columns SE_i* and CI_i* provide the standard error of $\widehat{\xi}_{si}$ and the 90% percentile confidential of ξ_{si} after removing the extreme values of $\widehat{\xi}_{si}$. Bold values in a parenthesis indicate that the interval does not include zero and hence ξ_{si} is significantly different from zero.

removing extreme values of $\widehat{\xi}_{si}$) and its estimate is generally greater than 0.5 regardless of horizons and countries. The limitation of the VAR-VECM approach is that it assumes that the true DGP follows a VAR (k) model – which may not be true. The advantage of the local-projection approach is its robustness to specifications in the true DGP. However, our results from local projections are consistent with those found in Cheung et al. (2004), implying that their main findings are not sensitive to a possible misspecification in the true DGP.

Fig. 1 plots the IRF of nominal exchange rates, prices and real exchange rates, respectively, under different shocks. Several interesting results are observed. First, IRFs from local projections are not as smooth as conventional IRFs based on the VAR-VECM such as those found in Cheung et al. (2004). This is due to the fact that VAR-VECM-based IRFs across horizons are constructed based on the same VAR coefficients, whereas local-projections-based IRFs are obtained from direct forecasting across horizons and new estimates are involved for each horizon. Second, the IRF of real exchange rates is close to that of nominal exchange rates, indicating the significance of nominal exchange-rate adjustments in the reversion toward PPP. Third, this paper finds that price adjustment toward equilibrium is highly persistent relative to nominal exchange-rate adjustments for all three countries, implying that nominal exchange-rate adjustments play a significant role for PPP convergence. Overall, findings from Fig. 1 indicate that nominal exchange-rate adjustments play a major role in the reversion toward PPP for all countries, which is in line with the results in Table 3.

Table 4 reports the estimated half-life of prices, nominal and real exchange rates under different shocks as well as the 90% confidence interval of the half-life. The confidence interval is obtained from

