PURCHASING POWER PARITY AND THE TAYLOR RULE

HYEONGWOO KIM,[a]* IPPEI FUJIWARA,[b] BRUCE E. HANSEN[c] AND MASAO OGAKI[d]

[a] *Department of Economics, Auburn University, AL, USA*
[b] *Crawford School of Public Policy and Centre for Applied Macroeconomic Analysis, Australian National University, Canberra, ACT, Australia*
[c] *Department of Economics, University of Wisconsin, Madison, WI, USA*
[d] *Department of Economics, Keio University, Tokyo, Japan*

SUMMARY

It is well known that there is a large degree of uncertainty around Rogoff's 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 non-parametric 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 3–5 years, with wide confidence intervals that extend to positive infinity. In contrast, the system method yields median-unbiased estimates that are typically shorter than 1 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. Copyright © 2014 John Wiley & Sons, Ltd.

📖 *All supporting information may be found in the online version of this article.*

1. INTRODUCTION

Reviewing the literature on purchasing power parity (PPP), which 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.

---

* Correspondence to: Hyeongwoo Kim, Department of Economics, Auburn University, 0339 Haley Center., Auburn, AL 36849, USA. E-mail: gmmkim@gmail.com

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

[2] Imbs *et al.* (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 *et al.* (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.

In this paper, we ask a different question: should we take Rogoff's remarkable consensus of 3- to 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; Molodtsova et al., 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 non-parametric bootstrap method to construct confidence intervals. For this purpose, we propose the grid bootstrap method for our generalized method of moments (GMM) 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 3–5 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 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 et al. (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

---

[3] Kehoe and Midrigan (2007) and Crucini et al. (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 et al. (2005) and Smets and Wouters (2007), among others, may be more efficient.

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.

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^* the log-foreign price level and e_t the log-nominal exchange rate as the price of one unit of the foreign currency in terms of the home currency. We denote by s_t 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 $\left[p_t \ p_t^* \ e_t \right]'$, 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:

$$s_{t+1} = d + \alpha s_t + \varepsilon_{t+1} \tag{1}$$

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 misspecified 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, equation (1) can be written as the following error correction model of real exchange rates with the cointegrating relation described earlier:

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

where $\mu = \mathbb{E} \left(p_t - p_t^* - e_t \right)$, $b = 1 - \alpha$, $d = -(1-\alpha)\mu$, $\varepsilon_{t+1} = \varepsilon_{1,t+1} + \varepsilon_{2,t+1} - \varepsilon_{3,t+1} = (e_{t+1} - \mathbb{E}_t e_{t+1}) + \left(p_{t+1}^* - \mathbb{E}_t p_{t+1}^* \right) - (p_{t+1} - \mathbb{E}_t p_{t+1})$, 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.

2.2. The Taylor Rule Model

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

$$\mathbb{E}_t \Delta e_{t+1} = i_t - i_t^* \tag{3}$$

---

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

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 $\left(i_t^T\right)$ 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 rate,[7] and γ_π and γ_x are the *long-run* Taylor rule coefficients on expected future inflation[8] $(\mathbb{E}_t \Delta p_{t+1})$ and current output deviations[9] (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:

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

Combining equations (3) and (4), we obtain the following:

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

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.

Now, let us rewrite equation (2) as the following equation in level variables:

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

Taking differences and rearranging, equation (2') can be rewritten as follows:

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

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

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

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

[5] The UIP often fails to hold when one tests it by estimating a single regression equation, $\Delta e_{t+1} = \beta \left(i_t - i_t^*\right) + \varepsilon_{t+1}$. 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.

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

[7] See Clarida *et al.* (1998, 2000) for details.

[8] It 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.

[9] If 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.

For notational simplicity, let us rewrite equation (7) in matrix form as follows:

$$\mathbf{A}\mathbb{E}_t\mathbf{y}_{t+1} = \mathbf{B}\mathbf{y}_t + \mathbf{x}_t \tag{7'}$$

and thus

$$\mathbb{E}_t\mathbf{y}_{t+1} = \mathbf{A}^{-1}\mathbf{B}\mathbf{y}_t + \mathbf{A}^{-1}\mathbf{x}_t \tag{8}$$
$$= \mathbf{D}\mathbf{y}_t + \mathbf{c}_t$$

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

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

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 equation (9) by $\mathbf{V}^{-1}$ and redefining variables:

$$\mathbb{E}_t\mathbf{z}_{t+1} = \mathbf{\Lambda}\mathbf{z}_t + \mathbf{h}_t \tag{10}$$

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 $\frac{\rho}{1-\gamma_\pi^s}\left(= \frac{\rho}{1-(1-\rho)\gamma_\pi}\right)$ is greater than unity as long as $1 < \gamma_\pi < \frac{1}{1-\rho}$. Therefore, if the long-run inflation 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 equation (7) satisfies the following relation (see Appendix A for the derivation):

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

where

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

---

[10] It is straightforward to show that $\mathbf{A}$ is non-singular and thus has a well-defined inverse.

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

$$f_t = -\left[i_t^* - \mathbb{E}_t \Delta p_{t+1}^*\right] + \frac{\gamma_x^s}{\gamma_\pi^s} x_t$$

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

and

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

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

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

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 equation (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 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. We believe, therefore, 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, 2000).

