

CAMA

Centre for Applied Macroeconomic Analysis

Purchasing Power Parity and the Taylor Rule

CAMA Working Paper 41/2013 June 2013

Hyeongwoo Kim

Department of Economics, Auburn University

Ippei Fujiwara

Crawford School of Public Policy, ANU and
Centre for Applied Macroeconomic Analysis (CAMA), ANU

Bruce E. Hansen

Department of Economics, University of Wisconsin

Masao Ogaki

Department of Economics, Keio University

Abstract

It is well-known that there is a large degree of uncertainty around Rogoff's (1996) consensus half-life of the real exchange rate. To obtain a more efficient estimator, we develop a system method that combines the Taylor rule and a standard exchange rate model to estimate half-lives. Further, we propose a median unbiased estimator for the system method based on the generalized method of moments with nonparametric grid bootstrap confidence intervals. Applying the method to real exchange rates of 18 developed countries against the US dollar, we find that most half-life estimates from the single equation method fall in the range of 3 to 5 years with wide confidence intervals that extend to positive infinity. In contrast, the system method yields median-unbiased estimates that are typically shorter than one year with much sharper 95% confidence intervals. Our Monte Carlo simulation results are consistent with an interpretation of these results that the true half-lives are short but long half-life estimates from single equation methods are caused by the high degree of uncertainty of these methods.

Keywords

Purchasing Power Parity; Taylor Rule; Half-Life of PPP Deviations; Median Unbiased Estimator; Grid-t Confidence Interval

JEL Classification

C32; E52; F31

Address for correspondence:

(E) cama.admin@anu.edu.au

[The Centre for Applied Macroeconomic Analysis](#) in the Crawford School of Public Policy has been established to build strong links between professional macroeconomists. It provides a forum for quality macroeconomic research and discussion of policy issues between academia, government and the private sector.

The Crawford School of Public Policy is the Australian National University's public policy school, serving and influencing Australia, Asia and the Pacific through advanced policy research, graduate and executive education, and policy impact.

Purchasing Power Parity and the Taylor Rule*

Hyeonwoo Kim[†], Ippei Fujiwara[‡], Bruce E. Hansen[§], and Masao Ogaki[¶]

June 2013

Abstract

It is well-known that there is a large degree of uncertainty around Rogoff's (1996) consensus half-life of the real exchange rate. To obtain a more efficient estimator, we develop a system method that combines the Taylor rule and a standard exchange rate model to estimate half-lives. Further, we propose a median unbiased estimator for the system method based on the generalized method of moments with nonparametric grid bootstrap confidence intervals. Applying the method to real exchange rates of 18 developed countries against the US dollar, we find that most half-life estimates from the single equation method fall in the range of 3 to 5 years with wide confidence intervals that extend to positive infinity. In contrast, the system method yields median-unbiased estimates that are typically shorter than one year with much sharper 95% confidence intervals. Our Monte Carlo simulation results are consistent with an interpretation of these results that the true half-lives are short but long half-life estimates from single equation methods are caused by the high degree of uncertainty of these methods.

Keywords: Purchasing Power Parity; Taylor Rule; Half-Life of PPP Deviations; Median Unbiased Estimator; Grid- t Confidence Interval

JEL Classification: C32; E52; F31

*Special thanks go to Lutz Kilian, Christian Murray, David Papell, Paul Evans, Eric Fisher, Henry Thompson, and two anonymous referees for helpful suggestions. We also thank seminar participants at Bank of Japan, Keio University, University of Houston, University of Michigan, Texas Tech University, Auburn University, University of Southern Mississippi, the 2009 ASSA Meeting, the 72nd Midwest Economics Association Annual Meeting, the Midwest Macro Meeting 2008, and the 2009 NBER Summer Institute (EFSF).

[†]Department of Economics, Auburn University, 0339 Haley Center., Auburn, AL 36849, USA. Tel: 1-334-844-2928. Fax: 1-334-844-4615. Email: gmmkim@gmail.com.

[‡]Crawford School of Public Policy and Centre for Applied Macroeconomic Analysis, Australian National University, Canberra ACT 0200, Australia. Tel: 61-2-6125-4763. Email: ippei.fujiwara@anu.edu.au.

[§]Department of Economics, University of Wisconsin, 1180 Observatory Drive, Madison, WI 53706-1393, USA. Tel: 1-608-263-3880. Fax: 1-608-263-3876. Email: behansen@wisc.edu.

[¶]Department of Economics, Keio University, 612 Mita-Kenkyu-shitsu 2-15-45 Mita Minato-ku, Tokyo, 108-8345, Japan. Tel: 81-3-5418-6403. Fax: 81-3-5427-1578. Email: mogaki@econ.keio.ac.jp.

1 Introduction

Reviewing the literature on Purchasing Power Parity (PPP) that uses single equation methods to estimate the half-lives of real exchange rate deviations from PPP, Rogoff (1996) found a remarkable consensus on 3 to 5 year half-life estimates. This formed an important piece of Rogoff's "PPP puzzle" as the question of how one might reconcile highly volatile short-run movements of real exchange rates with an extremely slow convergence rate to PPP.

Using Rogoff's consensus half-life as a starting point, various possible solutions to the PPP puzzle have been proposed in the literature.¹ One important discussion in this context relates to the aggregation bias that may generate upward bias in half-life estimates.² Another delicate issue is how one can aggregate micro evidence of price stickiness for dynamic aggregate models, such as in dynamic stochastic general equilibrium (DSGE) models, which Carvalho and Nechio (2011) have begun to investigate. Even though aggregation bias is an important potential problem, much more research seems necessary before a consensus is reached on whether or not the aggregation bias solves the PPP puzzle, and how we should aggregate for DSGE models.

In this paper, we ask a different question: Should we take Rogoff's remarkable consensus of 3-5 year half-life estimates as the starting point for aggregate CPI data? The consensus may at first seem to support the reliability of these estimates, but Kilian and Zha (2002), Murray and Papell (2002), and Rossi (2005) have all shown that there is a high degree of uncertainty around these point estimates. Murray and Papell (2002) conclude that single equation methods provide virtually no information regarding the size of the half-lives, indicating that it is not clear if the true half-lives are in fact as slow as Rogoff's remarkable consensus implies. If we apply a more efficient estimator to the real exchange rate data, it may be possible to find faster convergence rates.

For the purpose of obtaining a more efficient estimator, we develop a system method that combines the Taylor rule and a standard exchange rate model to estimate the half-life of the real exchange rate. Several recent papers have provided empirical evidence in favor of exchange rate models using Taylor rules (see Mark 2009, Engel and West 2005, 2006, Clarida and Waldman 2007, Molodtsova and Papell 2009, and Molodtsova, Nikolsko-Rzhevskyy and Papell 2008). Therefore, a system method using an exchange rate model with the Taylor rule is a promising way to improve on single equation methods to estimate the half-lives.

Because standard asymptotic theory usually does not provide adequate approximations for the estimation of half-lives of real exchange rates, we use a nonparametric bootstrap method to construct confidence intervals. For this purpose, we propose the grid bootstrap method for our GMM

¹See Murray and Papell (2002) for a discussion of these other solutions which take Rogoff's consensus half-life as a starting point.

²Imbs, Mumtaz, Ravn, and Rey (2005) point out that sectoral heterogeneity in convergence rates can cause upward bias in half-life estimates, and claim that this aggregation bias solves the PPP puzzle. While under certain conditions this is possible, the bias can be negligible under other conditions. For example, Chen and Engel (2005), Crucini and Shintani (2008), and Parsley and Wei (2007) have found negligible aggregation biases. Broda and Weinstein (2008) show that the aggregation bias of the form that Imbs, Mumtaz, Ravn, and Rey (2005) studied is small for their barcode data, even though the convergence coefficient rises as they move to aggregate indexes. These papers focus on purely statistical findings.

estimator along with its asymptotic distribution. Median unbiased estimates and bias corrected confidence bands are reported.³

We apply the system method to estimate the half-lives of real exchange rates of 18 developed countries against the US dollar. Most of the estimates from the single equation method fall in the range of 3 to 5 years, with wide confidence intervals that extend to positive infinity. In contrast, the system method yields median unbiased estimates that are typically substantially shorter than 3 years with much sharper confidence intervals, predominantly ranging from three quarters to 5 years. We implement an array of Monte Carlo simulations in order to understand why one might obtain much longer half-lives from single equation estimators than that of our system method. Our findings imply that the high estimates of the persistence parameter by single equation estimators in the literature may well be caused by large standard errors of the single equation estimators.

In the recent papers that use two-country exchange rate models with Taylor rules cited above, the authors assume that Taylor rules are adopted by the central banks of both countries. As some countries may not use Taylor rules, we remain agnostic about the monetary policy rule in the foreign country and assume that the Taylor rule is employed only by the home country. None of these papers with Taylor rules estimates the half-lives of real exchange rates.

Kim and Ogaki (2004), Kim (2005), and Kim, Ogaki, and Yang (2007) use system methods to estimate the half-lives of real exchange rates. However, they use conventional monetary models based on money demand functions without Taylor rules. Another important point of difference of these works from the present paper is that their inferences are based on asymptotic theory, while ours are based on the grid bootstrap.

The rest of the paper is organized as follows. Section 2 describes our baseline model. We construct a system of stochastic difference equations for the exchange rate and inflation, explicitly incorporating a forward looking Taylor rule into the system. Section 3 explains our estimation methods. In Section 4, we report our empirical results. Section 5 provides explanations on our Monte Carlo simulation schemes and findings. Section 6 presents our conclusions.

³Kehoe and Midrigan (2007) and Crucini, Shintani, and Tsuruga (2013) show that the persistence of the real exchange rate can be understood in the context of the New Keynesian Phillips Curve (NKPC) framework with Calvo (1983) pricing. That is, a higher degree of price inertia may cause more persistent real exchange rate deviations. Interestingly, the contrast between the single equation methods and our system method in the context of the PPP literature is similar to the contrast between single equation methods for the NKPC and system methods for DSGE models with the NKPC observed in the literature for closed economy models. Single equation methods such as Galí and Gertler's (1999) GMM yield small standard errors for the average price duration based on standard asymptotic theory. However, Kleibergen and Mavroeidis (2009), who take into account the weak identification problem of GMM, report that the upper bound of their 95% confidence interval for the price duration is infinity. The estimators of average price duration in system methods for DSGE models in Christiano, Eichenbaum, and Evans (2005) and Smets and Wouters (2007), among others, may be more efficient.

2 The Model

2.1 Gradual Adjustment Equation

We start with a univariate stochastic process of real exchange rates. Let p_t be the log domestic price level, p_t^* be the log foreign price level, and e_t be the log nominal exchange rate as the price of one unit of the foreign currency in terms of the home currency. And we denote s_t as the log of the real exchange rate, $p_t^* + e_t - p_t$.

We assume that PPP holds in the long-run. In other words, we assume that a cointegrating vector $[1 \ -1 \ -1]'$ exists for a vector $[p_t \ p_t^* \ e_t]'$, where p_t , p_t^* , and e_t are difference stationary processes. Under this assumption, the real exchange rate can be represented as the following stationary univariate autoregressive process of degree one.

$$s_{t+1} = d + \alpha s_t + \varepsilon_{t+1}, \quad (1)$$

where α is a positive persistence parameter that is less than one.

Admittedly, estimating the half-lives of real exchange rates with an AR(1) specification may not be ideal, because the AR(1) model is mis-specified and will lead to an inconsistent estimator if the true data generating process is a higher order autoregressive process, AR(p). It is interesting to see, however, that Rossi (2005) reported similar half-life estimates from both models. Later in Section 4, we confirm that this is roughly the case when we apply the single equation method to our exchange rate data. Thus, assuming AR(1) seems innocuous for the purpose of estimating the half life of most real exchange rates in our data. However, it is still possible that more general AR(p) models yield quite different half-lives for some exchange rates, particularly when the system method is used because we often observe hump-shaped responses (Steinsson, 2008). Even though this is an interesting question, we do not pursue this issue in the current paper because it is not easy to obtain informative saddle-path solutions for a higher order system of difference equations.

By rearranging and taking conditional expectations, the equation (1) can be written by the following error correction model of real exchange rates with the cointegrating relation described earlier.

$$\mathbb{E}_t \Delta p_{t+1} = b [\mu - (p_t - p_t^* - e_t)] + \mathbb{E}_t \Delta p_{t+1}^* + \mathbb{E}_t \Delta e_{t+1}, \quad (2)$$

where $\mu = \mathbb{E}(p_t - p_t^* - e_t)$, $b = 1 - \alpha$, $d = -(1 - \alpha)\mu$, $\varepsilon_{t+1} = \varepsilon_{1,t+1} + \varepsilon_{2,t+1} - \varepsilon_{3,t+1} = (e_{t+1} - \mathbb{E}_t e_{t+1}) + (p_{t+1}^* - \mathbb{E}_t p_{t+1}^*) - (p_{t+1} - \mathbb{E}_t p_{t+1})$, and $\mathbb{E}_t \varepsilon_{t+1} = 0$. $E(\cdot)$ denotes the unconditional expectation operator while $\mathbb{E}_t(\cdot)$ is the conditional expectation operator on I_t , the economic agent's information set at time t . Note that this model is consistent with a single-good version of Mussa's (1982) model.⁴ Note that b is the convergence rate ($= 1 - \alpha$), which is a positive constant less than unity by construction.

⁴We added a domestic price shock, $p_{t+1} - \mathbb{E}_t p_{t+1}$, that has a conditional expectation of zero given the information at time t .

2.2 The Taylor Rule Model

We assume that the uncovered interest parity (UIP) holds. That is,

$$\mathbb{E}_t \Delta e_{t+1} = i_t - i_t^*, \quad (3)$$

where i_t and i_t^* are domestic and foreign interest rates, respectively.⁵

The central bank in the home country is assumed to continuously set its optimal target interest rate (i_t^T) by the following forward looking Taylor Rule.⁶

$$i_t^T = \bar{r} + \gamma_\pi \mathbb{E}_t \Delta p_{t+1} + \gamma_x x_t,$$

where \bar{r} is a constant that includes a certain long-run equilibrium real interest rate along with a target inflation rate⁷, and γ_π and γ_x are the *long-run* Taylor Rule coefficients on expected future inflation⁸ ($\mathbb{E}_t \Delta p_{t+1}$) and current output deviations⁹ (x_t), respectively. We also assume that the central bank attempts to smooth the interest rate by the following rule.

$$i_t = (1 - \rho) i_t^T + \rho i_{t-1},$$

that is, the current actual interest rate is a weighted average of the target interest rate and the previous period's interest rate, where ρ is the smoothing parameter. Then, we can derive the forward looking version Taylor Rule equation with interest rate smoothing policy as follows.

$$i_t = (1 - \rho) \bar{r} + (1 - \rho) \gamma_\pi \mathbb{E}_t \Delta p_{t+1} + (1 - \rho) \gamma_x x_t + \rho i_{t-1} \quad (4)$$

Combining (3) and (4), we obtain the following.

$$\begin{aligned} \mathbb{E}_t \Delta e_{t+1} &= (1 - \rho) \bar{r} + (1 - \rho) \gamma_\pi \mathbb{E}_t \Delta p_{t+1} + (1 - \rho) \gamma_x x_t + \rho i_{t-1} - i_t^* \\ &= \iota + \gamma_\pi^s \mathbb{E}_t \Delta p_{t+1} + \gamma_x^s x_t + \rho i_{t-1} - i_t^*, \end{aligned} \quad (5)$$

where $\iota = (1 - \rho) \bar{r}$ is a constant, $\gamma_\pi^s = (1 - \rho) \gamma_\pi$ and $\gamma_x^s = (1 - \rho) \gamma_x$ are *short-run* Taylor Rule coefficients.