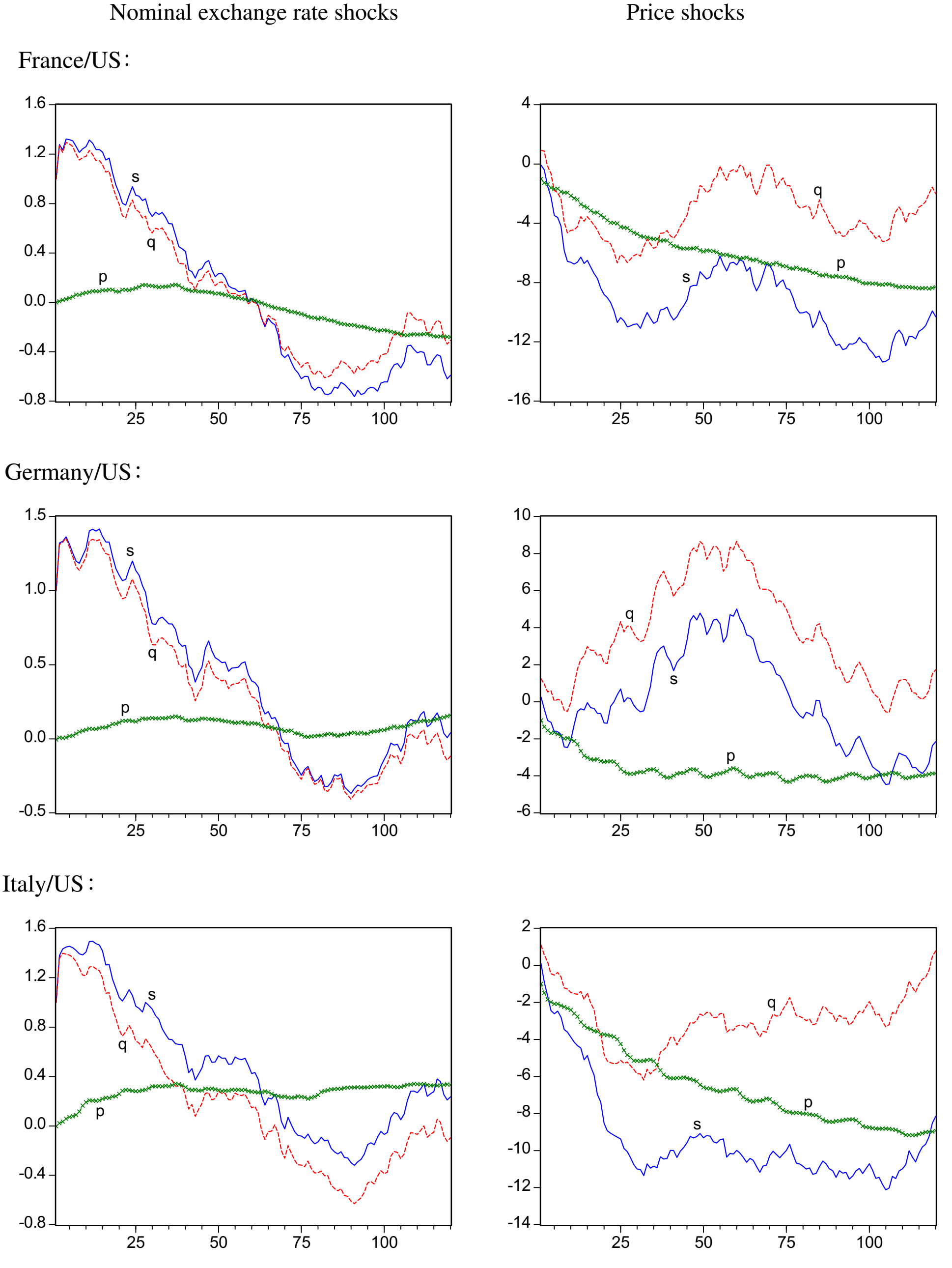

Fig. 1. IRFs of prices, nominal and real exchange rates with respect to shocks.

Table 4
Half-life estimation by local projections.

	FRN/US			GER/US			ITA/US		
	s_t	p_t	q_t	s_t	p_t	q_t	s_t	p_t	q_t
A nominal exchange-rate shock									
HL	38	95	28	36	98	33	31	104	28
0.05_bound	12	10	14	4	12	4	20	5	17
0.95_bound	108	119	110	104	118	113	118	116	118
A price shock									
HL	114	39	114	74	60	76	114	26	114
0.05_bound	30	2	39	20	4	3	5	4	28
0.95_bound	118	118	118	118	118	118	119	118	119

Note: FRN, GER and ITA denote France, Germany and Italy, respectively. HL denotes half-life. The long-run equilibrium for nominal exchange rates and prices is measured as the average value of the impulse response of these two variables at the 120th period under a shock. 0.05_bound and 0.95_bound are the lower and upper bound of a 90% confidence interval.

1000 block bootstrap replications with the block size being set to 8.[21] Following Cheung et al. (2004) and Engel and Morley (2001), this paper measures the equilibrium value for s_t and p_t as $(\mathrm{IR}_{si}|_{h=120} + \mathrm{IR}_{pi}|_{h=120})/2$, $i = 1, 2$.[22] Several interesting results are observed from Table 4. First, under a nominal exchange-rate shock, results from the upper panel of Table 4 indicate that the half-life of prices ranges from 8 to 9 years, but the half-life of nominal and real exchange rates ranges from 2.3 to 3.2 years. The half-life of real exchange rates in each country is close to that of nominal exchange rates. Although price adjustment is much more persistent than that of nominal exchange rates, nominal exchange-rate adjustments dominate in the reversion to PPP. This is the reason why the half-life of real exchange rates is close to that of nominal exchange rates. Second, under a price shock, the half-life of nominal and real exchange rates increases but that of prices decreases. The half-life varies from 6.2 to 9.5 years for real and nominal exchange rates and 2.2 to 5 years for prices. Again, the half-life of real rates is close to that of nominal rates. The much longer half-life of real and nominal exchange rates than that of prices under a price shock results from oscillatory IRFs with irregular cycles. Third, the confidence interval of the half-life is generally wide, regardless of shocks, which is consistent with Murray and Papell (2002). The reason could be that local-projection-based IRFs are constructed from direct forecasting across horizons and hence they are volatile relative to those constructed from VAR-VECM.

5. Conclusion

Ever since the seminal paper by Rogoff (1996), the role of price and nominal exchange-rate adjustments in the reversion toward PPP has been an interesting issue in relevant literature. Engel and Morley (2001) find that exchange-rate adjustment drives for PPP convergence at a torpid rate. Adopting a VAR-VECM approach, Cheung et al. (2004) find that “about 60–90% of real exchange-rate convergence is driven by nominal exchange rates and the contribution of prices to PPP convergence is relatively minor”. Their findings are in sharp contrast with the conventional explanation based on nominal rigidities and offer a new look at the theoretical modeling of real exchange-rate dynamics.