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.

---

[12]   If the system has a purely backward-looking solution, the conventional structural vector autoregressive (SVAR) estimation method may apply.

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 *et al.*, 2007).

3.1. Univariate Equation Approach

A univariate approach utilizes equation (1) or (2). For instance, the persistence parameter α in equation (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 equation (2) as follows:

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

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

3.2. GMM System Method

Our second estimation strategy combines equation (11) with (1). The estimation of 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 equation (11) as follows:

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

where

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

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 $\Omega_t = \{f_t, f_{t-1}, f_{t-2}, \ldots\}$. 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:

$$f_t = c(L)\nu_t \tag{13}$$

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

$$b(L) f_t = \nu_t \tag{14}$$

where $b(L) = c^{-1}(L) = 1 - b_1 L - b_2 L^2 - \ldots$. 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:

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

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, i.e. $b(L) = 1 - \sum_{j=1}^{r} b_j L^j$, where $r < \infty$.[13] It can then be shown that the coefficients of $\psi(L)$ can be computed recursively (see Sargent, 1987) as follows:

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

$$\psi_r = 0$$

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

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

$$\Delta e_{t+1} = \hat{\imath} + \frac{\alpha \gamma_\pi^s}{\alpha - \rho} \Delta p_{t+1} - \frac{\alpha \gamma_\pi^s}{\alpha - \rho} \Delta p_{t+1}^* + \frac{\alpha \gamma_\pi^s - (\alpha - \rho)}{\alpha - \rho} i_t^*$$
$$+ \frac{\gamma_\pi^s \left(\alpha \gamma_\pi^s - (\alpha - \rho)\right)}{(\alpha - \rho)\rho} (\psi_0 f_t + \psi_1 f_{t-1} + \ldots + \psi_{r-1} f_{t-r+1}) + \xi_{t+1} \tag{16}$$

$$f_{t+1} = k + b_1 f_t + b_2 f_{t-1} + \ldots + b_r f_{t-r+1} + \nu_{t+1} \tag{17}$$

---

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

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

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

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

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

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

where $\mathbf{x}_{1,t} = (1\ s_t)'$, $\mathbf{x}_{2,t} = (1\ f_t)'$ and $\tau = 0, 1, \ldots, p$.[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 that 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 equations (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, equations (2") and (16) can be rewritten as follows:

$$\Delta\tilde{p}_{t+1} = b\mu - b\left(\tilde{p}_t - e_t\right) + \Delta e_{t+1} + \tilde{\varepsilon}_{t+1}$$

$$\Delta e_{t+1} = \hat{\imath} + \frac{\alpha\gamma_\pi^s}{\alpha-\rho}\Delta\tilde{p}_{t+1} + \frac{\alpha\gamma_\pi^s - (\alpha-\rho)}{\alpha-\rho}i_t^*$$
$$+ \frac{\gamma_\pi^s(\alpha\gamma_\pi^s - (\alpha-\rho))}{(\alpha-\rho)\rho}\left(\psi_0 f_t + \ldots + \psi_{r-1}f_{t-r+1}\right) + \xi_{t+1}$$

Or, in matrix form:

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

---

[14] Recall 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.

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

[16] p does not necessarily coincide with r.

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

where $\mathbf{C}$ is a vector of constants and $\mathbf{S}$ is $\left[1 \ -1 \ \vdots \ -\frac{\alpha \gamma_\pi^s}{\alpha - \rho} \ 1\right]$.

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 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:

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

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

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

where $\mathbf{N}_c$, $\mathbf{\Omega}_c$, and $\mathbf{G}_c$ are defined in (C4), (C5) and (C6).

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

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

and the median unbiased estimator as

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

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 dollar 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 *et al.* 1998) and unemployment rate gaps (see

---

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

[19] See Hansen (1999) for detailed explanations.

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

[21] Among 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.

Table I. GMM estimation of the US Taylor rule estimation

Deviation	Sample period	γ_π (SE)	γ_x (SE)	ρ (SE)
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) Inflation is the quarterly change 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.

Boivin, 2006).[22,23] The data frequency is quarterly and from the IFS CD-ROM. The sample period is from 1979:Q3 to 1998:Q4 for Eurozone countries, and from 1979:Q3 to 2003:Q4 for the rest of the countries.

Based on the empirical evidence of the US Taylor rule, our sample period starts from 1979:Q3. As discussed in Section 2, the inflation and exchange rate dynamics may greatly depend on the 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 *et al.* (1998, 2000) provide important empirical evidence for the existence of a structural break in the US Taylor rule. Put 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 equation (4) as in Clarida *et al.* (2000),[24,25] with a longer sample period and report the results in Table I (see the note to Table I for a detailed explanation). We use two output gap measures for three different subsamples. 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.

We report our GMM version median unbiased estimates and the 95% grid-t confidence intervals in Table II. 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 III.

---

[22] We 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.
[23] The 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.
[24] They used GDP deflator inflation along with the CBO output gaps (and HP detrended gaps).
[25] Unlike Clarida *et al.* (2000), we assume that the Fed targets current output gap rather than future deviations. However, this does not make any significant changes to our results. Also, we include one lag of interest rate rather than two lags for simplicity.
[26] J-test statistics are available upon request.
[27] The results with quadratically detrended real GDP gaps were quantitatively similar.