⁵The UIP often fails to hold when one tests it by estimating a single regression equation, $\Delta e_{t+1} = \beta(i_t - i_t^*) + \varepsilon_{t+1}$. This indicates that it is not ideal to assume the UIP in our model, and future research should remove this assumption. We believe, however, that our initial attempt should start with the UIP, because it is difficult to write an exchange rate model with the Taylor rule without the UIP for our purpose of getting more information from the model. Further, Taylor rule-based exchange rate models in the literature often assumes the UIP.

⁶We remain agnostic about the policy rule of the foreign central bank, because the Taylor rule may not be employed in some countries.

⁷See Clarida, Galí, and Gertler (1998, 2000) for details.

⁸It may be more reasonable to use real-time data instead of final release data. However, doing so will introduce another complication as we need to specify the relation between the real-time price index and the consumer price index, which is frequently used in the PPP literature. Hence we leave the use of real-time data for future research.

⁹If we assume that the central bank responds to expected future output deviations rather than current deviations, we can simply modify the model by replacing x_t with $\mathbb{E}_t x_{t+1}$. However, this does not make any significant difference to our results.

Now, let's rewrite (2) as the following equation in level variables.

$$\mathbb{E}_t p_{t+1} = b\mu + \mathbb{E}_t e_{t+1} + (1-b)p_t - (1-b)e_t + \mathbb{E}_t p_{t+1}^* - (1-b)p_t^* \quad (2')$$

Taking differences and rearranging it, (2') can be rewritten as follows.

$$\mathbb{E}_t \Delta p_{t+1} = \mathbb{E}_t \Delta e_{t+1} + \alpha \Delta p_t - \alpha \Delta e_t + [\mathbb{E}_t \Delta p_{t+1}^* - \alpha \Delta p_t^* + \eta_t], \quad (6)$$

where $\alpha = 1 - b$ and $\eta_t = \eta_{1,t} + \eta_{2,t} - \eta_{3,t} = (e_t - \mathbb{E}_{t-1} e_t) + (p_t^* - \mathbb{E}_{t-1} p_t^*) - (p_t - \mathbb{E}_{t-1} p_t)$.

From (4), (5), and (6), we construct the following system of stochastic difference equations.

$$\begin{bmatrix} 1 & -1 & 0 \\ -\gamma_\pi^s & 1 & 0 \\ -\gamma_\pi^s & 0 & 1 \end{bmatrix} \begin{bmatrix} \mathbb{E}_t \Delta p_{t+1} \\ \mathbb{E}_t \Delta e_{t+1} \\ i_t \end{bmatrix} = \begin{bmatrix} \alpha & -\alpha & 0 \\ 0 & 0 & \rho \\ 0 & 0 & \rho \end{bmatrix} \begin{bmatrix} \Delta p_t \\ \Delta e_t \\ i_{t-1} \end{bmatrix} + \begin{bmatrix} \mathbb{E}_t \Delta p_{t+1}^* - \alpha \Delta p_t^* + \eta_t \\ \iota + \gamma_x^s x_t - i_t^* \\ \iota + \gamma_x^s x_t \end{bmatrix} \quad (7)$$

For notational simplicity, let's rewrite (7) in matrix form as follows.

$$\mathbf{A} \mathbb{E}_t \mathbf{y}_{t+1} = \mathbf{B} \mathbf{y}_t + \mathbf{x}_t, \quad (7')$$

and thus,

$$\begin{aligned} \mathbb{E}_t \mathbf{y}_{t+1} &= \mathbf{A}^{-1} \mathbf{B} \mathbf{y}_t + \mathbf{A}^{-1} \mathbf{x}_t \\ &= \mathbf{D} \mathbf{y}_t + \mathbf{c}_t, \end{aligned} \quad (8)$$

where $\mathbf{D} = \mathbf{A}^{-1} \mathbf{B}$ and $\mathbf{c}_t = \mathbf{A}^{-1} \mathbf{x}_t$.¹⁰ By eigenvalue decomposition, (8) can be rewritten as follows.

$$\mathbb{E}_t \mathbf{y}_{t+1} = \mathbf{V} \mathbf{\Lambda} \mathbf{V}^{-1} \mathbf{y}_t + \mathbf{c}_t, \quad (9)$$

where $\mathbf{D} = \mathbf{V} \mathbf{\Lambda} \mathbf{V}^{-1}$ and

$$\mathbf{V} = \begin{bmatrix} 1 & 1 & 1 \\ \frac{\alpha \gamma_\pi^s}{\alpha - \rho} & 1 & 1 \\ \frac{\alpha \gamma_\pi^s}{\alpha - \rho} & 1 & 0 \end{bmatrix}, \quad \mathbf{\Lambda} = \begin{bmatrix} \alpha & 0 & 0 \\ 0 & \frac{\rho}{1 - \gamma_\pi^s} & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

Pre-multiplying (9) by \mathbf{V}^{-1} and redefining variables,

$$\mathbb{E}_t \mathbf{z}_{t+1} = \mathbf{\Lambda} \mathbf{z}_t + \mathbf{h}_t, \quad (10)$$

where $\mathbf{z}_t = \mathbf{V}^{-1} \mathbf{y}_t$ and $\mathbf{h}_t = \mathbf{V}^{-1} \mathbf{c}_t$.

Note that, among non-zero eigenvalues in $\mathbf{\Lambda}$, α is between 0 and 1 by definition, while $\frac{\rho}{1 - \gamma_\pi^s}$ ($= \frac{\rho}{1 - (1 - \rho)\gamma_\pi}$) is greater than unity as long as $1 < \gamma_\pi < \frac{1}{1 - \rho}$. Therefore, if the long-run inflation

¹⁰It is straightforward to show that \mathbf{A} is nonsingular, and thus has a well-defined inverse.

coefficient γ_π is strictly greater than one, the system of stochastic difference equations (7) has a saddle path equilibrium, where rationally expected future fundamental variables enter in the exchange rate and inflation dynamics.¹¹ On the contrary, if γ_π is strictly less than unity, which might be true in the pre-Volker era in the US, the system would have a purely backward looking solution, where the solution would be determined by *past* fundamental variables and *any* martingale difference sequences.

Assuming γ_π is strictly greater than one, we can show that the solution to (7) satisfies the following relation (see Appendix A for the derivation).

$$\begin{aligned} \Delta e_{t+1} = & \hat{i} + \frac{\alpha\gamma_\pi^s}{\alpha - \rho} \Delta p_{t+1} - \frac{\alpha\gamma_\pi^s}{\alpha - \rho} \Delta p_{t+1}^* + \frac{\alpha\gamma_\pi^s - (\alpha - \rho)}{\alpha - \rho} i_t^* \\ & + \frac{\gamma_\pi^s(\alpha\gamma_\pi^s - (\alpha - \rho))}{(\alpha - \rho)\rho} \sum_{j=0}^{\infty} \left(\frac{1 - \gamma_\pi^s}{\rho} \right)^j \mathbb{E}_t f_{t+j+1} + \omega_{t+1}, \end{aligned} \quad (11)$$

where,

$$\begin{aligned} \hat{i} &= \frac{\alpha\gamma_\pi^s - (\alpha - \rho)}{(\alpha - \rho)(\gamma_\pi^s - (1 - \rho))} i_t^*, \\ f_t &= - [i_t^* - \mathbb{E}_t \Delta p_{t+1}^*] + \frac{\gamma_x^s}{\gamma_\pi^s} x_t \end{aligned}$$

$$\begin{aligned} \omega_{t+1} &= \frac{\gamma_\pi^s(\alpha\gamma_\pi^s - (\alpha - \rho))}{(\alpha - \rho)\rho} \sum_{j=0}^{\infty} \left(\frac{1 - \gamma_\pi^s}{\rho} \right)^j (\mathbb{E}_{t+1} f_{t+j+1} - \mathbb{E}_t f_{t+j+1}) \\ &+ \frac{\gamma_\pi^s}{\alpha - \rho} \eta_{t+1} - \frac{\alpha\gamma_\pi^s - (\alpha - \rho)}{\alpha - \rho} v_{t+1}, \end{aligned}$$

and,

$$\mathbb{E}_t \omega_{t+1} = 0$$

Or, (11) can be rewritten with full parameter specification as follows.

$$\begin{aligned} \Delta e_{t+1} = & \hat{i} + \frac{\alpha\gamma_\pi(1 - \rho)}{\alpha - \rho} \Delta p_{t+1} - \frac{\alpha\gamma_\pi(1 - \rho)}{\alpha - \rho} \Delta p_{t+1}^* + \frac{\alpha\gamma_\pi(1 - \rho) - (\alpha - \rho)}{\alpha - \rho} i_t^* \\ & + \frac{\gamma_\pi(1 - \rho)(\alpha\gamma_\pi(1 - \rho) - (\alpha - \rho))}{(\alpha - \rho)\rho} \sum_{j=0}^{\infty} \left(\frac{1 - \gamma_\pi(1 - \rho)}{\rho} \right)^j \mathbb{E}_t f_{t+j+1} + \omega_{t+1} \end{aligned} \quad (11')$$

Here, f_t is a proxy variable that summarizes the fundamental variables such as foreign *ex ante* real interest rates and domestic output deviations.

Note that if γ_π is strictly less than unity, the restriction in (11) may not be valid, since the system would have a backward looking equilibrium rather than a saddle path equilibrium.¹² In other words, exchange rate dynamics critically depends on the size of γ_π . However, as mentioned in

¹¹The condition $\gamma_\pi < \frac{1}{1-\rho}$ is easily met for all sample periods we consider in this paper.

¹²If the system has a purely backward looking solution, the conventional structural Vector Autoregressive (SVAR) estimation method may apply.

the introduction, we have some supporting empirical evidence of this requirement for the existence of a saddle path equilibrium, at least for the post-Volker era. So we believe that our specification remains valid for our purpose in this paper.

One related study, recently put forward by Clarida and Waldman (2007), investigates exchange rate dynamics when central banks employ Taylor rules in a small open economy framework (Svensson, 1999). In their paper, they derive the dynamics of real exchange rates by combining the Taylor Rule and the uncovered interest parity (or real interest parity), so that the real exchange rate is mainly determined by the *ex ante* real interest rate. In their model, the real interest rate follows an AR(1) process of which the autoregressive coefficient is a function of the Taylor rule coefficients. When the central bank responds to inflation more aggressively, the economy returns to its long-run equilibrium at a faster rate. Therefore, the half-life of PPP deviations is negatively affected by γ_π .

It should be noted that their model does not explicitly incorporate the commodity view of PPP in the sense that real exchange rate dynamics are mainly determined by the portfolio market equilibrium conditions. In contrast to their model, we combine a single good version of Mussa's (1982) model (2) with the UIP as well as the Taylor Rule. Under this framework, no policy parameters can affect the half-life of the PPP deviations because real exchange rate persistence is mainly driven by commodity arbitrages. On the other hand, policy parameters do affect volatilities of inflation and the nominal exchange rate in our model. For example, the more aggressively the central bank responds to inflation, the less volatile inflation is, which leads to a less volatile nominal exchange rate.

One interesting feature arises when another policy parameter, ρ , varies. As the value for ρ increases, the volatility of Δp_{t+1} decreases. This is due to the uncovered interest parity condition. A higher value of ρ , higher interest rate inertia, implies that the central bank changes the nominal interest rate less. Therefore, Δe_{t+1} should change less due to the uncovered interest parity. When $\alpha = \rho$, it can be shown that after the initial cost-push shock, price does not change at all (see Appendix B). That is, Δp_{t+1} instantly jumps and stays at its long-run equilibrium value of zero. Hence, the convergence toward long-run PPP should be carried over by the exchange rate adjustments. When $\alpha < \rho$, price must decrease after the initial cost-push shock, since the nominal exchange rate movement is limited by the uncovered interest parity and domestic interest rate inertia.

3 Estimation Methods

We discuss two estimation strategies here: a conventional univariate equation approach and the GMM system method (Kim, Ogaki, and Yang, 2007).

3.1 Univariate Equation Approach

A univariate approach utilizes the equations (1) or (2). For instance, the persistence parameter α in (1) can be consistently estimated by the conventional least squares method under the maintained

cointegrating relation assumption. Once we obtain the point estimate of α , the half-life of the real exchange rate can be calculated by $\frac{\ln(.5)}{\ln \alpha}$. Similarly, the regression equation for the convergence parameter b can be constructed from (2) as follows.

$$\Delta p_{t+1} = b[\mu - (p_t - p_t^* - e_t)] + \Delta p_{t+1}^* + \Delta e_{t+1} + \tilde{\varepsilon}_{t+1}, \quad (2'')$$

where $\tilde{\varepsilon}_{t+1} = -\varepsilon_{t+1} = -(e_{t+1} - \mathbb{E}_t e_{t+1}) - (p_{t+1}^* - \mathbb{E}_t p_{t+1}^*) + (p_{t+1} - \mathbb{E}_t p_{t+1})$ and $\mathbb{E}_t \tilde{\varepsilon}_{t+1} = 0$.

3.2 GMM System Method

Our second estimation strategy combines the equation (11) with (1). The estimation of the equation (11) is a challenging task, however, since it has an infinite sum of rationally expected discounted future fundamental variables. Following Hansen and Sargent (1980, 1982), we linearly project $\mathbb{E}_t(\cdot)$ onto Ω_t , the econometrician's information set at time t , which is a subset of I_t . Denoting $\hat{\mathbb{E}}_t(\cdot)$ as such a linear projection operator onto Ω_t , we can rewrite (11) as follows.

$$\begin{aligned} \Delta e_{t+1} = & \hat{\omega} + \frac{\alpha \gamma_\pi^s}{\alpha - \rho} \Delta p_{t+1} - \frac{\alpha \gamma_\pi^s}{\alpha - \rho} \Delta p_{t+1}^* + \frac{\alpha \gamma_\pi^s - (\alpha - \rho)}{\alpha - \rho} i_t^* \\ & + \frac{\gamma_\pi^s (\alpha \gamma_\pi^s - (\alpha - \rho))}{(\alpha - \rho) \rho} \sum_{j=0}^{\infty} \left(\frac{1 - \gamma_\pi^s}{\rho} \right)^j \hat{\mathbb{E}}_t f_{t+j+1} + \xi_{t+1}, \end{aligned} \quad (12)$$

where

$$\xi_{t+1} = \omega_{t+1} + \frac{\gamma_\pi^s (\alpha \gamma_\pi^s - (\alpha - \rho))}{(\alpha - \rho) \rho} \sum_{j=0}^{\infty} \left(\frac{1 - \gamma_\pi^s}{\rho} \right)^j \left(\mathbb{E}_t f_{t+j+1} - \hat{\mathbb{E}}_t f_{t+j+1} \right),$$

and

$$\hat{\mathbb{E}}_t \xi_{t+1} = 0,$$

by the law of iterated projections.

For appropriate instrumental variables that are in Ω_t , we assume $\Omega_t = \{f_t, f_{t-1}, f_{t-2}, \dots\}$. This assumption would be an innocent one under the stationarity assumption of the fundamental variable, f_t , and it can greatly lessen the burden in our GMM estimation by significantly reducing the number of coefficients to be estimated.

Assume, for now, that f_t is a zero mean covariance stationary, linearly indeterministic stochastic process so that it has the following Wold representation.

$$f_t = c(L)\nu_t, \quad (13)$$

where $\nu_t = f_t - \hat{\mathbb{E}}_{t-1} f_t$ and $c(L)$ is square summable. Assuming that $c(L) = 1 + c_1 L + c_2 L^2 + \dots$ is invertible, (13) can be rewritten as the following autoregressive representation.

$$b(L)f_t = \nu_t, \quad (14)$$

where $b(L) = c^{-1}(L) = 1 - b_1L - b_2L^2 - \dots$. Linearly projecting $\sum_{j=0}^{\infty} \left(\frac{1-\gamma_\pi^s}{\rho}\right)^j \mathbb{E}_t f_{t+j+1}$ onto Ω_t , Hansen and Sargent (1980) show that the following relation holds.

$$\sum_{j=0}^{\infty} \delta^j \hat{\mathbb{E}}_t f_{t+j+1} = \psi(L) f_t = \left[\frac{1 - (\delta^{-1} b(\delta))^{-1} b(L) L^{-1}}{1 - (\delta^{-1} L)^{-1}} \right] f_t, \quad (15)$$

where $\delta = \frac{1-\gamma_\pi^s}{\rho}$.