The VAR-VECM approach assumes that the true DGP follows a VAR(k) process which may be empirically restrictive. Jordà (2005) provides a local-projection approach to construct the IRF function of variables under investigation. Though it is robust to a misspecification in true DGP, Jordà’s (2005) method is not directly applicable in a non-stationary but cointegrated framework. This paper

[21] The half-life estimates are restricted to vary between 1 and 120 months since the equilibrium is assumed to be achieved at the 120th month. No extreme estimate is removed based on the criterion in footnote 19.

[22] Theoretically, there is no convergence for nominal exchange rates and relative prices in the impulse-response approach since they are non-stationary variables. In order to construct the half-life of each individual series, this paper follows Cheung et al. (2004) and Engel and Morley (2001) and presumes an “equilibrium value” for nominal exchange rates and relative prices at a sufficiently long time where PPP convergence has been completed, meaning that the paths of nominal exchange rates and relative prices are sufficiently close to each other.

therefore extends Jordà's method to the above framework and then constructs the IRF based on local projections. Therefore, our IRF analysis is robust to any misspecifications of the true DGP. This paper finds that nominal exchange-rate adjustments dominate in the reversion to PPP, regardless of nominal exchange-rate or price shocks and that the half-life of real exchange rates is close to that of nominal exchange rates. These results are consistent with those of Cheung et al. (2004). Therefore, the current research concludes that their main findings are robust to possible misspecifications in the true DGP.

Acknowledgements

We would like to appreciate Yi-Ting Chen, Chien-Ho Wang and especially an anonymous referee for their constructive suggestions and comments, which substantially improved the paper. Thanks are also due to the participants of seminar at National Sun Yat-sen University and Academia Sinica. Jyh-lin Wu appreciates financial support from National Science Council in Taiwan. All remaining errors are ours.

Appendix

In this appendix, this paper shows that (11) derived from a companion form (9) is indeed a vector-moving-average representation of $\Delta \mathbf{y}_t$. Johansen (1991) and Hansen (2005) have shown the MA representation of a cointegrated VAR system. This work therefore adopts Hansen's (2005) theorem 1 to facilitate our proof. For the convenience of our presentation, the current research re-produces Hansen's theorem 1 as follows.

Theorem 1. (*Granger Representation Theorem (Hansen, 2005)*) A cointegrated VAR process as equation (7) (where constant is the only deterministic term) has the following representation:

$$\mathbf{y}_t = \mathbf{C}\sum_{h=1}^{t} \epsilon_h + \mathbf{C}(L)\epsilon_t + \tau + \mathbf{C}(\mathbf{y}_0 - \mathbf{\Gamma}_1 \mathbf{y}_{-1} - \cdots - \mathbf{\Gamma}_{k-1}\mathbf{y}_{-k+1}), \tag{A1}$$

where $\mathbf{C} = \boldsymbol{\beta}_\perp(\boldsymbol{\alpha}'_\perp \mathbf{\Gamma}\boldsymbol{\beta}_\perp)^{-1}\boldsymbol{\alpha}'_\perp$ and the matrix $\boldsymbol{\alpha}'_\perp \mathbf{\Gamma}\boldsymbol{\beta}_\perp$ has full rank with $\mathbf{\Gamma} = \mathbf{I} - \sum_{i=1}^{k-1}\mathbf{\Gamma}_i$, $\tau = t\mathbf{C}\mathbf{\Phi} + \mathbf{C}(L)\mathbf{\Phi}$, and the coefficients in $\mathbf{C}(L)$ are given by the following recursive formula:

$$\Delta\mathbf{C}_h = \mathbf{\Pi}\mathbf{C}_{h-1} + \sum_{j=1}^{k-1}\mathbf{\Gamma}_j\Delta\mathbf{C}_{h-j}, \quad h = 1,2,\ldots \tag{A2}$$

with the convention of $\mathbf{C}_0 = \mathbf{I} - \mathbf{C}$, $\mathbf{C}_{-1} = \cdots = \mathbf{C}_{-k+1} = -\mathbf{C}$ and $\mathbf{\Pi}\mathbf{C} = \mathbf{0}$.