Table II. 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 dollar is the base currency.
(ii) Unemployment gaps are used for output deviations.
(iii) Sample periods are 1979:Q2–1998:Q4 (78 observations) for Eurozone countries and 1979:Q2–2003:Q4 (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 III. Univariate median unbiased estimates and grid-t confidence Intervals

Country	$\hat{\alpha}_{\text{LS}}$	$\text{CI}_{\text{grid-}t}$	HL	HL $\text{CI}_{\text{grid-}t}$
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 dollar is the base currency.
(ii) Sample periods are 1979:Q2–1998:Q4 (78 observations) for Eurozone countries and are 1979:Q2–2003:Q4 (98 observations) for non-Eurozone countries.
(iii) $\text{CI}_{\text{grid-}t}$ denotes the 95% confidence intervals that were obtained by 500 residual-based bootstrap replications on 30 grid points (Hansen, 1999).

Table IV. 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 AIC and modified BIC (Ng and Perron, 2001) with maximum 12 lags, respectively.
(ii) $HL_{AR(1)}$ refers to the half-life point estimates with an AR(1) specification and was replicated from Table III for comparison purposes.
(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.

We note that the system method provides much shorter half-life estimates compared with those resulting from the single-equation method (see Tables II and III). 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 corrected 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 Papell (2002) for most countries. [30] J-test accepts our model specification for all countries, with the exception of the UK.[31]

---

[28] Without 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.
[29] For 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.
[30] The 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.
[31] We also notice that our median-unbiased point estimate $\hat{\alpha}_{GMM,MUE}$ 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) implies that α coincides with the Calvo probability parameter.
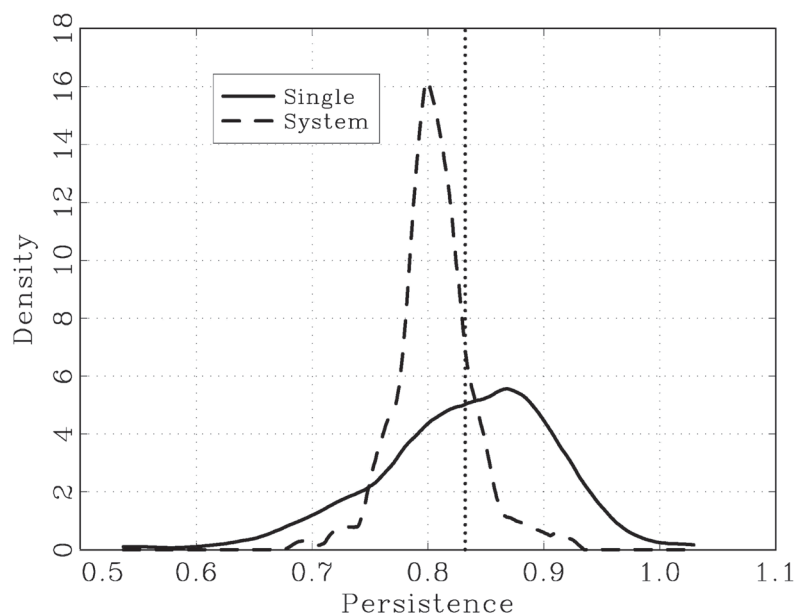
Notes: i) The dotted line is the true value of the persistence parameter, 0.832. ii) The Epanechnikov kernel is used.
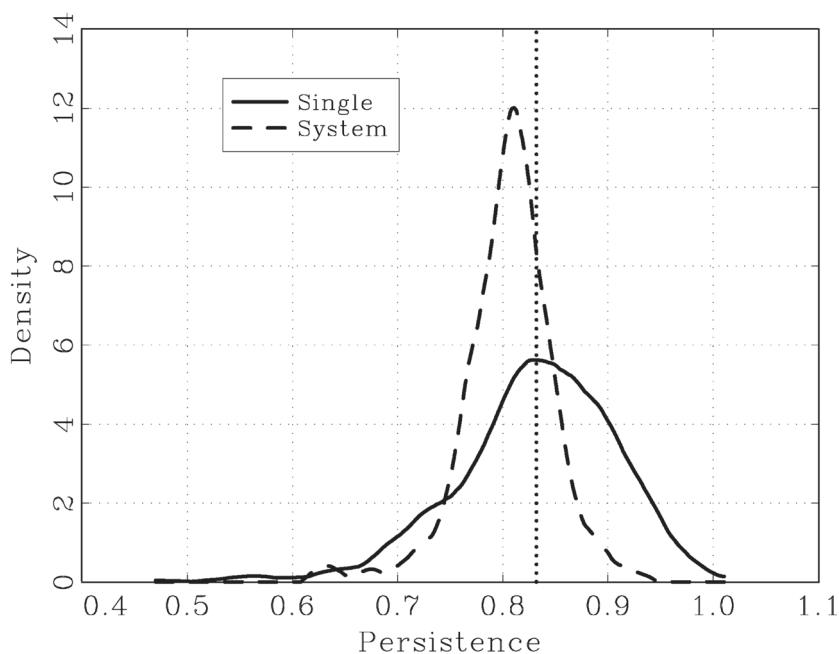
Figure 1.   Kernel density estimate of the persistence parameter: no UIP shock $\left(\sigma_v^2 = 0\right)$ and $T = 78$.