For actual estimation, we assume that f_t can be represented by a finite order AR(r) process, that is, $b(L) = 1 - \sum_{j=1}^r b_j L^j$, where $r < \infty$.¹³ It can then be shown that the coefficients of $\psi(L)$ can be computed recursively (see Sargent 1987) as follows.

$$\psi_0 = (1 - \delta b_1 - \dots - \delta^r b_r)^{-1}$$

$$\psi_r = 0$$

$$\psi_{j-1} = \delta \psi_j + \delta \psi_0 b_j,$$

where $j = 1, 2, \dots, r$. We then obtain the following two orthogonality conditions.

$$\begin{aligned} \Delta e_{t+1} &= \hat{i} + \frac{\alpha \gamma_\pi^s}{\alpha - \rho} \Delta p_{t+1} - \frac{\alpha \gamma_\pi^s}{\alpha - \rho} \Delta p_{t+1}^* + \frac{\alpha \gamma_\pi^s - (\alpha - \rho)}{\alpha - \rho} i_t^* \\ &+ \frac{\gamma_\pi^s (\alpha \gamma_\pi^s - (\alpha - \rho))}{(\alpha - \rho) \rho} (\psi_0 f_t + \psi_1 f_{t-1} + \dots + \psi_{r-1} f_{t-r+1}) + \xi_{t+1}, \end{aligned} \quad (16)$$

$$f_{t+1} = k + b_1 f_t + b_2 f_{t-1} + \dots + b_r f_{t-r+1} + \nu_{t+1}, \quad (17)$$

where k is a constant scalar and $\hat{\mathbb{E}}_t \nu_{t+1} = 0$.^{14,15}

Finally, the system method (GMM) estimation utilizes all aforementioned orthogonality conditions, (2''), (16), and (17). That is, a GMM estimation can be implemented by the following $2(p+2)$ orthogonality conditions.

$$\mathbb{E} \mathbf{x}_{1,t} (s_{t+1} - d - \alpha s_t) = 0 \quad (18)$$

$$\hat{\mathbb{E}} \mathbf{x}_{2,t-\tau} \left(\begin{aligned} &\Delta e_{t+1} - \hat{i} - \frac{\alpha \gamma_\pi^s}{\alpha - \rho} \Delta p_{t+1} + \frac{\alpha \gamma_\pi^s}{\alpha - \rho} \Delta p_{t+1}^* - \frac{\alpha \gamma_\pi^s - (\alpha - \rho)}{\alpha - \rho} i_t^* \\ &- \frac{\gamma_\pi^s (\alpha \gamma_\pi^s - (\alpha - \rho))}{(\alpha - \rho) \rho} (\psi_0 f_t + \psi_1 f_{t-1} + \dots + \psi_{r-1} f_{t-r+1}) \end{aligned} \right) = 0 \quad (19)$$

$$\hat{\mathbb{E}} \mathbf{x}_{2,t-\tau} (f_{t+1} - k - b_1 f_t - b_2 f_{t-1} - \dots - b_r f_{t-r+1}) = 0, \quad (20)$$

¹³We can use conventional Akaike Information criteria or Bayesian Information criteria in order to choose the degree of such autoregressive processes.

¹⁴Recall that Hansen and Sargent (1980) assume a zero-mean covariance stationary process. If the variable of interest has a non-zero unconditional mean, we can either demean it prior to the estimation or include a constant but leave its coefficient unconstrained. West (1989) showed that the further efficiency gain can be obtained by imposing additional restrictions on the deterministic term. However, the imposition of such an additional restriction is quite burdensome, so we simply add a constant here.

¹⁵In actual estimations, we normalized (16) by multiplying $(\alpha - \rho)$ to each side in order to reduce nonlinearity.

where $\mathbf{x}_{1,t} = (1 \ s_t)'$, $\mathbf{x}_{2,t} = (1 \ f_t)'$, and $\tau = 0, 1, \dots, p$.^{16,17}

3.3 Median Unbiased Estimator and Grid- t Confidence Intervals

We correct for the bias in our α estimates by the grid- t method, which is similar to the one by Hansen (1999) for the least squares estimator. It is straightforward to generate pseudo samples for the orthogonality condition (20) by the conventional residual-based bootstrapping. However, there are some complications in obtaining samples directly from (18) and (19), since p_t^* is treated as a forcing variable in our model. We deal with this problem as follows.

In order to generate pseudo samples for the orthogonality conditions (18) and (19), we denote \tilde{p}_t as the relative price index $p_t - p_t^*$. Then, (2'') and (16) can be rewritten as follows.

$$\begin{aligned} \Delta \tilde{p}_{t+1} &= b\mu - b(\tilde{p}_t - e_t) + \Delta e_{t+1} + \tilde{\varepsilon}_{t+1} \\ \Delta e_{t+1} &= \hat{i} + \frac{\alpha \gamma_\pi^s}{\alpha - \rho} \Delta \tilde{p}_{t+1} + \frac{\alpha \gamma_\pi^s - (\alpha - \rho)}{\alpha - \rho} i_t^* \\ &\quad + \frac{\gamma_\pi^s (\alpha \gamma_\pi^s - (\alpha - \rho))}{(\alpha - \rho) \rho} (\psi_0 f_t + \dots + \psi_{r-1} f_{t-r+1}) + \xi_{t+1} \end{aligned}$$

Or, in matrix form,

$$\begin{aligned} \begin{bmatrix} \Delta \tilde{p}_{t+1} \\ \Delta e_{t+1} \end{bmatrix} &= \mathbf{C} + \mathbf{S}^{-1} \begin{bmatrix} -(1 - \alpha) \\ 0 \end{bmatrix} [\tilde{p}_t - e_t] \\ &\quad + \mathbf{S}^{-1} \begin{bmatrix} 0 \\ \frac{\alpha \gamma_\pi^s - (\alpha - \rho)}{\alpha - \rho} i_t^* + \frac{\gamma_\pi^s (\alpha \gamma_\pi^s - (\alpha - \rho))}{(\alpha - \rho) \rho} \\ (\psi_0 f_t + \dots + \psi_{r-1} f_{t-r+1}) \end{bmatrix} + \mathbf{S}^{-1} \begin{bmatrix} \tilde{\varepsilon}_{t+1} \\ \xi_{t+1} \end{bmatrix}, \end{aligned} \quad (21)$$

where \mathbf{C} is a vector of constants and \mathbf{S} is $\begin{bmatrix} 1 & -1 & \dots & -\frac{\alpha \gamma_\pi^s}{\alpha - \rho} & 1 \end{bmatrix}$.

Then, by treating each grid point $\alpha \in [\alpha_{\min}, \alpha_{\max}]$ as a true value, we can generate pseudo samples of $\Delta \tilde{p}_{t+1}$ and Δe_{t+1} through the conventional bootstrapping.¹⁸ The level variables \tilde{p}_t and e_t are obtained by numerical integration. It should be noted that all other parameters are treated as nuisance parameters (η).¹⁹ Following Hansen (1999), we define the grid- t statistic at each grid point $\alpha \in [\alpha_{\min}, \alpha_{\max}]$ as follows.

$$t_n(\alpha) = \frac{\hat{\alpha}_{\text{GMM}} - \alpha}{se(\hat{\alpha}_{\text{GMM}})}, \quad (22)$$

where $se(\hat{\alpha}_{\text{GMM}})$ denotes the robust GMM standard error at the GMM estimate $\hat{\alpha}_{\text{GMM}}$. Implementing GMM estimations for B bootstrap iterations at each of N grid point of α , we obtain the

¹⁶ p does not necessarily coincide with r .

¹⁷In actual estimations, we again use the aforementioned normalization.

¹⁸Historical data was used for the initial values and the foreign interest rate i_t^* .

¹⁹See Hansen (1999) for detailed explanations.

(β quantile) grid- t bootstrap quantile functions, $q_{n,\beta}^*(\alpha) = q_{n,\beta}^*(\alpha, \eta(\alpha))$. Note that each function is evaluated at each grid point α rather than at the point estimate.²⁰

In Appendix C, we derive the asymptotic distribution of the grid- t statistic (22) as follows. Under the local to unity ($\alpha = 1 + c/n$) framework,

$$t_n(\alpha) \Rightarrow \frac{\mathbf{S}'_1 (\mathbf{G}'_c \boldsymbol{\Omega}_c^{-1} \mathbf{G}_c)^{-1} \mathbf{G}'_c \boldsymbol{\Omega}_c^{-1} \mathbf{N}_c}{\left(\mathbf{S}'_1 (\mathbf{G}'_c \boldsymbol{\Omega}_c^{-1} \mathbf{G}_c)^{-1} \mathbf{S}_1 \right)^{1/2}}$$

where \mathbf{N}_c , $\boldsymbol{\Omega}_c$, and \mathbf{G}_c are defined in (c4), (c5), and (c6).

Finally, we define the 95% grid- t confidence interval as follows.

$$\{\alpha \in R : q_{n,2.5\%}^*(\alpha) \leq t_n(\alpha) \leq q_{n,97.5\%}^*(\alpha)\}, \quad (23)$$

and the median unbiased estimator as,

$$\alpha_{MUE} = \alpha \in R, \text{ s.t. } t_n(\alpha) = q_{n,50\%}^*(\alpha) \quad (24)$$

In Appendix C, we also show that the grid bootstrap confidence bands are correctly sized under some regularity conditions described in Assumption 1.

4 Empirical Results

This section reports estimates of the persistence parameter α (or convergence rate parameter b) and their implied half-lives resulting from the two estimation strategies discussed above.

We use CPIs to construct real exchange rates with the US\$ as a base currency. We consider 19 industrialized countries that provide 18 real exchange rates.²¹ For interest rates, we use quarterly money market interest rates that are short-term interbank call rates rather than conventional short-term treasury bill rates, since we incorporate the Taylor Rule in the model where a central bank sets its target short-term market rate. For output deviations, we consider two different measures of output gaps, quadratically detrended real GDP gap (see Clarida, Galí, and Gertler 1998) and unemployment rate gaps (see Boivin, 2006).^{22,23} The data frequency is quarterly and from the IFS CD-ROM. The sample period is from 1979:III to 1998:IV for Eurozone countries, and from 1979:III to 2003:IV for the rest of the countries.

Based on the empirical evidence of the US Taylor Rule, our sample period starts from 1979:III. As discussed in Section II, the inflation and exchange rate dynamics may greatly depend on the

²⁰If they are evaluated at the point estimate, the quantile functions correspond to the Efron and Tibshirani's (1993) bootstrap- t quantile functions.

²¹Among the 23 industrialized countries classified by IMF, we dropped Greece, Iceland, and Ireland due to lack of reasonable number of observations. Luxembourg was also dropped because of its currency union with Belgium.

²²We also tried the same analysis with the cyclical components of real GDP series from the HP-filter with 1600 of smoothing parameter. The results were quantitatively similar.

²³The unemployment gap is defined as a 5 year backward moving average subtracted by the current unemployment rate. This specification makes its sign consistent with that of the conventional output gap.

size of the central bank’s reaction coefficient to expected inflation. We showed that the rationally expected *future* fundamental variables appear in the exchange rate and inflation dynamics only when the long-run inflation coefficient γ_π is strictly greater than unity. Clarida, Galí, and Gertler (1998, 2000) provide important empirical evidence for the existence of a structural break in the US Taylor Rule. Put it differently, they show that γ_π was strictly less than one during the pre-Volker era, while it became strictly greater than unity in the post-Volker era.

We implement similar GMM estimations for (4) as in Clarida, Galí, and Gertler (2000)^{24,25} with a longer sample period and report the results in Table 1 (see the note on Table 1 for detailed explanation). We use two output gap measures for three different sub-samples. Most coefficients were highly significant and specification tests by J -test were not rejected.²⁶ More importantly, our requirement for the existence of a saddle path equilibrium was met in the post-Volker era rather than the pre-Volker era. Therefore, we may conclude that this provides an empirical justification for the choice of our sample period.

Insert Table 1 Here

We report our GMM version median unbiased estimates and the 95% grid- t confidence intervals in Table 2. We implemented estimations using both gap measures, but report the full estimates with unemployment gaps in order to save space.²⁷ We chose $N = 30$ and $B = 500$ totaling 15,000 GMM simulations for each exchange rate. We chose $p = r = 8$ by the conventional Bayesian Information Criteria, and standard errors were adjusted using the QS kernel estimator with automatic bandwidth selection in order to deal with unknown serial correlation problems. For comparison, we report the corresponding estimates by the least squares in Table 3.

We note that the system method provides much shorter half-life estimates compared with those resulting from the single equation method (see Tables 2 and 3). The median value of the half-life estimate was 3.42 years from the univariate estimations after adjusting for the median bias using the grid- t bootstrap. However, the median value of the GMM median unbiased estimates was still below 1 year, 0.94 year, when we correct for the bias.²⁸ Our estimates are roughly consistent with the average half-life estimates from the micro-data evidence by Crucini and Shintani (2008)²⁹ and the differences of the point estimates for different countries are very similar to those of Murray and

²⁴They used GDP deflator inflation along with the CBO output gaps (and HP detrended gaps).

²⁵Unlike Clarida, Galí, and Gertler (2000), we assume that the Fed targets current output gap rather than future deviations. However, this doesn’t make any significant changes to our results. And we include one lag of interest rate rather than two lags for simplicity.

²⁶ J -test statistics are available upon request.

²⁷The results with quadratically detrended real GDP gaps were quantitatively similar.

²⁸Without bias correction, the median value of the half-life estimate was 2.59 years from the univariate estimations and 0.90 year from the system method. All estimates and the conventional 95% bootstrap confidence intervals are available from authors upon request.

²⁹For the OECD countries, their baseline half-life estimates for traded good prices were 1.5 years, and 1.58 and 2.00 years for all and non-traded good prices.

Papell (2002) for most countries.³⁰ J -test accepts our model specification for all countries with an exception of the UK.³¹

Regarding efficiency, we obtained substantial efficiency gains from the system method over the single equation method. Murray and Papell (2002) report a version of the grid- α confidence intervals (Hansen, 1999)³² of which upper limits of their half-life estimates are infinity for every exchange rates they consider. Based on such results, they conclude that single equation methods may provide virtually no useful information due to wide confidence intervals.

Our grid- t confidence intervals from the single equation method were consistent with such a view (see Table 3). The upper limits are infinity for most real exchange rates. However, when we implement estimations by the system method, our 95% GMM version grid- t confidence intervals were very compact. Our results can be also considered as great improvement over Kim, Ogaki, and Yang (2007) who acquired limited success in efficiency gains.

Insert Table 2 Here

Insert Table 3 Here

Lastly, we compare univariate half-life estimates from an AR(1) specification with those from a more general AR(p) specification. Following Rossi (2005), we choose the number of lags by the modified Akaike Information criteria (MAIC, Ng and Perron, 2001) with a maximum 12 lags. We also estimate the lag length by the modified Bayesian Information criteria (MBIC, Ng and Perron, 2001), which yields $p = 1$ for most real exchange rates. The MAIC chooses $p = 1$ for 6 out of 18 real exchange rates. For the remaining 12 real exchange rates, we implement the impulse-response analysis to estimate the half-lives of PPP deviations. As can be seen in Table 4, allowing higher order AR(p) processes results in very different half-life estimates from those of the AR(1) specification for some countries such as Italy, Portugal, and Spain. This implies that one has to be careful in interpreting the results based on AR(1) models for these exchange rates. For many other real exchange rates, however, half-life estimates do not change much, implying that the AR(1) process is not a bad approximation.