From Theorem 1, Hansen (2005) immediately has the following corollary results:

$$\Delta\mathbf{y}_t = \mathbf{C}_\Delta(L)\epsilon_t + \mathbf{C}\mathbf{\Phi}, \tag{A3}$$

where $\mathbf{C}_{\Delta,h} = \mathbf{C}_h - \mathbf{C}_{h-1}$, and $\mathbf{C}_h$, $h = 0,1,\ldots$, are coefficients of the polynomials in Theorem 1.[23] Furthermore, Hansen (2005) shows that (Lemma A.5, p.36) $\Delta\mathbf{C}_h$ can be calculated as

$$\Delta\mathbf{C}_h = \kappa_{2,h}, \quad h = 0,1,2,\ldots, \tag{A4}$$

where $\kappa_h = (\kappa'_{1,h}, \ldots, \kappa'_{k,h})' = \mathbf{F}^h\mathbf{E}$, in which $\mathbf{F}$ is the matrix defined in the companion form of equation (9) and $\mathbf{E} = (\mathbf{I}_n, \mathbf{I}_n, \mathbf{0}, \ldots, \mathbf{0})'$. Based on equation (A4), it is straightforward to derive that $\Delta\mathbf{C}_h = (\mathbf{F}^h_{2,1} + \mathbf{F}^h_{2,2})$. This shows that coefficients in the moving average representation derived from the companion form of equation (9) are identical with those derived from Hansen (2005).

To show that the coefficient matrices of $\mathbf{y}_{t-1}$ and $\Delta\mathbf{y}_{t-j}$ in equation (10) converge to zero, i.e. $\mathbf{F}^{h+1}_{2,j} \to \mathbf{0}\ \forall j = 1,\ldots,k$, this paper first shows that $\mathbf{F}^h$ does not diverge as $h \to \infty$. Following Phillips (1998), this work denotes

[23] As is claimed in footnote 6, a shock at time t changes $\mathbf{y}_t$ with $\mathbf{C} + \mathbf{C}_0$ from (A1), that is identical with the change of the same shock to $\Delta\mathbf{y}_t$ as in (A3), i.e. $\mathbf{C}_0 - \mathbf{C}_{-1} = \mathbf{C}_0 + \mathbf{C}$.

$$\mathbf{D} = \begin{bmatrix} \mathbf{\Pi}_1 & \mathbf{\Pi}_2 & \mathbf{\Pi}_3 & \cdots & \mathbf{\Pi}_{k-1} & \mathbf{\Pi}_k \\ \mathbf{I} & \mathbf{0} & \mathbf{0} & \cdots & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{I} & \mathbf{0} & \cdots & \mathbf{0} & \mathbf{0} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \cdots & \mathbf{I} & \mathbf{0} \end{bmatrix},$$

where $\mathbf{D}$ have eigenvalues equal to the inverse of the roots of $\det(\mathbf{A}(z)) = 0$.[24] The current research further denotes

$$\mathbf{K} = \begin{bmatrix} \mathbf{I} & \mathbf{0} & \mathbf{0} & \cdots & \mathbf{0} & \mathbf{0} \\ \mathbf{I} & -\mathbf{I} & \mathbf{0} & \cdots & \mathbf{0} & \mathbf{0} \\ \mathbf{I} & -\mathbf{I} & -\mathbf{I} & \cdots & \mathbf{0} & \mathbf{0} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ \mathbf{I} & -\mathbf{I} & -\mathbf{I} & \cdots & -\mathbf{I} & -\mathbf{I} \end{bmatrix}; \quad \mathbf{K}^{-1} = \begin{bmatrix} \mathbf{I} & \mathbf{0} & \mathbf{0} & \cdots & \mathbf{0} & \mathbf{0} \\ \mathbf{I} & -\mathbf{I} & \mathbf{0} & \cdots & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{I} & -\mathbf{I} & \cdots & \mathbf{0} & \mathbf{0} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \cdots & \mathbf{I} & -\mathbf{I} \end{bmatrix}.$$

This paper then shows that matrix $\mathbf{F}$ in equation (9) and matrix $\mathbf{D}$ are similar matrices since $\mathbf{F} = \mathbf{K}^{-1}\mathbf{D}\mathbf{K}$. Therefore, matrix $\mathbf{F}$ has the same eigenvalues as matrix $\mathbf{D}$. Since this work has assumed that eigenvalues of matrix $\mathbf{D}$ are either one or smaller than one. Thus, $\mathbf{F}^h$ will not diverge as $h \to \infty$.