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 III). 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 a great improvement over Kim *et al.* (2007), who acquired limited success in efficiency gains.

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 criterion (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 IV, 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

---

[32]   Their confidence intervals are constructed following Andrews (1993) and Andrews and Chen (1994), which are identical to 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.

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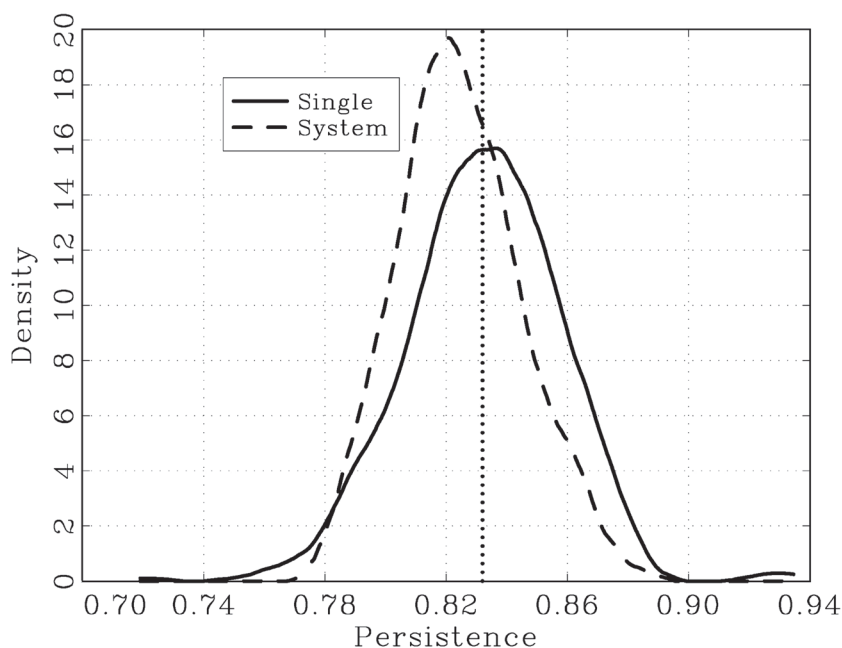0.832. ii) The Epanechnikov kernel is used.

Figure 2. Kernel density estimate of the persistence parameter: with UIP shock $\left(\sigma_v^2 = \sigma_u^2\right)$ and $T = 78$.

rates, however, half-life estimates do not change much, implying that the AR(1) process is not a bad approximation.

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 $\left(\sigma_v^2\right)$ relative to that of the monetary shock $\left(\sigma_u^2\right)$, i.e. $\sigma_v^2 = 0, \sigma_v^2 = \sigma_u^2$, and $\sigma_v^2 = 5\sigma_v^2$. 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, the more severe is the misspecification of the system method. 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 rescaled so that they match with calibrated variances.

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 $\left(\sigma_v^2 = \sigma_u^2\right)$ and $T = 500$.

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 Figures 1–3. We also report various statistical properties of singe and system estimators in Table V. 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 V), 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 $\sigma_v^2 = \sigma_v^2$, 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.

---

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

Table V. Statistics of the persistence parameter estimates from simulated data

T	UIP shock	Distribution	Estimator	Mean	SD	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 3 degrees of freedom, respectively, for the underlying distribution of structural shocks. Standard deviations are scaled to match each of calibrated variance of shocks.

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 USA 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 non-parametric 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.

---

[34] Results from standard bootstrap are available upon request.

When we use the grid bootstrap method, most of the (approximately) median unbiased estimates from the single-equation method fall in the range of 3–5 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 1 year, with much sharper confidence intervals, most of which range from three 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 1 year, which is far more reasonable than the consensus 3–5 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 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 modeled. 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.

REFERENCES

Andrews DWK. 1993. Exactly median-unbiased estimation of first order autoregressive/unit root models. *Econometrica* **61**: 139–165.

Andrews DWK, Chen HY. 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.

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, Weinstein DE. 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, Nechio F. 2011. Aggregation and the PPP puzzle in a sticky-price model. *American Economic Review* **101**: 2391–2424.

---

[35] It is also interesting to see that our half-life estimates imply about four to six 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 US data for the NKPC and DSGE models.

Chen SS, Engel C. 2005. Does 'aggregation bias' explain the PPP puzzle? *Pacific Economic Review* **10**: 49–72.

Christiano LJ, Eichenbaum M, Evans CL. 2005. Nominal rigidities and the dynamic effects of a shock to monetary policy. *Journal of Political Economy* **113**: 1–45.

Clarida R, Galí J, Gertler M. 1998. Monetary policy rules in practice: some international evidence. *European Economic Review* **42**: 1033–1067.

Clarida R, Galí J, Gertler M. 2000. Monetary policy rules and macroeconomic stability: evidence and some theory. *Quarterly Journal of Economics* **115**: 147–180.

Clarida R, Waldman D. 2007. Is bad news about inflation good news for the exchange rate? *NBER Working Paper No 13010*.