Insert Table 4 Here

³⁰The exceptions to this similarity are Japan and the UK, as our point estimates for these countries are much smaller than others. Using the same sample period of Murray and Papell (2002), however, we obtained the α estimates of 0.89 and 0.82 for Japan and the UK, respectively, indicating that these exceptions seem to have arisen from the difference in the sample periods.

³¹We also notice that our median-unbiased point estimate $\hat{\alpha}_{\text{GMM,MUE}}$ is consistent with the price-stickiness parameter estimates by Galí and Gertler (1999) who use the New Keynesian Phillips Curve specification with Calvo pricing. Recall that a single-good version model by Kehoe and Midrigan (2007) or Kim (2009) implies that α coincides with the Calvo probability parameter.

³²Their confidence intervals are constructed following Andrews (1993) and Andrews and Chen (1994), which are identical to the Hansen's (1999) grid- α confidence intervals if we assume that the errors are drawn from the empirical distribution rather than the i.i.d. normal distribution.

5 Monte Carlo Simulation Studies

The empirical results in the previous section are consistent with two possible interpretations. One is that the true half-lives are short, and long half-life estimates given by single equation methods are due to their high degree of uncertainty. Another is that the true half-lives are long, and short half-life estimates obtained by the system method are due to the bias caused by the misspecification of the model. For the purpose of obtaining evidence as to which interpretation is more appropriate, this section provides Monte Carlo simulations based on the DSGE model described in Appendix D, which is consistent with the model equations above that are used for our estimation.

For the purpose of examining the impact of misspecification, we introduce the UIP shock in addition to the monetary policy shock. We investigate three possible values for the size of the variance of the UIP shock (σ_v^2) relative to that of the monetary shock (σ_u^2), that is, $\sigma_v^2 = 0$, $\sigma_v^2 = \sigma_u^2$, and $\sigma_v^2 = 5\sigma_u^2$. Recall that our saddle-path equation was derived in the absence of the UIP shock. Putting it differently, the greater the value for σ_v^2 is, the more severe the misspecification of the system method is. We also consider 78 observations (T) for each simulated series that match those of the Eurozone countries, while $T = 500$ is also employed in order to see what happens in large samples. We further consider errors from the standard normal distribution as well as errors from the student- t distribution with three degrees of freedom (t_3). Variances, 1 and 3 for the standard normal and t_3 are re-scaled so that they match with calibrated variances.

From 500 sets of simulated observations for each case, we estimated kernel density estimates of the persistence parameter via the single equation method and the system method. All estimates are corrected for the median-bias before we estimate baseline statistics and density function estimates.³³

We report estimated distributions of the persistence parameter in Figure 1 through 3. We also report various statistical properties of single and system estimators in Table 5.

We note that the system method is substantially more efficient than the single equation method when the number of observations is small ($T = 78$) as we can see in Figures 1 and 2. Even though the single equation estimator has somewhat better empirical properties in mean and median values (see Table 5), the distributions of the single equation estimators are flatter than those of the system method estimators. Therefore, high estimates of the persistence parameter by the single equation method in the literature may well be caused by high standard errors. We also note that these results are fairly robust to the size of the UIP shock and to the underlying distributional assumption of the shocks.

When misspecification of the system method is very large and the sample size is much larger than that of the available data, then the cost of misspecification can offset the benefit of efficiency of the system method. For instance, when $T = 500$ and $\sigma_v^2 = \sigma_u^2$, the difference of standard deviations becomes quite small so that the gain of using the system method decreases. However, with reasonable size of misspecification and realistic sample size, it is likely that the cost of misspecification is much smaller than the benefit of efficiency.

³³We use interpolations using the estimates from up to 10 grid points to correct for bias in GMM estimates.

Insert Figures 1, 2, and 3 Here

Insert Table 5 here

6 Conclusion

It is a well-known fact that there is a high degree of uncertainty around Rogoff's (1996) consensus half-life of the real exchange rate. In response to this fact, this paper proposed a system method that combines the Taylor Rule and a standard exchange rate model. We estimated the half-lives of real exchange rates for 18 developed countries against the US and obtained much shorter half-life estimates than those obtained using the single equation method. Our Monte Carlo simulation results are consistent with an interpretation that the large uncertainty of the single equation estimators is responsible for the high estimates of the persistence parameter from single equation methods in the literature.

We used two types of nonparametric bootstrap methods to construct confidence intervals: the standard bootstrap and the grid bootstrap for our GMM estimator, where we also demonstrate the asymptotic properties of the grid bootstrap method. The standard bootstrap evaluates bootstrap quantiles at the point estimate of the AR(1) coefficient, which implicitly assumes that the bootstrap quantile functions are constant functions. This assumption does not hold for the AR model, and the grid bootstrap method, which avoids this assumption, has better coverage properties. In our applications, we often obtain very different confidence intervals for these two methods.³⁴ Therefore, the violation of the assumption is deemed quantitatively important.

When we use the grid bootstrap method, most of the (approximately) median unbiased estimates from the single equation method fall in the range of 3 to 5 years with wide confidence intervals that extend to positive infinity. In contrast, the system method yields median unbiased estimates that are typically substantially less than one year with much sharper confidence intervals, most of which range from 3 quarters to 5 years.

These results indicate that monetary variables from the exchange rate model based on the Taylor rule provide useful information about the half-lives of real exchange rates. Confidence intervals are much narrower than those from a single equation method indicating that the estimators from the system method are significantly sharper. Approximately median unbiased estimates of the half-lives are typically about one year, which is far more reasonable than the consensus 3 to 5 years from single equation methods.³⁵

Our paper is the first step toward a system method with the exchange rate model based on the Taylor rule. We followed most of the papers in the literature with this type of model by

³⁴Results from standard bootstrap are available upon request.

³⁵It is also interesting to see that our half-life estimates imply about 4 to 6 quarters of average price duration in the context of the Calvo pricing model. Our 95% confidence intervals of the half-lives of real exchange rates are consistent with most of the estimates of average price durations for aggregate U.S. data for the NKPC and DSGE models.

using the uncovered interest parity to connect the Taylor rule to the exchange rate. Because the uncovered interest parity for short-term interest rates is rejected by the data, one future direction is to modify the model by removing the uncovered interest parity. This is a challenging task because no consensus has emerged as to how the deviation from the uncovered interest parity should be modelled. Even though the AR(1) specification seems to be a good approximation for most real exchange rates, it is possible that more general AR(p) models yield quite different half-lives for some exchange rates. This is another challenging task in our system approach, as it is not easy to obtain informative saddle-path solutions for a higher order system of difference equations.

A Derivation of (11)

Since \mathbf{A} in (10) is diagonal, assuming $0 < \alpha < 1$ and $1 < \gamma_\pi < \frac{1}{1-\rho}$, we can solve the system as follows.

$$z_{1,t} = \sum_{j=0}^{\infty} \alpha^j h_{1,t-j-1} + \sum_{j=0}^{\infty} \alpha^j u_{t-j} \quad (\text{a1})$$

$$z_{2,t} = - \sum_{j=0}^{\infty} \left(\frac{1 - \gamma_\pi^s}{\rho} \right)^{j+1} \mathbb{E}_t h_{2,t+j} \quad (\text{a2})$$

$$z_{3,t} = h_{3,t-1} + v_t, \quad (\text{a3})$$

where u_t and v_t are *any* martingale difference sequences.

Since $\mathbf{y}_t = \mathbf{V}\mathbf{z}_t$,

$$\begin{bmatrix} \Delta p_t \\ \Delta e_t \\ i_{t-1} \end{bmatrix} = \begin{bmatrix} 1 & 1 & 1 \\ \frac{\alpha\gamma_\pi^s}{\alpha-\rho} & 1 & 1 \\ \frac{\alpha\gamma_\pi^s}{\alpha-\rho} & 1 & 0 \end{bmatrix} \begin{bmatrix} z_{1,t} \\ z_{2,t} \\ z_{3,t} \end{bmatrix} \quad (\text{a4})$$

From first and second rows of (a4), we get the following.

$$\Delta e_t = \frac{\alpha\gamma_\pi^s}{\alpha-\rho} \Delta p_t - \frac{\alpha\gamma_\pi^s - (\alpha-\rho)}{\alpha-\rho} z_{2,t} - \frac{\alpha\gamma_\pi^s - (\alpha-\rho)}{\alpha-\rho} z_{3,t} \quad (\text{a5})$$

Now, we find the analytic solutions for \mathbf{z}_t . Since $\mathbf{h}_t = \mathbf{V}^{-1}\mathbf{c}_t$,

$$\mathbf{h}_t = \frac{1}{1-\gamma_\pi^s} \begin{bmatrix} -\frac{\alpha-\rho}{\alpha\gamma_\pi^s - (\alpha-\rho)} & \frac{\alpha-\rho}{\alpha\gamma_\pi^s - (\alpha-\rho)} & 0 \\ \frac{\alpha\gamma_\pi^s}{\alpha\gamma_\pi^s - (\alpha-\rho)} & -\frac{\alpha\gamma_\pi^s}{\alpha\gamma_\pi^s - (\alpha-\rho)} & 1 \\ 0 & 1 & -1 \end{bmatrix} \begin{bmatrix} \mathbb{E}_t \Delta p_{t+1}^* - \alpha \Delta p_t^* + \eta_t + \iota + \gamma_x^s x_t - i_t^* \\ \gamma_\pi^s (\mathbb{E}_t \Delta p_{t+1}^* - \alpha \Delta p_t^* + \eta_t) + \iota + \gamma_x^s x_t - i_t^* \\ \gamma_\pi^s (\mathbb{E}_t \Delta p_{t+1}^* - \alpha \Delta p_t^* + \eta_t) + \iota + \gamma_x^s x_t - \gamma_\pi^s i_t^* \end{bmatrix},$$

and thus,

$$h_{1,t} = -\frac{\alpha-\rho}{\alpha\gamma_\pi^s - (\alpha-\rho)} (\mathbb{E}_t \Delta p_{t+1}^* - \alpha \Delta p_t^* + \eta_t) \quad (\text{a6})$$

$$h_{2,t} = \frac{1}{1-\gamma_\pi^s} \left[\frac{\rho\gamma_\pi^s}{\alpha\gamma_\pi^s - (\alpha-\rho)} (\mathbb{E}_t \Delta p_{t+1}^* - \alpha \Delta p_t^* + \eta_t) + \iota + \gamma_x^s x_t - \gamma_\pi^s i_t^* \right] \quad (\text{a7})$$

$$h_{3,t} = -i_t^* \quad (\text{a8})$$

Plugging (a6) into (a1),

$$\begin{aligned} z_{1,t} &= -\frac{\alpha-\rho}{\alpha\gamma_\pi^s - (\alpha-\rho)} \sum_{j=0}^{\infty} \alpha^j (\Delta p_{t-j}^* - \alpha \Delta p_{t-j-1}^* + \eta_{t-j-1}) + \sum_{j=0}^{\infty} \alpha^j u_{t-j} \\ &= -\frac{\alpha-\rho}{\alpha\gamma_\pi^s - (\alpha-\rho)} \Delta p_t^* + \sum_{j=0}^{\infty} \alpha^j u_{t-j} - \frac{\alpha-\rho}{\alpha\gamma_\pi^s - (\alpha-\rho)} \sum_{j=0}^{\infty} \alpha^j \eta_{t-j-1} \end{aligned} \quad (\text{a9})$$

Plugging (a7) into (a2)³⁶,

$$\begin{aligned}
z_{2,t} &= -\frac{\gamma_\pi^s}{\alpha\gamma_\pi^s - (\alpha - \rho)} \sum_{j=0}^{\infty} \left(\frac{1 - \gamma_\pi^s}{\rho}\right)^j (\mathbb{E}_t \Delta p_{t+j+1}^* - \alpha \mathbb{E}_t \Delta p_{t+j}^* + \mathbb{E}_t \eta_{t+j}) \\
&\quad - \frac{1}{\rho} \sum_{j=0}^{\infty} \left(\frac{1 - \gamma_\pi^s}{\rho}\right)^j (\iota + \gamma_x^s \mathbb{E}_t x_{t+j} - \gamma_\pi^s \mathbb{E}_t i_{t+j}^*) \\
&= \frac{\alpha\gamma_\pi^s}{\alpha\gamma_\pi^s - (\alpha - \rho)} \Delta p_t^* - \frac{\gamma_\pi^s}{\alpha\gamma_\pi^s - (\alpha - \rho)} \eta_t - \frac{\iota}{\gamma_\pi^s - (1 - \rho)} \\
&\quad - \frac{\gamma_\pi^s}{\rho} \sum_{j=0}^{\infty} \left(\frac{1 - \gamma_\pi^s}{\rho}\right)^j \mathbb{E}_t \Delta p_{t+j+1}^* - \frac{\gamma_\pi^s}{\rho} \sum_{j=0}^{\infty} \left(\frac{1 - \gamma_\pi^s}{\rho}\right)^j \left(\frac{\gamma_x^s}{\gamma_\pi^s} \mathbb{E}_t x_{t+j} - \mathbb{E}_t i_{t+j}^*\right)
\end{aligned}$$

Then, denoting f_t as $-(i_t^* - \mathbb{E}_t \Delta p_{t+1}^*) + \frac{\gamma_x^s}{\gamma_\pi^s} x_t = -(i_t^* - \mathbb{E}_t \Delta p_{t+1}^*) + \frac{\gamma_x}{\gamma_\pi} x_t$,

$$z_{2,t} = \frac{\alpha\gamma_\pi^s}{\alpha\gamma_\pi^s - (\alpha - \rho)} \Delta p_t^* - \frac{\gamma_\pi^s}{\alpha\gamma_\pi^s - (\alpha - \rho)} \eta_t - \frac{\iota}{\gamma_\pi^s - (1 - \rho)} - \frac{\gamma_\pi^s}{\rho} \sum_{j=0}^{\infty} \left(\frac{1 - \gamma_\pi^s}{\rho}\right)^j \mathbb{E}_t f_{t+j} \quad (\text{a10})$$

Finally, plugging (a8) into (a3),

$$z_{3,t} = -i_{t-1}^* + v_t \quad (\text{a11})$$

Now, plugging (a10) and (a11) into (a5),

$$\begin{aligned}
\Delta e_t &= \frac{\alpha\gamma_\pi^s}{\alpha - \rho} \Delta p_t - \frac{\alpha\gamma_\pi^s}{\alpha - \rho} \Delta p_t^* + \frac{\gamma_\pi^s}{\alpha - \rho} \eta_t + \frac{\alpha\gamma_\pi^s - (\alpha - \rho)}{(\alpha - \rho)(\gamma_\pi^s - (1 - \rho))} \iota \\
&\quad + \frac{\gamma_\pi^s(\alpha\gamma_\pi^s - (\alpha - \rho))}{(\alpha - \rho)\rho} \sum_{j=0}^{\infty} \left(\frac{1 - \gamma_\pi^s}{\rho}\right)^j \mathbb{E}_t f_{t+j} + \frac{\alpha\gamma_\pi^s - (\alpha - \rho)}{\alpha - \rho} i_{t-1}^* - \frac{\alpha\gamma_\pi^s - (\alpha - \rho)}{\alpha - \rho} v_t
\end{aligned} \quad (\text{a12})$$