Based on our assumption on the roots of $\det(\mathbf{A}(z)) = 0$ and the fact that $\mathbf{\Pi} = \boldsymbol{\alpha}\boldsymbol{\beta}'$, this work concludes that the roots of $\det[\mathbf{I}(1-z) - \boldsymbol{\alpha}\boldsymbol{\beta}'z - \sum_{i=1}^{k-1}\mathbf{\Gamma}_i(1-z)z^i] = 0$ are outside the unit circle implying that $\Delta\mathbf{C}_h$ in (A2) is stable, and hence $\Delta\mathbf{C}_h \to \mathbf{0}$ as $h \to \infty$. It is then straightforward to show that $(\mathbf{F}_{2,1}^{h+1} + \mathbf{F}_{2,2}^{h+1}) \to \mathbf{0}$ as $h \to \infty$ since $\Delta\mathbf{C}_h = (\mathbf{F}_{2,1}^h + \mathbf{F}_{2,2}^h)$ as indicated by equation (A4). The structure of $\mathbf{F}$ yields the following two equations:

$$\mathbf{F}_{2,i}^{h+1} = (\mathbf{F}_{2,1}^h + \mathbf{F}_{2,2}^h)\mathbf{\Gamma}_{i-1} + \mathbf{F}_{2,i+1}^h, \quad \forall i = 2, \ldots, k-1; \tag{A5}$$

$$\mathbf{F}_{2,k}^{h+1} = (\mathbf{F}_{2,1}^h + \mathbf{F}_{2,2}^h)\mathbf{\Gamma}_{k-1}. \tag{A6}$$

Solving equations (A5) and (A6) with recursive substitution yields:

$$\mathbf{F}_{2,j}^{h+1} = (\mathbf{F}_{2,1}^h + \mathbf{F}_{2,2}^h)\mathbf{\Gamma}_{j-1} + (\mathbf{F}_{2,1}^{h-1} + \mathbf{F}_{2,2}^{h-1})\mathbf{\Gamma}_j + \cdots + (\mathbf{F}_{2,1}^{h-k+j} + \mathbf{F}_{2,2}^{h-k+j})\mathbf{\Gamma}_{k-1}, \forall j = 2, \ldots, k, \tag{A7}$$

where $\mathbf{F}_{2,1}^{h-k+j} + \mathbf{F}_{2,2}^{h-k+j} = \mathbf{0}$ if $h-k+j < 0$. Equation (A7) indicates that $\mathbf{F}_{2,j}^{h+1}$ for $j = 2,\ldots,k$, can be expressed as a function of ${\mathbf{F}_{2,1}}^{h-p} + {\mathbf{F}_{2,2}}^{h-p}$ for a finite $p = 0, \ldots, P$. Since this work has shown that $(\mathbf{F}_{2,1}^{h+1} + \mathbf{F}_{2,2}^{h+1}) \to \mathbf{0}$ as $h \to \infty$, which implies that $\mathbf{F}_{2,j}^{h+1} \to \mathbf{0}$ as $h \to \infty$, for $j = 2,\ldots,k$. This would also imply that $\mathbf{F}_{2,1}^{h+1} \to \mathbf{0}$. This paper concludes that for all $j = 1,\ldots,k$, $\mathbf{F}_{2,j}^{h+1} \to \mathbf{0}$ as $h \to \infty$ in equation (10). Therefore, equation (11) is indeed a MA representation.

References

Abuaf, N., Jorion, P., 1990. Purchasing power parity in the long run. Journal of Finance 45, 157–174.
Brockwell, P.J., Davis, R.A., 1991. Time Series: Theory and Methods, second ed. Springer-Verlag, New York.
Chang, P.L., Sakata, S., 2007. Estimation of impulse response functions using long autoregression. The Econometrics Journal 10, 453–469.
Cheung, Y.W., Lai, K.S., 1993. Finite-sample sizes of Johansen's likelihood ratio tests for cointegration. Oxford Bulletin of Economics and Statistics 55, 313–328.
Cheung, Y.W., Lai, K.S., Bergman, M., 2004. Dissecting the PPP puzzle: the unconventional roles of nominal exchange rate and price adjustments. Journal of International Economics 64, 135–150.
Elliott, G., Rothenberg, T.J., Stock, J.H., 1996. Efficient tests for an autoregressive unit root. Econometrica 64, 813–836.
Engle, R.F., Granger, C.W.J., 1987. Cointegration and error correction representation: estimation and testing. Econometrica 55, 251–276.
Engel, C., Morley, J.C., 2001. The adjustment of prices and the adjustment of the exchange rate. NBER Working Paper no.8550.
Engle, R.F., Yoo, S., 1987. Forecasting and testing in co-integrated systems. Journal of Econometrics 35, 143–159.