Crucini MJ, Shintani M. 2008. Persistence in law-of-one-price deviations: evidence from micro-data. *Journal of Monetary Economics* **55**: 629–644.

Crucini MJ, Shintani M, Tsuruga T. 2013. Do Sticky Prices Increase Real Exchange Rate Volatility at the Sector Level? *European Economic Review* **62**: 58–72.

Efron B, Tibshirani RJ. 1993. *An Introduction to the Bootstrap*. Chapman and Hall/CRC: London, UK.

Engel C, West KD. 2005. Exchange rates and fundamentals. *Journal of Political Economy* **113**: 485–517.

Engel C, West KD. 2006. Taylor rules and the deutschmark-dollar exchange rate. *Journal of Money Credit, and Banking* **38**: 1175–1194.

Galí J, Gertler M. 1999. Inflation dynamics: a structural econometric analysis. *Journal of Monetary Economics* **44**: 195–222.

Hansen BE. 1995. Rethinking the univariate approach to unit root testing. *Econometric Theory* **11**: 1148–1171.

Hansen BE. 1999. The grid bootstrap and the autoregressive model. *Review of Economics and Statistics* **81**: 594–607.

Hansen LP, Sargent TJ. 1980. Formulating and estimating dynamic linear rational expectations models. *Journal of Economic Dynamics and Control* **2**: 7–46.

Hansen LP, Sargent TJ. 1982. Instrumental variables procedures for estimating linear rational expectations models. *Journal of Monetary Economics* **9**: 263–296.

Imbs J, Mumtaz H, Ravn MO, Rey H. 2005. PPP strikes back: aggregation and the real exchange rates. *Quarterly Journal of Economics* **120**: 1–43.

Kehoe PJ, Midrigan V. 2007. Stikcy prices and sectoral real exchange Rates. *Working Paper No 656*, Federal Reserve Bank of Minneapolis.

Kilian L, Zha T. 2002. Quantifying the uncertainty about the half-life of deviations from PPP. *Journal of Applied Econometrics* **17**: 107–125.

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, Ogaki M. 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, Ogaki M, Yang MS. 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, Mavroeidis S. 2009. Weak instrument robust tests in GMM and the new Keynesian Phillips curve. *Journal of Business and Economic Statistics* **27**: 293–311.

Mark NC. 2009. Changing monetary policy rules, learning, and real exchange rate dynamics. *Journal of Money Credit and Banking* **41**: 1047–1070.

Molodtsova T, Nikolsko-Rzhevskyy A, Papell DH. 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, Papell DH. 2009. Out-of-sample exchange rate predictability with Taylor rule fundamentals. *Journal of International Economics* **77**: 167–180.

Murray CJ, Papell DH. 2002. The purchasing power parity persistence paradigm. *Journal of International Economics* **56**: 1–19.

Mussa M. 1982. A model of exchange rate dynamics. *Journal of Political Economy* **90**: 74–104.

Ng S, Perron P. 2001. Lag length selection and the construction of unit root tests with good size and power. *Econometrica* **69**: 1519–1554.

Parsley DC, Wei SJ. 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 TJ. 1987. *Macroeconomic Theory*. Academic Press: New York.

Smets F, Wouters R. 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.

Svensson LE. 2000. Open-economy inflation targeting. *Journal of International Economics* **50**: 155–183.

West KD. 1989. Estimation of linear rational expectations models in the presence of deterministic terms. *Journal of Monetary Economics* **24**: 437–442.

APPENDIX A: DERIVATION OF EQUATION (11)

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

$$z_{1,t} = \sum_{j=0}^{\infty} \alpha^j h_{1,t-j-1} + \sum_{j=0}^{\infty} \alpha^j u_{t-j} \tag{A1}$$

$$z_{2,t} = -\sum_{j=0}^{\infty} \left(\frac{1-\gamma_\pi^s}{\rho} \right)^{j+1} \mathbb{E}_t h_{2,t+j} \tag{A2}$$

$$z_{3,t} = h_{3,t-1} + \upsilon_t \tag{A3}$$

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

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

$$\begin{bmatrix} \Delta p_t \\ \Delta e_t \\ i_{t-1} \end{bmatrix} = \begin{bmatrix} 1 & 1 & 1 \\ \frac{\alpha \gamma_\pi^s}{\alpha - \rho} & 1 & 1 \\ \frac{\alpha \gamma_\pi^s}{\alpha - \rho} & 1 & 0 \end{bmatrix} \begin{bmatrix} z_{1,t} \\ z_{2,t} \\ z_{3,t} \end{bmatrix} \tag{A4}$$

From first and second rows of equation (A4), we get the following:

$$\Delta e_t = \frac{\alpha \gamma_\pi^s}{\alpha - \rho} \Delta p_t - \frac{\alpha \gamma_\pi^s - (\alpha - \rho)}{\alpha - \rho} z_{2,t} - \frac{\alpha \gamma_\pi^s - (\alpha - \rho)}{\alpha - \rho} z_{3,t} \tag{A5}$$