Updating (a12) once and applying the law of iterated expectations,

$$\begin{aligned}
\Delta e_{t+1} &= \hat{\iota} + \frac{\alpha\gamma_\pi^s}{\alpha - \rho} \Delta p_{t+1} - \frac{\alpha\gamma_\pi^s}{\alpha - \rho} \Delta p_{t+1}^* + \frac{\alpha\gamma_\pi^s - (\alpha - \rho)}{\alpha - \rho} i_t^* \\
&\quad + \frac{\gamma_\pi^s(\alpha\gamma_\pi^s - (\alpha - \rho))}{(\alpha - \rho)\rho} \sum_{j=0}^{\infty} \left(\frac{1 - \gamma_\pi^s}{\rho}\right)^j \mathbb{E}_t f_{t+j+1} + \omega_{t+1},
\end{aligned} \quad (\text{a13})$$

where

$$\hat{\iota} = \frac{\alpha\gamma_\pi^s - (\alpha - \rho)}{(\alpha - \rho)(\gamma_\pi^s - (1 - \rho))} \iota,$$

$$\begin{aligned}
\omega_{t+1} &= \frac{\gamma_\pi^s(\alpha\gamma_\pi^s - (\alpha - \rho))}{(\alpha - \rho)\rho} \sum_{j=0}^{\infty} \left(\frac{1 - \gamma_\pi^s}{\rho}\right)^j (\mathbb{E}_{t+1} f_{t+j+1} - \mathbb{E}_t f_{t+j+1}) \\
&\quad + \frac{\gamma_\pi^s}{\alpha - \rho} \eta_{t+1} - \frac{\alpha\gamma_\pi^s - (\alpha - \rho)}{\alpha - \rho} v_{t+1},
\end{aligned}$$

³⁶We use the fact $\mathbb{E}_t \eta_{t+j} = 0$, $j = 1, 2, \dots$.

and,

$$\mathbb{E}_t \omega_{t+1} = 0$$

B The Solution When $\alpha = \rho$

When α equals ρ , we have the following system of difference equations.

$$\begin{bmatrix} 1 & -1 & 0 \\ -\gamma_\pi^s & 1 & 0 \\ -\gamma_\pi^s & 0 & 1 \end{bmatrix} \begin{bmatrix} \mathbb{E}_t \Delta p_{t+1} \\ \mathbb{E}_t \Delta e_{t+1} \\ i_t \end{bmatrix} = \begin{bmatrix} \rho & -\rho & 0 \\ 0 & 0 & \rho \\ 0 & 0 & \rho \end{bmatrix} \begin{bmatrix} \Delta p_t \\ \Delta e_t \\ i_{t-1} \end{bmatrix} + \begin{bmatrix} \mathbb{E}_t \Delta p_{t+1}^* - \rho \Delta p_t^* + \eta_t \\ \iota + \gamma_x^s x_t - i_t^* \\ \iota + \gamma_x^s x_t \end{bmatrix}, \quad (\text{b1})$$

which can be represented by the following.

$$\mathbb{E}_t \mathbf{y}_{t+1} = \mathbf{V} \mathbf{\Lambda} \mathbf{V}^{-1} \mathbf{y}_t + \mathbf{c}_t, \quad (\text{b2})$$

where

$$\mathbf{V} = \begin{bmatrix} 0 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 0 \end{bmatrix}, \quad \mathbf{\Lambda} = \begin{bmatrix} \rho & 0 & 0 \\ 0 & \frac{\rho}{1-\gamma_\pi^s} & 0 \\ 0 & 0 & 0 \end{bmatrix}, \quad \mathbf{V}^{-1} = \begin{bmatrix} -1 & 1 & 0 \\ 1 & -1 & 1 \\ 0 & 1 & -1 \end{bmatrix}$$

The system yields the same eigenvalues, $\alpha = \rho$ and $\frac{\rho}{1-(1-\rho)\gamma_\pi}$. Therefore, when γ_π is greater than one, we have the saddle-path equilibrium as before. By pre-multiplying both sides of (b2) by \mathbf{V}^{-1} , we get,

$$\mathbb{E}_t \mathbf{z}_{t+1} = \mathbf{\Lambda} \mathbf{z}_t + \mathbf{h}_t, \quad (\text{b3})$$

where $\mathbf{V}^{-1} \mathbf{y}_t = \mathbf{z}_t$ and $\mathbf{V}^{-1} \mathbf{c}_t = \mathbf{h}_t$.

We solve the system as follows.

$$z_{1,t} = \sum_{j=0}^{\infty} \rho^j h_{1,t-j-1} + \sum_{j=0}^{\infty} \rho^j u_{t-j} \quad (\text{b4})$$

$$z_{2,t} = - \sum_{j=0}^{\infty} \left(\frac{1-\gamma_\pi^s}{\rho} \right)^{j+1} \mathbb{E}_t h_{2,t+j} \quad (\text{b5})$$

$$z_{3,t} = h_{3,t-1} + v_t, \quad (\text{b6})$$

where u_t and v_t are *any* martingale difference sequences.

Since $\mathbf{y}_t = \mathbf{V} \mathbf{z}_t$,

$$\begin{bmatrix} \Delta p_t \\ \Delta e_t \\ i_{t-1} \end{bmatrix} = \begin{bmatrix} 0 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 0 \end{bmatrix} \begin{bmatrix} z_{1,t} \\ z_{2,t} \\ z_{3,t} \end{bmatrix} \quad (\text{b7})$$

Now, we find the analytical solutions for \mathbf{z}_t . Since $\mathbf{h}_t = \mathbf{V}^{-1} \mathbf{c}_t$,

$$\mathbf{h}_t = \begin{bmatrix} -1 & 1 & 0 \\ 1 & -1 & 1 \\ 0 & 1 & -1 \end{bmatrix} \begin{bmatrix} (\mathbb{E}_t \Delta p_{t+1}^* - \rho \Delta p_t^* + \eta_t) + \iota + \gamma_x^s x_t - i_t^* \\ \gamma_\pi^s (\mathbb{E}_t \Delta p_{t+1}^* - \rho \Delta p_t^* + \eta_t) + \iota + \gamma_x^s x_t - i_t^* \\ \gamma_\pi^s (\mathbb{E}_t \Delta p_{t+1}^* - \rho \Delta p_t^* + \eta_t) + \iota + \gamma_x^s x_t - \gamma_\pi^s i_t^* \end{bmatrix},$$

thus,

$$h_{1,t} = -(1 - \gamma_\pi^s) (\mathbb{E}_t \Delta p_{t+1}^* - \rho \Delta p_t^* + \eta_t) \quad (\text{b8})$$

$$h_{2,t} = \mathbb{E}_t \Delta p_{t+1}^* - \rho \Delta p_t^* + \eta_t + \iota + \gamma_x^s x_t - \gamma_\pi^s i_t^* \quad (\text{b9})$$

$$h_{3,t} = -(1 - \gamma_\pi^s) i_t^* \quad (\text{b10})$$

From (b4) and (b8),

$$\begin{aligned} z_{1,t} &= -(1 - \gamma_\pi^s) \sum_{j=0}^{\infty} \rho^j (\Delta p_{t-j}^* - \rho \Delta p_{t-j-1}^* + \eta_{t-j-1}) + \sum_{j=0}^{\infty} \rho^j u_{t-j} \\ &= -(1 - \gamma_\pi^s) \Delta p_t^* + \sum_{j=0}^{\infty} \rho^j u_{t-j} - (1 - \gamma_\pi^s) \sum_{j=0}^{\infty} \rho^j \eta_{t-j-1} \end{aligned} \quad (\text{b11})$$

From (b5) and (b9),

$$\begin{aligned} z_{2,t} &= - \sum_{j=0}^{\infty} \left(\frac{1 - \gamma_\pi^s}{\rho} \right)^{j+1} (\mathbb{E}_t \Delta p_{t+j+1}^* - \rho \mathbb{E}_t \Delta p_{t+j}^* + \mathbb{E}_t \eta_{t+j} + \iota + \gamma_x^s \mathbb{E}_t x_{t+j} - \gamma_\pi^s \mathbb{E}_t i_{t+j}^*) \\ &= (1 - \gamma_\pi^s) \Delta p_t^* - \left(\frac{1 - \gamma_\pi^s}{\rho} \right) \eta_t - \frac{(1 - \gamma_\pi^s) \iota}{\rho - (1 - \gamma_\pi^s)} \\ &\quad - \gamma_\pi^s \sum_{j=0}^{\infty} \left(\frac{1 - \gamma_\pi^s}{\rho} \right)^{j+1} \left(\mathbb{E}_t \Delta p_{t+j+1}^* + \frac{\gamma_x^s}{\gamma_\pi^s} \mathbb{E}_t x_{t+j} - \mathbb{E}_t i_{t+j}^* \right) \end{aligned} \quad (\text{b12})$$

Denoting f_t as $-(i_t^* - \mathbb{E}_t \Delta p_{t+1}^*) + \frac{\gamma_x^s}{\gamma_\pi^s} x_t = -(i_t^* - \mathbb{E}_t \Delta p_{t+1}^*) + \frac{\gamma_x^s}{\gamma_\pi^s} x_t$,

$$z_{2,t} = (1 - \gamma_\pi^s) \Delta p_t^* - \gamma_\pi^s \sum_{j=0}^{\infty} \left(\frac{1 - \gamma_\pi^s}{\rho} \right)^{j+1} \mathbb{E}_t f_{t+j} - \left(\frac{1 - \gamma_\pi^s}{\rho} \right) \eta_t - \frac{(1 - \gamma_\pi^s) \iota}{\rho - (1 - \gamma_\pi^s)} \quad (\text{b13})$$

From (b6) and (b10),

$$z_{3,t} = -(1 - \gamma_\pi^s) i_{t-1}^* + v_t \quad (\text{b14})$$

From (b7), (b13), and (b14),

$$\begin{aligned} \Delta p_t &= (1 - \gamma_\pi^s) \Delta p_t^* - \gamma_\pi^s \sum_{j=0}^{\infty} \left(\frac{1 - \gamma_\pi^s}{\rho} \right)^{j+1} \mathbb{E}_t f_{t+j} \\ &\quad - \left(\frac{1 - \gamma_\pi^s}{\rho} \right) \eta_t + \frac{(1 - \gamma_\pi^s)}{(1 - \gamma_\pi^s) - \rho} \iota - (1 - \gamma_\pi^s) i_{t-1}^* + v_t \end{aligned} \quad (\text{b15})$$

Updating (b15) once and applying the law of iterated expectations,

$$\Delta p_{t+1} = \hat{\iota} + (1 - \gamma_\pi^s) \Delta p_{t+1}^* - (1 - \gamma_\pi^s) i_t^* - \gamma_\pi^s \sum_{j=0}^{\infty} \left(\frac{1 - \gamma_\pi^s}{\rho} \right)^{j+1} \mathbb{E}_t f_{t+j} + \omega_{t+1}, \quad (\text{b16})$$

where

$$\hat{l} = \frac{(1 - \gamma_\pi^s)}{(1 - \gamma_\pi^s) - \rho} l,$$

$$\omega_{t+1} = -\gamma_\pi^s \sum_{j=0}^{\infty} \left(\frac{1 - \gamma_\pi^s}{\rho} \right)^j (\mathbb{E}_{t+1} f_{t+j+1} - \mathbb{E}_t f_{t+j+1}) - \left(\frac{1 - \gamma_\pi^s}{\rho} \right) \eta_{t+1} + v_{t+1},$$

and

$$\mathbb{E}_t \omega_{t+1} = 0$$

Note that there is no inertia for domestic inflation in this solution, since there is no backward looking component. Put it differently, when there is a shock, Δp_{t+1} instantly *jumps* to its long-run equilibrium.

On the contrary, Δe_{t+1} does have inertia. From (b7),

$$\Delta e_t = z_{1,t} + \Delta p_t \tag{b17}$$

Plug (b11) into (b17) and update it once to get,

$$\Delta e_{t+1} = \Delta p_{t+1} - (1 - \gamma_\pi^s) \Delta p_{t+1}^* + \sum_{j=0}^{\infty} \rho^j u_{t-j+1} - (1 - \gamma_\pi^s) \sum_{j=0}^{\infty} \rho^j \eta_{t-j}, \tag{b18}$$

where Δp_{t+1} contains rational expectation of future fundamentals as defined in (b16). Note that Δe_{t+1} exhibits inertia due to the presence of the martingale difference sequences.

In a nutshell, in the special case of $\rho = \alpha$, domestic inflation instantly jumps to its long-run equilibrium and all convergence will be carried over by the exchange rate adjustments.

C GMM with a Near Unit Root and the Grid Bootstrap

C.1 Asymptotic Distribution

When the variables are jointly stationary, then the t -ratio $t_n(\alpha)$ is asymptotically normal and both conventional inference and the grid bootstrap method provide valid methods for confidence interval coverage. We are interested in the case where the persistence parameter α is large and possibly equal to one. The appropriate way to incorporate this into an asymptotic distribution theory is to model α as local to 1, e.g.

$$\alpha = 1 + c/n. \quad (\text{c1})$$

With this reparameterization, the localizing parameter c indexes the degree of persistence.

Set $\beta = (\alpha, d, \phi)$ where ϕ are the parameters in (19)-(20) in addition to α and d . Let $\mathbf{m}_{t+1}(\beta)$ be the list of moment functions in (19)-(20) and set

$$\mathbf{g}_t(\beta) = \begin{pmatrix} s_t (s_{t+1} - d - \alpha s_t) \\ s_{t+1} - d - \alpha s_t \\ \mathbf{m}_{t+1}(\beta) \end{pmatrix},$$

which is the set of moment functions (18)-(20). Define

$$\begin{aligned} \bar{\mathbf{g}}_n(\beta) &= \frac{1}{n} \sum_{t=1}^n \mathbf{g}_t(\beta) \\ \mathbf{\Omega}_n(\beta) &= \frac{1}{n} \sum_{t=1}^n \mathbf{g}_t(\beta) \mathbf{g}_t(\beta)' \\ \mathbf{G}_n(\beta) &= \frac{1}{n} \sum_{t=1}^n \frac{\partial}{\partial \beta'} \mathbf{g}_t(\beta). \end{aligned}$$

Let \mathbf{m}_{t+1} , \mathbf{g}_t , $\bar{\mathbf{g}}_n$, $\mathbf{\Omega}_n$ and \mathbf{G}_n denote these functions evaluated at the true β . Also, define the moments $\sigma_\varepsilon^2 = \mathbb{E}\varepsilon_{t+1}^2$, $\mu = \mathbb{E}\mathbf{m}_{t+1}\varepsilon_{t+1}$, $\mathbf{Q} = \mathbb{E}\frac{\partial}{\partial \phi'} \mathbf{m}_{t+1}(\beta)$ and $\mathbf{M} = \mathbb{E}\mathbf{m}_{t+1}\mathbf{m}'_{t+1}$.

Given a preliminary estimator $\tilde{\beta}$, the GMM estimator $\hat{\beta}$ minimizes $\bar{\mathbf{g}}_n(\beta)' \mathbf{\Omega}_n(\tilde{\beta})^{-1} \bar{\mathbf{g}}_n(\beta)$.

It is well known that under standard conditions, the GMM estimator has the asymptotic linear representation

$$\sqrt{n} (\hat{\beta} - \beta) = (\mathbf{G}'_n \mathbf{\Omega}_n^{-1} \mathbf{G}_n)^{-1} \mathbf{G}'_n \mathbf{\Omega}_n^{-1} \sqrt{n} \bar{\mathbf{g}}_n + o_p(1). \quad (\text{c2})$$

To obtain an asymptotic distribution under the local-to-unity assumption (c1) we have to introduce additional scale factors so that the moment matrices have non-degenerate limiting distributions.