[24] See for example, Schott (1997, p. 250, Theorem 7.3).

Froot, K.A., Rogoff, K., 1995. Perspectives on PPP and long-run real exchange rates. In: Grossman, G., Rogoff, K. (Eds.), Handbook of International Economics, vol. 3. North-Holland, Amsterdam, pp. 1647–1688.

Fuller, W.A., 1996. An Introduction to Statistical Time Series, second ed. John Wiley and Sons.

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

Hansen, P.R., 2005. Granger's representation theorem: a closed-form expression for I(1) processes. Econometrics Journal 8, 23–38.

Ing, C.K., 2003. Multistep prediction in autoregressive processes. Econometric Theory 19, 254–279.

Johansen, S., 1991. Estimation and hypothesis testing of cointegration vectors in Gaussian vector autoregressive models. Econometrica 59, 1551–1580.

Johansen, S., 2002. A small sample correction for the test of cointegrating rank in the vector autoregressive model. Econometrica 70, 1929–1961.

Jordà, Ò, 2005. Estimation and inference of impulse responses by local projections. American Economic Review 95, 161–182.

Jordà, Ò, 2009. Simultaneous confidence regions for impulse responses. The Review of Economics and Statistics 91, 629–674.

Kazimi, C., Brownstone, D., 1999. Bootstrap confidence bands for shrinkage estimators. Journal of Econometrics 90, 99–127.

Kilian, L., Kim, Y.J., 2009. Do local projections solve the bias problem in impulse inference. Working Paper, Department of Economics, University of Michigan.

Kim, T.-H., White, H., 2004. On more robust estimation of skewness and kurtosis. Finance Research Letters 1, 56–73.

Koop, G.M., Pesaran, H., Potter, S., 1996. Impulse response analysis in nonlinear multivariate models. Journal of Econometrics 74, 119–147.

Künsch, H.R., 1989. The Jackknife and the bootstrap for general stationary observations. The Annals of Statistics 17, 1217–1241.

Lin, J.L., Tsay, R.S., 1996. Co-integration constraint and forecasting: an empirical examination. Journal of Applied Econometrics 11, 519–538.

Liu, R., Singh, K., 1992. Moving blocks Jackknife and bootstrap capture weak dependence. In: LePage, R., Billiard, L. (Eds.), Exploring the Limits of the Bootstrap. Wiley, New York.

Mark, N., 1990. Real and nominal exchange rates in the long run: an empirical investigation. Journal of International Economics 28, 115–136.

Michael, P., Nobay, A.R., Peel, D.A., 1997. Transaction costs and nonlinear adjustment in real exchange rates: an empirical investigation. Journal of Political Economy 105, 862–879.

Murray, C.J., Papell, D.H., 2002. The purchasing power parity persistence paradigm. Journal of International Economics 56 (1), 1–19.

O'Connell, P.G.J., 1998. Market frictions and real exchange rates. Journal of International Money and Finance 17, 71–95.

Pesaran, M.H., Shin, Y., 1998. Generalized impulse response analysis in linear multivariate models. Economics Letters 58, 17–29.

Phillips, P.C.B., 1998. Impulse response and forecast error variance asymptotics in nonstationary VARs. Journal of Econometrics 83, 21–56.

Romano, J.P., Thombs, L.A., 1996. Inference for autocorrelations under weak assumptions. Journal of the American Statistical Association 91, 590–600.

Reinsel, G.C., Ahn, S.K., 1988. Asymptotic properties of the likelihood ratio test for cointegration in the nonstationary vector AR model. Technical Report, Department of Statistics, University of Wisconsin-Madison.

Rogoff, K., 1996. The purchasing power parity puzzle. Journal of Economic Literature 34, 647–668.

Schott, J.R., 1997. Matrix Analysis for Statistics. Wiley, New York.

Taylor, M.P., Sarno, L., 1998. The behavior of real exchange rates during the post-Bretton Woods period. Journal of International Economics 46, 281–312.

Weiss, A.A., 1991. Multi-step estimation and forecasting in dynamic models. Journal of Econometrics 48, 135–149.