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

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

and thus

$$h_{1,t} = -\frac{\alpha-\rho}{\alpha\gamma_\pi^s-(\alpha-\rho)} \left(\mathbb{E}_t \Delta p_{t+1}^* - \alpha \Delta p_t^* + \eta_t \right) \tag{A6}$$

$$h_{2,t} = \frac{1}{1-\gamma_\pi^s} \left[\frac{\rho\gamma_\pi^s}{\alpha\gamma_\pi^s-(\alpha-\rho)} \left(\mathbb{E}_t \Delta p_{t+1}^* - \alpha \Delta p_t^* + \eta_t \right) + \iota + \gamma_x^s x_t - \gamma_\pi^s i_t^* \right] \tag{A7}$$

$$h_{3,t} = -i_t^* \tag{A8}$$

Plugging equation (A6) into equation (A1):

$$
\begin{aligned}
z_{1,t} &= -\frac{\alpha - \rho}{\alpha \gamma_\pi^s - (\alpha - \rho)} \sum_{j=0}^{\infty} \alpha^j \left(\Delta p_{t-j}^* - \alpha \Delta p_{t-j-1}^* + \eta_{t-j-1} \right) + \sum_{j=0}^{\infty} \alpha^j u_{t-j} \\
&= -\frac{\alpha - \rho}{\alpha \gamma_\pi^s - (\alpha - \rho)} \Delta p_t^* + \sum_{j=0}^{\infty} \alpha^j u_{t-j} - \frac{\alpha - \rho}{\alpha \gamma_\pi^s - (\alpha - \rho)} \sum_{j=0}^{\infty} \alpha^j \eta_{t-j-1}
\end{aligned}
\tag{A9}
$$

Plugging equation (A7) into equation (A2):[36]

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

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

$$
z_{2,t} = \frac{\alpha \gamma_\pi^s}{\alpha \gamma_\pi^s - (\alpha - \rho)} \Delta p_t^* - \frac{\gamma_\pi^s}{\alpha \gamma_\pi^s - (\alpha - \rho)} \eta_t - \frac{\iota}{\gamma_\pi^s - (1 - \rho)} - \frac{\gamma_\pi^s}{\rho} \sum_{j=0}^{\infty} \left(\frac{1 - \gamma_\pi^s}{\rho} \right)^j E_t f_{t+j} \tag{A10}
$$

Finally, plugging equation (A8) into equation (A3):

$$z_{3,t} = -i_{t-1}^* + \upsilon_t \tag{A11}$$

Now, plugging equations (A10) and (A11) into equation (A5):

$$
\begin{aligned}
\Delta e_t &= \frac{\alpha \gamma_\pi^s}{\alpha - \rho} \Delta p_t - \frac{\alpha \gamma_\pi^s}{\alpha - \rho} \Delta p_t^* + \frac{\gamma_\pi^s}{\alpha - \rho} \eta_t + \frac{\alpha \gamma_\pi^s - (\alpha - \rho)}{(\alpha - \rho)\left(\gamma_\pi^s - (1 - \rho) \right)} \iota \\
&\quad + \frac{\gamma_\pi^s \left(\alpha \gamma_\pi^s - (\alpha - \rho) \right)}{(\alpha - \rho)\rho} \sum_{j=0}^{\infty} \left(\frac{1 - \gamma_\pi^s}{\rho} \right)^j \mathbb{E}_t f_{t+j} + \frac{\alpha \gamma_\pi^s - (\alpha - \rho)}{\alpha - \rho} i_{t-1}^* - \frac{\alpha \gamma_\pi^s - (\alpha - \rho)}{\alpha - \rho} \upsilon_t
\end{aligned}
\tag{A12}
$$

Updating equation (A12) once and applying the law of iterated expectations:

$$
\begin{aligned}
\Delta e_{t+1} &= \hat{\iota} + \frac{\alpha \gamma_\pi^s}{\alpha - \rho} \Delta p_{t+1} - \frac{\alpha \gamma_\pi^s}{\alpha - \rho} \Delta p_{t+1}^* + \frac{\alpha \gamma_\pi^s - (\alpha - \rho)}{\alpha - \rho} i_t^* \\
&\quad + \frac{\gamma_\pi^s (\alpha \gamma_\pi^s - (\alpha - \rho))}{(\alpha - \rho)\rho} \sum_{j=0}^{\infty} \left(\frac{1 - \gamma_\pi^s}{\rho} \right)^j \mathbb{E}_t f_{t+j+1} + \omega_{t+1}
\end{aligned}
\tag{A13}
$$

---

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

where

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

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

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

APPENDIX 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:

$$z_{1,t} = \sum_{j=0}^{\infty}\rho^j h_{1,t-j-1} + \sum_{j=0}^{\infty}\rho^j u_{t-j} \quad \text{(B4)}$$

$$z_{2,t} = -\sum_{j=0}^{\infty}\left(\frac{1 - \gamma_\pi^s}{\rho}\right)^{j+1}\mathbb{E}_t h_{2,t+j} \quad \text{(B5)}$$

$$z_{3,t} = h_{3,t-1} + \upsilon_t \tag{B6}$$

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

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

$$\begin{bmatrix} \Delta p_t \\ \Delta e_t \\ i_{t-1} \end{bmatrix} = \begin{bmatrix} 0 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 0 \end{bmatrix} \begin{bmatrix} z_{1,t} \\ z_{2,t} \\ z_{3,t} \end{bmatrix} \tag{B7}$$