We define

$$D_n = \begin{bmatrix} n^{1/2} & 0 \\ 0 & \mathbf{I}_{\ell+1} \end{bmatrix}$$

where $\ell = \dim(\mathbf{m}_t)$, and

$$\delta_n = \begin{bmatrix} n^{1/2} & 0 \\ 0 & \mathbf{I}_{p+1} \end{bmatrix}$$

where $p = \dim(\beta)$. Then we can write (c2) equivalently as

$$\sqrt{n}\delta_n(\hat{\beta} - \beta) = \left(\overline{\mathbf{G}}_n' \overline{\boldsymbol{\Omega}}_n^{-1} \overline{\mathbf{G}}_n\right)^{-1} \overline{\mathbf{G}}_n' \overline{\boldsymbol{\Omega}}_n^{-1} \sqrt{n}\delta_n^{-1} \overline{\mathbf{g}}_n + o_p(1) \quad (\text{c3})$$

where

$$\overline{\boldsymbol{\Omega}}_n = D_n^{-1} \boldsymbol{\Omega}_n D_n^{-1}$$

and

$$\overline{\mathbf{G}}_n = D_n^{-1} \mathbf{G}_n \delta_n^{-1}.$$

Since the errors ε_{t+1} and \mathbf{m}_{t+1} are martingale differences, then

$$\frac{1}{\sqrt{n}} \sum_{t=1}^{[nr]} \begin{pmatrix} \varepsilon_{t+1} \\ \mathbf{m}_{t+1} \end{pmatrix} \Rightarrow W(r),$$

a Brownian motion with covariance matrix

$$\mathbb{E} \begin{pmatrix} \varepsilon_{t+1}^2 & \varepsilon_{t+1} \mathbf{m}'_{t+1} \\ \mathbf{m}_{t+1} \varepsilon_{t+1} & \mathbf{m}_{t+1} \mathbf{m}'_{t+1} \end{pmatrix} = \begin{pmatrix} \sigma_\varepsilon^2 & \mu \\ \mu & \mathbf{M} \end{pmatrix},$$

Partition $W(r) = (W_1(r), W_2(r))$. Under the local-to-unity assumption (c1)

$$n^{-1/2} s_{[nr]} \Rightarrow W_{1c}(r)$$

where $dW_{1c}(r) = cW_{1c}(r) + dW_1(r)$ is a standard diffusion process.

It follows that

$$\begin{aligned} \sqrt{n}\delta_n^{-1} \overline{\mathbf{g}}_n &= \begin{pmatrix} \frac{1}{n} \sum_{t=1}^n s_t \varepsilon_{t+1} \\ \frac{1}{\sqrt{n}} \sum_{t=1}^n \varepsilon_{t+1} \\ \frac{1}{\sqrt{n}} \sum_{t=1}^n \mathbf{m}_{t+1} \end{pmatrix} \\ &\Rightarrow \begin{pmatrix} \int_0^1 W_{1c} dW_1 \\ W_1(1) \\ W_2(1) \end{pmatrix} \\ &\equiv \mathbf{N}_c, \end{aligned} \quad (\text{c4})$$

$$\begin{aligned}
\bar{\Omega}_n &= \begin{pmatrix} \frac{1}{n^2} \sum_{t=1}^n s_t^2 \varepsilon_{t+1}^2 & \frac{1}{n^{3/2}} \sum_{t=1}^n s_t \varepsilon_{t+1}^2 & \frac{1}{n} \sum_{t=1}^n s_t \varepsilon_{t+1} \mathbf{m}'_{t+1} \\ \frac{1}{n^{3/2}} \sum_{t=1}^n s_t \varepsilon_{t+1}^2 & \frac{1}{n} \sum_{t=1}^n \varepsilon_{t+1} & \frac{1}{n} \sum_{t=1}^n \varepsilon_{t+1} \mathbf{m}'_{t+1} \\ \frac{1}{n} \sum_{t=1}^n s_t \mathbf{m}_{t+1} \varepsilon_{t+1} & \frac{1}{n} \sum_{t=1}^n \mathbf{m}_{t+1} \varepsilon_{t+1} & \frac{1}{n} \sum_{t=1}^n \mathbf{m}_{t+1} \mathbf{m}'_{t+1} \end{pmatrix} \\
&\Rightarrow \begin{pmatrix} \int_0^1 W_{1c}^2 \sigma_\varepsilon^2 & \int_0^1 W_{1c} \sigma_\varepsilon^2 & \int_0^1 W_{1c} \mu' \\ \int_0^1 W_{1c} \sigma_\varepsilon^2 & \sigma_\varepsilon^2 & \mu' \\ \int_0^1 W_{1c} \mu & \mu & \mathbf{M} \end{pmatrix} \\
&\equiv \mathbf{\Omega}_c,
\end{aligned} \tag{c5}$$

and

$$\begin{aligned}
\bar{\mathbf{G}}_n &= \begin{pmatrix} -\frac{1}{n^2} \sum_{t=1}^n s_t^2 & -\frac{1}{n^{3/2}} \sum_{t=1}^n s_t & 0 \\ -\frac{1}{n^{3/2}} \sum_{t=1}^n s_t & -1 & 0 \\ \frac{1}{n^{3/2}} \sum_{t=1}^n \frac{\partial}{\partial \alpha} \mathbf{m}_{t+1}(\beta) & 0 & \frac{1}{n} \sum_{t=1}^n \frac{\partial}{\partial \phi'} \mathbf{m}_{t+1}(\beta) \end{pmatrix} \\
&\Rightarrow \begin{pmatrix} -\int_0^1 W_{1c}^2 & -\int_0^1 W_{1c} & 0 \\ -\int_0^1 W_{1c} & -1 & 0 \\ 0 & 0 & \mathbf{Q} \end{pmatrix} \\
&\equiv \mathbf{G}_c.
\end{aligned} \tag{c6}$$

Applying these distributional results to (c3), we find

$$\sqrt{n} \delta_n (\hat{\beta} - \beta) \Rightarrow (\mathbf{G}'_c \mathbf{\Omega}_c^{-1} \mathbf{G}_c)^{-1} \mathbf{G}'_c \mathbf{\Omega}_c^{-1} \mathbf{N}_c. \tag{c7}$$

The asymptotic distribution of $\hat{\alpha}$ is obtained by taking the first element of this vector. Let $\mathbf{S}_1 = (1 \ 0)'$ be a $(p+2) \times 1$ unit vector. Then

$$n(\hat{\alpha} - \alpha) \Rightarrow \mathbf{S}'_1 (\mathbf{G}'_c \mathbf{\Omega}_c^{-1} \mathbf{G}_c)^{-1} \mathbf{G}'_c \mathbf{\Omega}_c^{-1} \mathbf{N}_c. \tag{c8}$$

The standard error for $\hat{\alpha}$ is

$$\begin{aligned}
n \text{ se}(\hat{\alpha}) &= \left(n \mathbf{S}'_1 (\mathbf{G}'_n \mathbf{\Omega}_n^{-1} \mathbf{G}_n)^{-1} \mathbf{S}_1 \right)^{1/2} \\
&= \left(\mathbf{S}'_1 (\bar{\mathbf{G}}'_n \bar{\mathbf{\Omega}}_n^{-1} \bar{\mathbf{G}}_n)^{-1} \mathbf{S}_1 \right)^{1/2} \\
&\Rightarrow \left(\mathbf{S}'_1 (\mathbf{G}'_c \mathbf{\Omega}_c^{-1} \mathbf{G}_c)^{-1} \mathbf{S}_1 \right)^{1/2}.
\end{aligned}$$

Thus the t-ratio for α has the asymptotic distribution

$$t_n(\alpha) = \frac{\hat{\alpha} - \alpha}{se(\hat{\alpha})} \Rightarrow \frac{\mathbf{S}'_1 (\mathbf{G}'_c \boldsymbol{\Omega}_c^{-1} \mathbf{G}_c)^{-1} \mathbf{G}'_c \boldsymbol{\Omega}_c^{-1} \mathbf{N}_c}{\left(\mathbf{S}'_1 (\mathbf{G}'_c \boldsymbol{\Omega}_c^{-1} \mathbf{G}_c)^{-1} \mathbf{S}_1 \right)^{1/2}}.$$

We state this formally.

Proposition 1: Under (c1),

$$t_n(\alpha) \Rightarrow \frac{\mathbf{S}'_1 (\mathbf{G}'_c \boldsymbol{\Omega}_c^{-1} \mathbf{G}_c)^{-1} \mathbf{G}'_c \boldsymbol{\Omega}_c^{-1} \mathbf{N}_c}{\left(\mathbf{S}'_1 (\mathbf{G}'_c \boldsymbol{\Omega}_c^{-1} \mathbf{G}_c)^{-1} \mathbf{S}_1 \right)^{1/2}} \quad (\text{c9})$$

where \mathbf{N}_c , $\boldsymbol{\Omega}_c$, and \mathbf{G}_c are defined in (c4), (c5), and (c6).

In the special case that ε_{t+1} and \mathbf{m}_{t+1} are uncorrelated, then $\mu = 0$ and both $\boldsymbol{\Omega}_c$ and \mathbf{G}_c are block diagonal. Then $\hat{\alpha}$ is asymptotically independent of $\hat{\beta}$ and $t_n(\alpha)$ has a classic Dickey-Fuller distribution.

However, when ε_{t+1} and \mathbf{m}_{t+1} are correlated so that $\mu \neq 0$ then $\hat{\alpha}$ and $\hat{\beta}$ are not asymptotically independent. In this case the asymptotic distribution in Proposition 1 is a mixture of a non-standard Dickey-Fuller and a standard normal, similar to the result by Hansen (1995) for the case of unit root testing with covariates. The situation is actually quite similar, as the GMM estimator is a combination of the (non-standard) least-squares estimator of α with a set of classic moment restrictions.

C.2 Grid Bootstrap

As discussed in Beran (1987) and Hansen (1999) Proposition 1, conventional bootstrap confidence intervals have asymptotic first-order correct coverage if the parameter estimates (used to construct the bootstrap distribution) are consistent for the true values, and the asymptotic distribution is continuous in the parameters. Furthermore, the conventional bootstrap generically fails to have asymptotic first-order correct coverage if these conditions fail.

This theory, plus the distribution theory of Proposition 1 above, helps us understand why the conventional bootstrap will not have correct coverage. The asymptotic distribution (c9) depends on the parameters c , σ_ε^2 , μ , \mathbf{M} , and \mathbf{Q} . The parameter $c = n(\alpha - 1)$ is estimated by $\hat{c} = n(\hat{\alpha} - 1)$ which is inconsistent as shown in (c8). Consequently, the conventional bootstrap will not have correct coverage.

In contrast, as discussed in Hansen (1999) Proposition 1, the grid bootstrap confidence interval for α has asymptotic first-order correct coverage if the remaining parameter estimates are consistent for the true values and the asymptotic distribution of $t_n(\alpha)$ is continuous in the parameters. First, we see by direct examination that the distribution in (c9) is a continuous function of the parameters c , σ_ε^2 , μ , \mathbf{M} , and \mathbf{Q} . Second, the moments σ_ε^2 , μ , \mathbf{M} , and \mathbf{Q} are identified and are consistently

estimated by sample averages. For fixed α (equivalently, fixed c) the residual bootstrap method will consistently estimate these population moments under the auxiliary assumption that the underlying errors are iid. This meets the conditions for the grid bootstrap and we conclude that the interval for α has asymptotic first-order correct coverage.

Assumption 1: The error vector $(\varepsilon_{t+1}, \xi_{t+1}, \nu_{t+1})$ is independent and identically distributed, and has finite $2 + \delta$ moments for some $\delta > 0$. The local-to-unity condition (c1) holds, the autoregressive roots of (14) lie outside the unit circle, and the set of moment equations (18)-(19)-(20) satisfy the standard conditions for GMM estimation.

Proposition 2: Let A denote the grid bootstrap confidence interval defined in (23). Under Assumption 1, $\mathbb{P}(\alpha \in A) \rightarrow 0.95$ as $n \rightarrow \infty$.

We are slightly informal here regarding the regularity conditions and therefore state this result as a proposition rather than as a formal theorem.

There are two important caveats regarding this result.

First, the grid bootstrap confidence interval only works for α , not for the other parameters. This is because the asymptotic distribution (c7) suggests that the distribution of the entire estimator vector is non-standard and a function of c , and the grid bootstrap method only “solves” the confidence interval problem for the single parameter which is the source of the non-pivotalness, in this case α . In the present context this is satisfactory, as our interest focuses on the persistence parameter α .

Second, our grid bootstrap method relies on the residual bootstrap, and is therefore critically dependent on the assumption that the errors are iid. In particular, this excludes conditional heteroskedasticity. As our application concerns quarterly observations where the degree of heteroskedasticity is mild, we believe that this limitation is not too restrictive.

D Stochastic Simulations

In this Appendix, we show the derivation of the model used for the Monte Carlo simulations.

D.1 Household

Households in the domestic country maximize

$$\mathbb{E}_0 \sum_{t=0}^{\infty} \beta^t \ln(C_t),$$

subject to the budget constraint:

$$B_t + E_t B_t^* + P_t C_t = (1 + i_{t-1}) B_{t-1} + (1 + i_{t-1}^*) E_t B_{t-1}^* + Y_t + T_t.$$

We denote consumption by C_t , domestic bond holdings by B_t , nominal exchange rate by E_t , foreign bond holdings by B_t^* , the consumer price index by P_t , domestic nominal interest rate by i_t , foreign nominal interest rate by i_t^* , output by Y_t and the lump-sum tax by T_t .

From the optimality conditions in both home and foreign households' maximization problem, we can derive the UIP condition:

$$i_t = \mathbb{E}_t \Delta e_{t+1} + i_t^* + v_t, \quad (\text{d1})$$

where v_t denotes the UIP shock. Lower case variables indicates the logarithm of upper case variables.

D.2 Firms (Price Setting)

We assume that firms' price setting is summarized in the gradual adjustment model by Mussa (1982):

$$\mathbb{E}_t \Delta p_{t+1} = b [\mu - (p_t - p_t^* - e_t)] + \mathbb{E}_t \Delta p_{t+1}^* + \mathbb{E}_t \Delta e_{t+1}, \quad (\text{d2})$$

which could be transformed into

$$\begin{aligned} \mathbb{E}_t \Delta p_{t+1} &= (1-b) \Delta p_t - (1-b) \Delta p_t^* - (1-b) \Delta e_t + \mathbb{E}_t \Delta p_{t+1}^* + \mathbb{E}_t \Delta e_{t+1} \\ &\quad + (p_t^* - \mathbb{E}_{t-1} p_t^*) + (e_t - \mathbb{E}_{t-1} e_t) - (p_t - \mathbb{E}_{t-1} p_t). \end{aligned} \quad (\text{25})$$

D.3 Central Bank

The domestic central banks follows the Taylor type rule:

$$i_t = (1 - \rho) (\gamma_\pi \mathbb{E}_t \Delta \hat{p}_{t+1} + \gamma_x x_t) + \rho i_{t-1} + u_t, \quad (\text{d4})$$

where x_t and u_t denote output gap and the monetary policy shock.