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

$$\mathbf{h}_t = \begin{bmatrix} -1 & 1 & 0 \\ 1 & -1 & 1 \\ 0 & 1 & -1 \end{bmatrix} \begin{bmatrix} \left(\mathbb{E}_t \Delta p^*_{t+1} - \rho \Delta p^*_t + \eta_t \right) + \iota + \gamma^s_x x_t - i^*_t \\ \gamma^s_\pi (\mathbb{E}_t \Delta p^*_{t+1} - \rho \Delta p^*_t + \eta_t) + \iota + \gamma^s_x x_t - i^*_t \\ \gamma^s_\pi \left(\mathbb{E}_t \Delta p^*_{t+1} - \rho \Delta p^*_t + \eta_t \right) + \iota + \gamma^s_x x_t - \gamma^s_\pi i^*_t \end{bmatrix}$$

thus:

$$h_{1,t} = -\left(1 - \gamma^s_\pi\right)\left(\mathbb{E}_t \Delta p^*_{t+1} - \rho \Delta p^*_t + \eta_t\right) \tag{B8}$$

$$h_{2,t} = \mathbb{E}_t \Delta p^*_{t+1} - \rho \Delta p^*_t + \eta_t + \iota + \gamma^s_x x_t - \gamma^s_\pi i^*_t \tag{B9}$$

$$h_{3,t} = -\left(1 - \gamma^s_\pi\right) i^*_t \tag{B10}$$

From equations (B4) and (B8):

$$\begin{aligned} z_{1,t} &= -\left(1 - \gamma^s_\pi\right) \sum_{j=0}^{\infty} \rho^j \left(\Delta p^*_{t-j} - \rho \Delta p^*_{t-j-1} + \eta_{t-j-1}\right) + \sum_{j=0}^{\infty} \rho^j u_{t-j} \\ &= -\left(1 - \gamma^s_\pi\right) \Delta p^*_t + \sum_{j=0}^{\infty} \rho^j u_{t-j} - \left(1 - \gamma^s_\pi\right) \sum_{j=0}^{\infty} \rho^j \eta_{t-j-1} \end{aligned} \tag{B11}$$

From equations (B5) and (B9):

$$\begin{aligned} z_{2,t} &= -\sum_{j=0}^{\infty} \left(\frac{1 - \gamma^s_\pi}{\rho}\right)^{j+1} \left(\mathbb{E}_t \Delta p^*_{t+j+1} - \rho \mathbb{E}_t \Delta p^*_{t+j} + \mathbb{E}_t \eta_{t+j} + \iota + \gamma^s_x \mathbb{E}_t x_{t+j} - \gamma^s_\pi \mathbb{E}_t i^*_{t+j}\right) \\ &= \left(1 - \gamma^s_\pi\right) \Delta p^*_t - \left(\frac{1 - \gamma^s_\pi}{\rho}\right) \eta_t - \frac{\left(1 - \gamma^s_\pi\right) \iota}{\rho - \left(1 - \gamma^s_\pi\right)} \\ &\quad - \gamma^s_\pi \sum_{j=0}^{\infty} \left(\frac{1 - \gamma^s_\pi}{\rho}\right)^{j+1} \left(\mathbb{E}_t \Delta p^*_{t+j+1} + \frac{\gamma^s_x}{\gamma^s_\pi} \mathbb{E}_t x_{t+j} - \mathbb{E}_t i^*_{t+j}\right) \end{aligned} \tag{B12}$$

Denoting f_t as $-\left(i^*_t - \mathbb{E}_t \Delta p^*_{t+1}\right) + \frac{\gamma^s_x}{\gamma^s_\pi} x_t = -\left(i^*_t - \mathbb{E}_t \Delta p^*_{t+1}\right) + \frac{\gamma_x}{\gamma_\pi} x_t$:

$$z_{2,t} = \left(1 - \gamma_\pi^s\right) \Delta p_t^* - \gamma_\pi^s \sum_{j=0}^{\infty} \left(\frac{1 - \gamma_\pi^s}{\rho}\right)^{j+1} \mathbb{E}_t f_{t+j} - \left(\frac{1 - \gamma_\pi^s}{\rho}\right) \eta_t - \frac{\left(1 - \gamma_\pi^s\right)}{\rho - \left(1 - \gamma_\pi^s\right)} \iota \quad \text{(B13)}$$

From equations (B6) and (B10):

$$z_{3,t} = -\left(1 - \gamma_\pi^s\right) i_{t-1}^* + \upsilon_t \quad \text{(B14)}$$

From equations (B7), (B13) and (B14):

$$\Delta p_t = \left(1 - \gamma_\pi^s\right) \Delta p_t^* - \gamma_\pi^s \sum_{j=0}^{\infty} \left(\frac{1 - \gamma_\pi^s}{\rho}\right)^{j+1} \mathbb{E}_t f_{t+j}$$
$$- \left(\frac{1 - \gamma_\pi^s}{\rho}\right) \eta_t + \frac{\left(1 - \gamma_\pi^s\right)}{\left(1 - \gamma_\pi^s\right) - \rho} \iota - \left(1 - \gamma_\pi^s\right) i_{t-1}^* + \upsilon_t \quad \text{(B15)}$$

Updating equation (B15) once and applying the law of iterated expectations:

$$\Delta p_{t+1} = \hat{\iota} + \left(1 - \gamma_\pi^s\right) \Delta p_{t+1}^* - \left(1 - \gamma_\pi^s\right) i_t^* - \gamma_\pi^s \sum_{j=0}^{\infty} \left(\frac{1 - \gamma_\pi^s}{\rho}\right)^{j+1} \mathbb{E}_t f_{t+j} + \omega_{t+1} \quad \text{(B16)}$$

where

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

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

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 differently, when there is a shock, Δp_{t+1} instantly *jumps* to its long-run equilibrium.