D.4 System of Equations

From equations (d1), (d2), (25), and (d4), we can derive the system of equations for π_t^e , q_t , s_t^e , $\Delta\hat{p}_t$, $\Delta\hat{e}_t$, i_t and π_t^{*e} as follows:

$$\begin{aligned}\pi_t^e &= q_t + \mathbb{E}_t \Delta p_{t+1}^* + s_t^e, \\ \Delta\hat{p}_t &= (1-b)\pi_{t-1}^e - (1-b)\Delta\hat{e}_t + s_t^e + (\Delta\hat{p}_t^* - \pi_{t-1}^{*e}) + (\Delta\hat{e}_t - s_{t-1}^e) - (1-b)\Delta\hat{p}_t^* + \mathbb{E}_t \Delta\hat{p}_{t+1}^*, \quad (\text{d5}) \\ i_t &= (1-\rho)\gamma_\pi \pi_t^e + \rho i_{t-1} + (1-\rho)\gamma_x x_t + u_t, \\ i_t &= s_t^e + i_t^* + v_t, \\ \pi_t^e &\equiv \mathbb{E}_t \Delta p_{t+1}, \\ \pi_t^{*e} &\equiv \mathbb{E}_t \Delta p_{t+1}^*, \\ s_t^e &\equiv \mathbb{E}_t \Delta e_{t+1},\end{aligned}$$

where we define the real exchange rate q_t :

$$q_t \equiv p_t^* + e_t - p_t - \mu.$$

It is assumed that the price is set one period in advance when deriving equation (d5). Three exogenous variables i_t^* , Δp_t^* and x_t are assumed to follow the VAR(2) model:

$$\begin{bmatrix} i_t^* \\ \Delta p_t^* \\ x_t \end{bmatrix} = A \begin{bmatrix} i_{t-1}^* \\ \Delta p_{t-1}^* \\ x_{t-1} \end{bmatrix} + B \begin{bmatrix} i_{t-2}^* \\ \Delta p_{t-2}^* \\ x_{t-3} \end{bmatrix} + \begin{bmatrix} e_{i,t} \\ e_{\Delta p^*,t} \\ e_{x,t} \end{bmatrix}.$$

Parameter matrices A and B as well as the variances of shocks are estimated for the sample from 1979:III to 1998: IV. The US is supposed to be the home country. Hence, i_t^* and Δp_t^* are weighted average of policy interest rates and CPI inflation rates for Australia, Austria, Belgium, Canada, Finland, France, Germany, Italy, Japan, Netherlands, Portugal, Spain, Sweden, Switzerland and United Kingdom. We use the weight by the FRB for the effective exchange rate. For x_t , we use unemployment rate gap in Boivin (2006).

D.5 Calibration

We use the estimated parameters for simulation. They are set as follows.

b	γ_π	γ_x	ρ	σ_u
0.168	2.435	0.162	0.796	0.109

References

- ANDREWS, D. W. K. (1993): “Exactly Median-Unbiased Estimation of First Order Autoregressive/Unit Root Models,” *Econometrica*, 61, 139–165.
- ANDREWS, D. W. K., AND H.-Y. CHEN (1994): “Approximately Median-Unbiased Estimation of Autoregressive Models,” *Journal of Business and Economic Statistics*, 12, 187–204.
- BERAN, R. (1987): “Prepivoting to Reduce Level Error of Confidence Sets,” *Biometrika*, 74, 457–468.
- BILS, M., AND P. J. KLENOW (2005): “Some Evidence on the Importance of Sticky Prices,” *Journal of Political Economy*, 112, 947–985.
- BOIVIN, J. (2006): “Has US Monetary Policy Changed? Evidence from Drifting Coefficients and Real Time Data,” *Journal of Money, Credit, and Banking*, 38, 1149–1173.
- BRODA, C., AND D. E. WEINSTEIN (2008): “Understanding International Price Differences Using Barcode Data,” NBER Working Paper No. 14017.
- CALVO, G. (1983): “Staggered Prices in a Utility Maximizing Framework,” *Journal of Monetary Economics*, 12, 383–398.
- CARVALHO, C., AND F. NECHIO (2011): “Aggregation and the PPP Puzzle in a Sticky-Price Model,” *American Economic Review*, forthcoming.
- CHEN, S.-S., AND C. ENGEL (2005): “Does ‘Aggregation Bias’ Explain the PPP Puzzle?,” *Pacific Economic Review*, 10, 49–72.
- CHRISTIANO, L. J., M. EICHENBAUM, AND C. L. EVANS (2005): “Nominal Rigidities and the Dynamic Effects of a Shock to Monetary Policy,” *Journal of Political Economy*, 113, 1–45.
- CLARIDA, R., J. GALÍ, AND M. GERTLER (1998): “Monetary Policy Rules in Practice: Some International Evidence,” *European Economic Review*, 42, 1033–1067.
- (2000): “Monetary Policy Rules and Macroeconomic Stability: Evidence and Some Theory,” *Quarterly Journal of Economics*, 115, 147–180.
- CLARIDA, R., AND D. WALDMAN (2007): “Is Bad News About Inflation Good News for the Exchange Rate?,” NBER Working Paper No. 13010.
- CRUCINI, M. J., AND M. SHINTANI (2008): “Persistence in Law-of-One-Price Deviations: Evidence from Micro-Data,” *Journal of Monetary Economics*, 55, 629–644.
- DEL NEGRO, M., AND F. SCHORFHEIDE (2008): “Forming Priors for DSGE Models (and How It Affects the Assessment of Nominal Rigidities),” *Journal of Monetary Economics*, 55, 1191–1208.

- EFRON, B., AND R. J. TIBSHIRANI (1993): *An Introduction to the Bootstrap*. Chapman and Hall/CRC, London, UK.
- EICHENBAUM, M., AND J. D. FISHER (2007): “Estimating the Frequency of Price Re-optimization in Calvo-Style Models,” *Journal of Monetary Economics*, 54, 2032–2047.
- ENGEL, C., AND K. D. WEST (2005): “Exchange Rates and Fundamentals,” *Journal of Political Economy*, 113, 485–517.
- (2006): “Taylor Rules and the Deutschmark-Dollar Exchange Rate,” *Journal of Money, Credit, and Banking*, 38, 1175–1194.
- GALÍ, J., AND M. GERTLER (1999): “Inflation Dynamics: A Structural Econometric Analysis,” *Journal of Monetary Economics*, 44, 195–222.
- HANSEN, B. E. (1995): “Rethinking the Univariate Approach to Unit Root Testing,” *Econometric Theory*, 11, 1148–1171.
- (1999): “The Grid Bootstrap and the Autoregressive Model,” *Review of Economics and Statistics*, 81, 594–607.
- HANSEN, L. P. (1982): “Large Sample Properties of Generalized Method of Moments Estimators,” *Econometrica*, 50, 1029–1054.
- HANSEN, L. P., AND T. J. SARGENT (1980): “Formulating and Estimating Dynamic Linear Rational Expectations Models,” *Journal of Economic Dynamics and Control*, 2, 7–46.
- (1981): “A Note on Wiener-Kolmogorov Prediction Formulas for Rational Expectations Models,” Research Department Staff Report 69, Federal Reserve Bank of Minneapolis, Minneapolis, MN.
- (1982): “Instrumental Variables Procedures for Estimating Linear Rational Expectations Models,” *Journal of Monetary Economics*, 9, 263–296.
- IMBS, J., H. MUMTAZ, M. O. RAVN, AND H. REY (2005): “PPP Strikes Back: Aggregation and the Real Exchange Rates,” *Quarterly Journal of Economics*, 120, 1–43.
- KEHOE, P. J., AND V. MIDRIGAN (2007): “Sticky Prices and Sectoral Real Exchange Rates,” Federal Reserve Bank of Minneapolis Working Paper 656.
- KILIAN, L., AND T. ZHA (2002): “Quantifying the Uncertainty about the Half-Life of Deviations from PPP,” *Journal of Applied Econometrics*, 17, 107–125.
- KIM, H. (2006): “Essays on Exchange Rate Models under a Taylor Rule Type Monetary Policy,” Ph.D. Dissertation, Ohio State University.

- KIM, J. (2005): “Convergence Rates to PPP for Traded and Non-Traded Goods: A Structural Error Correction Model Approach,” *Journal of Business and Economic Statistics*, 23, 76–86.
- KIM, J., AND M. OGAKI (2004): “Purchasing Power Parity for Traded and Non-Traded Goods: A Structural Error Correction Model Approach,” *Monetary and Economic Studies*, 22, 1–26.
- KIM, J., M. OGAKI, AND M.-S. YANG (2007): “Structural Error Correction Models: A System Method for Linear Rational Expectations Models and an Application to an Exchange Rate Model,” *Journal of Money, Credit, and Banking*, 39, 2057–2075.
- KLEIBERGEN, F., AND S. MAVROEIDIS (2009): “Weak Instrument Robust Tests in GMM and the New Keynesian Phillips Curve,” *Journal of Business and Economic Statistics*, 27, 293–311.
- MAGNUSSON, L. M., AND S. MAVROEIDIS (2010): “Identification-Robust Minimum Distance Estimation of the New Keynesian Phillips Curve,” *Journal of Money, Credit and Banking*, 42, 465–481.
- MARK, N. C. (2001): *International Macroeconomics and Finance: Theory and Econometric Methods*. Blackwell Publishers, Oxford, UK.
- (2009): “Changing Monetary Policy Rules, Learning, and Real Exchange Rate Dynamics,” *Journal of Money, Credit and Banking*, 41, 1047–1070.
- MOLODTSOVA, T., A. NIKOLSKO-RZHEVSKYY, AND D. H. PAPELL (2008): “Taylor Rule with Real-Time Data: A Tale of Two Countries and One Exchange Rate,” *Journal of Monetary Economics*, 55, S63–S79.
- MOLODTSOVA, T., AND D. H. PAPELL (2009): “Out-of-Sample Exchange Rate Predictability with Taylor Rule Fundamentals,” *Journal of International Economics*, 77, 167–180.
- MURRAY, C. J., AND D. H. PAPELL (2002): “The Purchasing Power Parity Persistence Paradigm,” *Journal of International Economics*, 56, 1–19.
- (2005): “Do Panels Help Solve the Purchasing Power Parity Puzzle?,” *Journal of Business and Economic Statistics*, 23, 410–415.
- MUSSA, M. (1982): “A Model of Exchange Rate Dynamics,” *Journal of Political Economy*, 90, 74–104.
- NAKAMURA, E., AND J. STEINSSON (2008): “Five Facts About Prices: A Reevaluation of Menu Cost Models,” *Quarterly Journal of Economics*, 123, 1415–1464.
- NG, S., AND P. PERRON (2001): “Lag Length Selection and the Construction of Unit Root Tests with Good Size and Power,” *Econometrica*, 69, 1519–1554.

- OGAKI, M. (1993): “Generalized Method of Moments: Econometric Applications,” in *Handbook of Statistics, Vol. II*, ed. by G. S. Maddala, C. R. Rao, and H. D. Vinod, pp. 455–488, Amsterdam. North-Holland.
- PARSLEY, D. C., AND S.-J. WEI (2007): “A Prism into the PPP Puzzles: The Micro-Foundations of Big Mac Real Exchange Rates,” *Economic Journal*, 117, 1336–1356.
- ROGOFF, K. (1996): “The Purchasing Power Parity Puzzle,” *Journal of Economic Literature*, 34, 647–668.
- ROSSI, B. (2005): “Confidence Intervals for Half-Life Deviations From Purchasing Power Parity,” *Journal of Business and Economic Statistics*, 23, 432–442.
- SARGENT, T. J. (1987): *Macroeconomic Theory*. Academic Press, New York, second edn.
- SBORDONE, A. M. (2002): “Prices and Unit Labor Costs: A New Test of Price Stickiness,” *Journal of Monetary Economics*, 49, 265–292.
- (2005): “Do Expected Future Marginal Costs Drive Inflation Dynamics?,” *Journal of Monetary Economics*, 52, 1183–1197.
- SMETS, F., AND R. WOUTERS (2007): “Shocks and Frictions in U.S. Business Cycles: A Bayesian DSGE Approach,” *American Economic Review*, 97, 586–606.
- STEINSSON, J. (2008): “The Dynamic Behavior of the Real Exchange Rate in Sticky-Price Models,” *American Economic Review*, 98, 519–533.
- STOCK, J. H., AND M. W. WATSON (1993): “A Simple Estimator of Cointegrating Vectors in Higher Order Integrated Systems,” *Econometrica*, 61, 783–820.
- SVENSSON, L. E. (2000): “Open-Economy Inflation Targeting,” *Journal of International Economics*, 50, 155–183.
- WEST, K. D. (1987): “A Specification Test for Speculative Bubbles,” *Quarterly Journal of Economics*, 102, 553–580.
- (1989): “Estimation of Linear Rational Expectations Models, In the Presence of Deterministic Terms,” *Journal of Monetary Economics*, 24, 437–442.

Table 1. GMM Estimation of the US Taylor Rule Estimation

Deviation	Sample Period	γ_π (s.e.)	γ_x (s.e.)	ρ (s.e.)
Real GDP	1959:Q1-2003:Q4	1.466 (0.190)	0.161 (0.054)	0.820 (0.029)
	1959:Q1-1979:Q2	0.605 (0.099)	0.577 (0.183)	0.708 (0.056)
	1979:Q3-2003:Q4	2.517 (0.306)	0.089 (0.218)	0.806 (0.034)
Unemployment	1959:Q1-2003:Q4	1.507 (0.217)	0.330 (0.079)	0.847 (0.028)
	1959:Q1-1979:Q2	0.880 (0.096)	0.217 (0.072)	0.710 (0.057)
	1979:Q3-2003:Q4	2.435 (0.250)	0.162 (0.078)	0.796 (0.034)

Notes: i) Inflations are quarterly changes in log CPI level ($\ln p_t - \ln p_{t-1}$). ii) Quadratically detrended gaps are used for real GDP output deviations. iii) Unemployment gaps are 5 year backward moving average unemployment rates minus current unemployment rates. iv) The set of instruments includes four lags of federal funds rate, inflation, output deviation, long-short interest rate spread, commodity price inflation, and M2 growth rate.

Table 2. GMM Median Unbiased Estimates and 95% Grid- t Confidence Intervals

Country	$\hat{\alpha}_{\text{GMM}}$	$\text{CI}_{\text{grid-}t}$	HL	HL $\text{CI}_{\text{grid-}t}$	J (pv)
Australia	0.884	[0.837,0.943]	1.404	[0.977,2.953]	5.532 (0.700)
Austria	0.804	[0.786,0.826]	0.793	[0.721,0.904]	8.173 (0.417)
Belgium	0.816	[0.794,0.844]	0.852	[0.751,1.019]	7.942 (0.439)
Canada	1.000	[0.967,1.000]	∞	[5.109, ∞)	4.230 (0.836)
Denmark	0.937	[0.874,1.000]	2.675	[1.290, ∞)	6.272 (0.617)
Finland	0.948	[0.897,1.000]	3.235	[1.587, ∞)	7.460 (0.488)
France	0.799	[0.777,0.822]	0.772	[0.688,0.885]	8.517 (0.385)
Germany	0.786	[0.767,0.809]	0.721	[0.652,0.819]	9.582 (0.296)
Italy	0.832	[0.806,0.864]	0.945	[0.805,1.181]	4.228 (0.836)
Japan	0.754	[0.729,0.782]	0.613	[0.549,0.706]	9.800 (0.279)
Netherlands	0.838	[0.798,0.883]	0.984	[0.766,1.388]	6.638 (0.576)
New Zealand	0.805	[0.786,0.828]	0.799	[0.718,0.918]	6.874 (0.550)
Norway	0.873	[0.785,0.971]	1.271	[0.716,5.983]	8.225 (0.412)
Portugal	0.792	[0.779,0.806]	0.741	[0.694,0.803]	6.132 (0.633)
Spain	0.896	[0.856,0.943]	1.581	[1.114,2.954]	6.738 (0.565)
Sweden	1.000	[0.945,1.000]	∞	[3.088, ∞)	7.107 (0.525)
Switzerland	0.831	[0.795,0.870]	0.937	[0.755,1.240]	9.136 (0.331)
UK	0.778	[0.756,0.806]	0.690	[0.620,0.801]	17.49 (0.025)
Median	0.832	[0.795,0.867]	0.941	[0.753,1.211]	-

Notes: i) The US\$ is the base currency. ii) Unemployment gaps are used for output deviations. iii) Sample periods are 1979.II-1998.IV (78 observations) for Eurozone countries and are 1979.II-2003.IV (98 observations) for non-Eurozone countries. iv) $\text{CI}_{\text{grid-}t}$ denotes the 95% confidence intervals that were obtained by 500 residual-based bootstrap replications on 30 grid points (Hansen, 1999). v) J denotes the J -statistic and pv is its associated p -values.

Table 3. Univariate Median Unbiased Estimates and Grid-t Confidence Intervals

Country	$\hat{\alpha}_{LS}$	CI_{grid-t}	HL	HL CI_{grid-t}
Australia	0.972	[0.891,1.000]	6.173	[1.494, ∞)
Austria	0.945	[0.866,1.000]	3.087	[1.205, ∞)
Belgium	0.924	[0.847,1.000]	2.203	[1.045, ∞)
Canada	1.000	[0.946,1.000]	∞	[3.122, ∞)
Denmark	0.942	[0.866,1.000]	2.886	[1.200, ∞)
Finland	0.959	[0.883,1.000]	4.107	[1.390, ∞)
France	0.931	[0.847,1.000]	2.432	[1.044, ∞)
Germany	0.950	[0.852,1.000]	3.349	[1.078, ∞)
Italy	0.943	[0.859,1.000]	2.932	[1.138, ∞)
Japan	0.952	[0.886,1.000]	3.511	[1.428, ∞)
Netherlands	0.936	[0.839,1.000]	2.619	[0.990, ∞)
New Zealand	0.959	[0.923,0.997]	4.089	[2.174,61.29]
Norway	0.934	[0.851,1.000]	2.529	[1.073, ∞)
Portugal	0.975	[0.913,1.000]	6.765	[1.904, ∞)
Spain	0.959	[0.898,1.000]	4.129	[1.604, ∞)
Sweden	0.959	[0.891,1.000]	4.089	[1.497, ∞)
Switzerland	0.951	[0.862,1.000]	3.481	[1.168, ∞)
UK	0.932	[0.845,1.000]	2.442	[1.028, ∞)
Median	0.951	[0.866,1.000]	3.415	[1.203, ∞)

Notes: i) The US\$ is the base currency. ii) Sample periods are 1979.II-1998.IV (78 observations) for Eurozone countries and are 1979.II-2003.IV (98 observations) for non-Eurozone countries. iii) CI_{grid-t} denotes the 95% confidence intervals that were obtained by 500 residual-based bootstrap replications on 30 grid points (Hansen, 1999).

Table 4. Univariate Median Unbiased Half-Life Estimates: AR(1) vs. AR(p)

Country	p_{MAIC}	p_{MBIC}	$HL_{AR(1)}$	HL_{IRF}
Australia	1	1	6.173	6.173
Austria	1	1	3.087	3.087
Belgium	4	1	2.203	2.884
Canada	6	1	∞	∞
Denmark	4	1	2.886	3.883
Finland	6	2	4.107	3.631
France	1	1	2.432	2.432
Germany	6	1	3.349	3.386
Italy	3	1	2.932	∞
Japan	1	1	3.511	3.511
Netherlands	6	1	2.619	2.882
New Zealand	9	1	4.089	3.895
Norway	1	1	2.529	2.529
Portugal	6	1	6.765	∞
Spain	2	1	4.129	12.13
Sweden	4	4	4.089	3.387
Switzerland	1	1	3.481	3.481
UK	3	1	2.442	3.129
Median	3.5	1	3.415	3.496

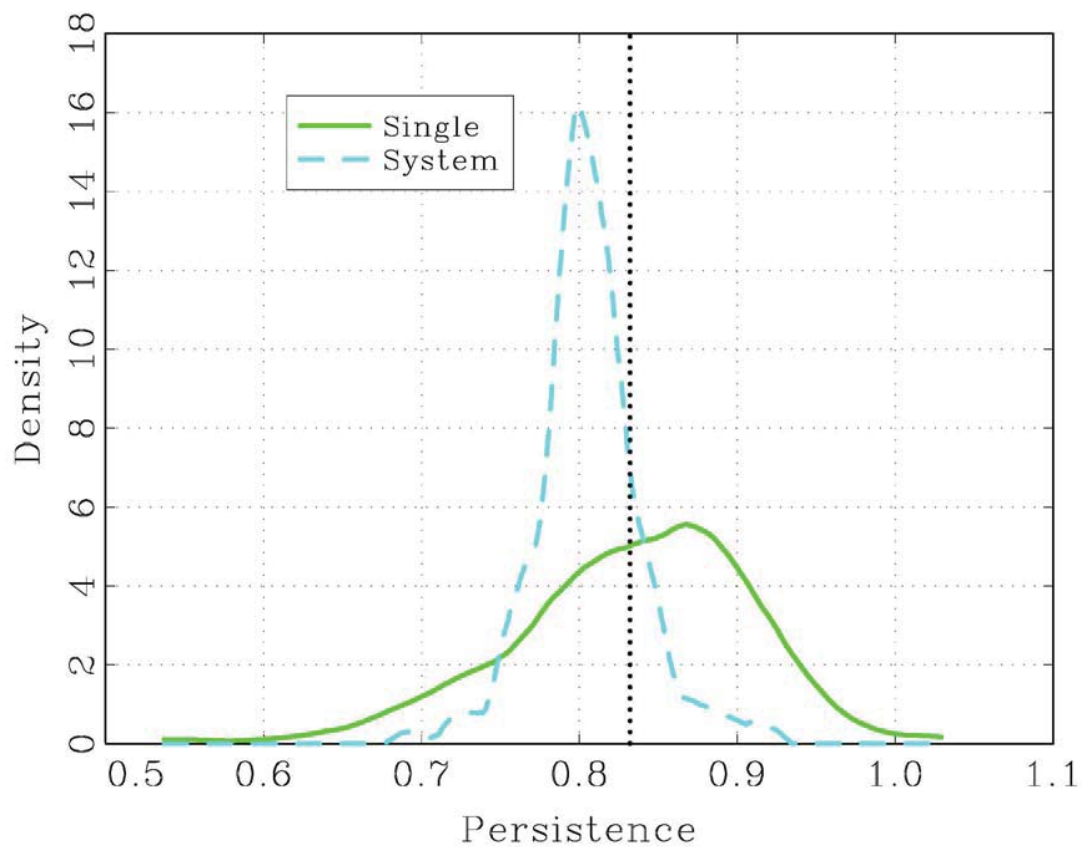
Notes: i) p_{MAIC} and p_{MBIC} denote the lag length chosen by the modified Akaike Information criteria and the modified Bayesian Information criteria (Ng and Perron, 2001) with maximum 12 lags, respectively. ii) $HL_{AR(1)}$ refers the half-life point estimates with an AR(1) specification and was replicated from Table 3 for a comparison purpose. iii) HL_{IRF} denotes the half-life point estimates obtained from the impulse-response function with the lag length chosen by p_{MAIC} . HL_{IRF} with p_{MBIC} is not reported because the estimates are virtually the same as $HL_{AR(1)}$. iv) We correct the median bias of each autoregressive coefficient for higher order AR(p) conditioning on all other coefficients.

Table 5. Statistics of the Persistence Parameter Estimates from Simulated Data

T	UIP Shock	Distribution	Estimator	Mean	Std. Dev	Min	Median	Max
78	$\sigma_v^2 = 0$	Normal	Single	0.832	0.080	0.511	0.844	1.004
			System	0.793	0.041	0.529	0.797	0.943
		t_3	Single	0.833	0.076	0.537	0.842	1.030
			System	0.805	0.033	0.691	0.804	0.920
78	$\sigma_v^2 = \sigma_u^2$	Normal	Single	0.829	0.081	0.459	0.841	1.005
			System	0.800	0.050	0.540	0.804	0.940
		t_3	Single	0.831	0.077	0.470	0.837	1.011
			System	0.805	0.044	0.629	0.809	0.927
78	$\sigma_v^2 = 5\sigma_u^2$	Normal	Single	0.829	0.079	0.526	0.841	1.006
			System	0.809	0.056	0.569	0.817	0.970
		t_3	Single	0.827	0.079	0.521	0.838	1.009
			System	0.809	0.048	0.640	0.809	0.957
500	$\sigma_v^2 = 0$	Normal	Single	0.833	0.026	0.751	0.835	0.892
			System	0.814	0.021	0.749	0.812	0.894
		t_3	Single	0.833	0.026	0.751	0.833	0.952
			System	0.823	0.019	0.756	0.822	0.888
500	$\sigma_v^2 = \sigma_u^2$	Normal	Single	0.831	0.026	0.745	0.832	0.891
			System	0.819	0.023	0.765	0.815	0.894
		t_3	Single	0.832	0.025	0.719	0.833	0.935
			System	0.824	0.020	0.779	0.823	0.884
500	$\sigma_v^2 = 5\sigma_u^2$	Normal	Single	0.830	0.026	0.727	0.832	0.894
			System	0.831	0.026	0.742	0.833	0.895
		t_3	Single	0.831	0.025	0.715	0.832	0.915
			System	0.834	0.023	0.753	0.834	0.885

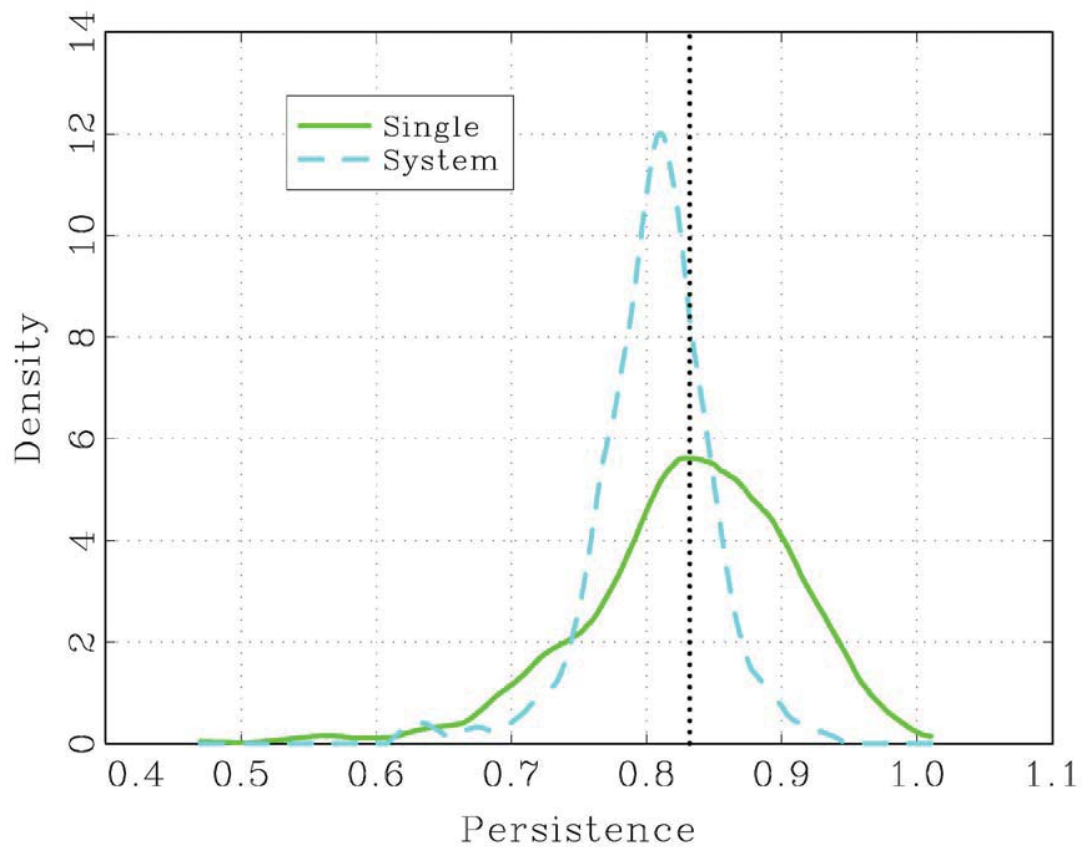
Notes: i) We obtained these summary statistics from 500 simulated samples. ii) T is the number of observations. We set T be 78 from our Eurozone data to see small sample properties, while $T = 500$ for large samples. iii) We studied three possible values for the size of the UIP shock, $\sigma_v^2 = 0$, $\sigma_v^2 = \sigma_u^2$, $\sigma_v^2 = 5\sigma_u^2$, where $\sigma_u^2 = 0.33^2$ is the calibrated variance of the monetary policy shock. iv) Normal and t_3 are the standard normal distribution and the t distribution with the 3 degrees of freedom, respectively, for the underlying distribution of structural shocks. Standard deviations are scaled to match each of calibrated variance of shocks.

Figure 1. Kernel Density Estimate of the Persistence Parameter: No UIP Shock ($\sigma_v^2 = 0$) and $T = 78$



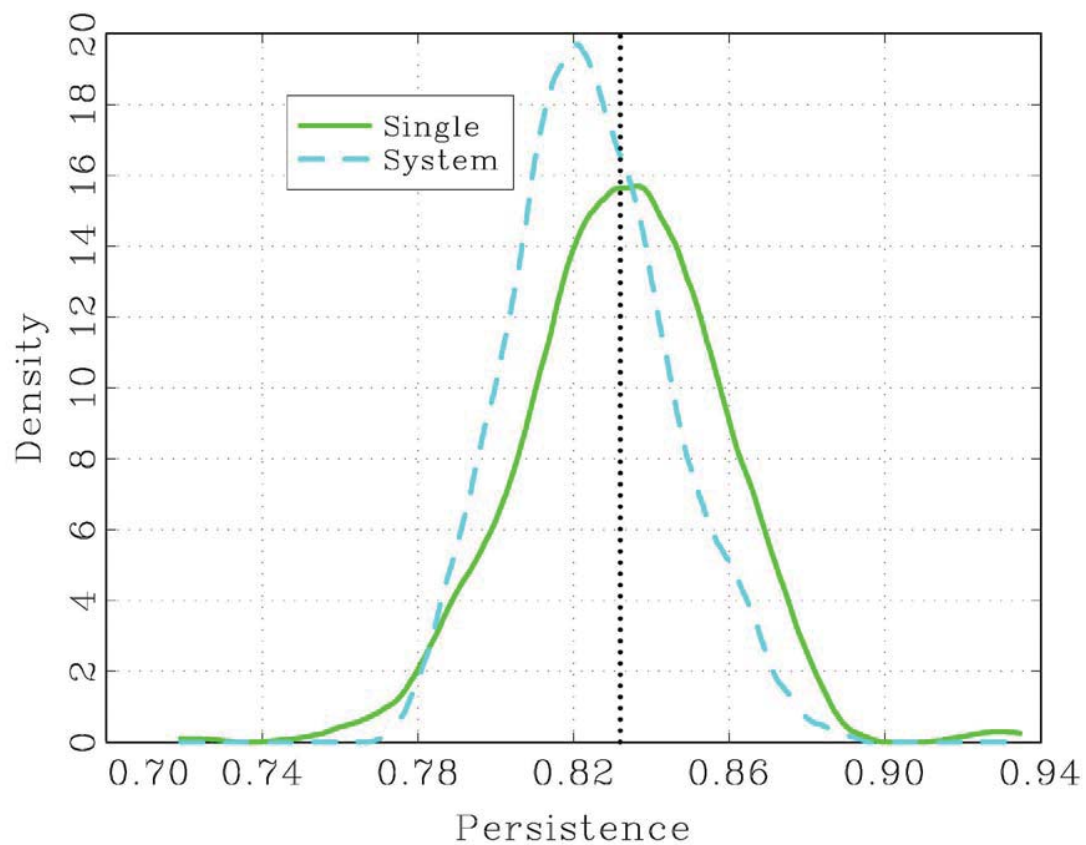
Notes: i) The dotted line is the true value of the persistence parameter, 0.832. ii) The Epanechnikov kernel is used.

Figure 2. Kernel Density Estimate of the Persistence Parameter: With UIP Shock
($\sigma_v^2 = \sigma_u^2$) and $T = 78$



Notes: i) The dotted line is the true value of the persistence parameter, 0.832. ii) The Epanechnikov kernel is used.

Figure 3. Kernel Density Estimate of the Persistence Parameter: With UIP Shock
($\sigma_v^2 = \sigma_u^2$) and $T = 500$



Notes: i) The dotted line is the true value of the persistence parameter, 0.832. ii) The Epanechnikov kernel is used.