Conversely, Δe_{t+1} does have inertia. From equation (B7):

$$\Delta e_t = z_{1,t} + \Delta p_t \quad \text{(B17)}$$

Plug equation (B11) into (B17) and update it once to get

$$\Delta e_{t+1} = \Delta p_{t+1} - \left(1 - \gamma_\pi^s\right) \Delta p_{t+1}^* + \sum_{j=0}^{\infty} \rho^j u_{t-j+1} - \left(1 - \gamma_\pi^s\right) \sum_{j=0}^{\infty} \rho^j \eta_{t-j} \quad \text{(B18)}$$

where Δp_{t+1} contains a rational expectation of future fundamentals as defined in equation (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.

APPENDIX C: GMM WITH A NEAR-UNIT ROOT AND THE GRID BOOTSTRAP

APPENDIX 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; for example:

$$\alpha = 1 + c/n \tag{C1}$$

With this reparametrization, 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

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

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

$$\bar{\mathbf{g}}_n(\beta) = \frac{1}{n} \sum_{t=1}^{n} \mathbf{g}_t(\beta)$$

$$\boldsymbol{\Omega}_n(\beta) = \frac{1}{n} \sum_{t=1}^{n} \mathbf{g}_t(\beta) \mathbf{g}_t(\beta)'$$

$$\mathbf{G}_n(\beta) = \frac{1}{n} \sum_{t=1}^{n} \frac{\partial}{\partial \beta'} \mathbf{g}_t(\beta)$$

Let $\mathbf{m}_{t+1}$, $\mathbf{g}_t$, $\bar{\mathbf{g}}_n$, $\boldsymbol{\Omega}_n$ 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 $\widetilde{\beta}$, the GMM estimator $\widehat{\beta}$ minimizes $\bar{\mathbf{g}}_n(\beta)' \boldsymbol{\Omega}_n \left(\widetilde{\beta}\right)^{-1} \bar{\mathbf{g}}_n(\beta)$.

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

$$\sqrt{n}\left(\widehat{\beta} - \beta\right) = \left(\mathbf{G}_n' \boldsymbol{\Omega}_n^{-1} \mathbf{G}_n\right)^{-1} \mathbf{G}_n' \boldsymbol{\Omega}_n^{-1} \sqrt{n}\bar{\mathbf{g}}_n + o_p(1) \tag{C2}$$

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 $p = \dim(\beta)$. We can then write (C2) equivalently as

$$\sqrt{n}\delta_n\left(\widehat{\beta} - \beta\right) = \left(\overline{\mathbf{G}}_n'\overline{\mathbf{\Omega}}_n^{-1}\overline{\mathbf{G}}_n\right)^{-1}\overline{\mathbf{G}}_n'\overline{\mathbf{\Omega}}_n^{-1}\sqrt{n}\delta_n^{-1}\overline{\mathbf{g}}_n + o_p(1) \tag{C3}$$

where

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

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

$$\sqrt{n}\delta_n^{-1}\overline{\mathbf{g}}_n = \begin{pmatrix} \frac{1}{n}\sum_{t=1}^n s_t\varepsilon_{t+1} \\ \frac{1}{\sqrt{n}}\sum_{t=1}^n \varepsilon_{t+1} \\ \frac{1}{\sqrt{n}}\sum_{t=1}^n \mathbf{m}_{t+1} \end{pmatrix}$$
$$\Rightarrow \begin{pmatrix} \int_0^1 W_{1c}dW_1 \\ W_1(1) \\ W_2(1) \end{pmatrix} \tag{C4}$$
$$\equiv \mathbf{N}_c$$

$$\overline{\mathbf{\Omega}}_n = \begin{pmatrix} \frac{1}{n^2}\sum_{t=1}^n s_t^2\varepsilon_{t+1}^2 & \frac{1}{n^{3/2}}\sum_{t=1}^n s_t\varepsilon_{t+1}^2 & \frac{1}{n}\sum_{t=1}^n s_t\varepsilon_{t+1}\mathbf{m}_{t+1}' \\ \frac{1}{n^{3/2}}\sum_{t=1}^n s_t\varepsilon_{t+1}^2 & \frac{1}{n}\sum_{t=1}^n \varepsilon_{t+1} & \frac{1}{n}\sum_{t=1}^n \varepsilon_{t+1}\mathbf{m}_{t+1}' \\ \frac{1}{n}\sum_{t=1}^n s_t\mathbf{m}_{t+1}\varepsilon_{t+1} & \frac{1}{n}\sum_{t=1}^n \mathbf{m}_{t+1}\varepsilon_{t+1} & \frac{1}{n}\sum_{t=1}^n \mathbf{m}_{t+1}\mathbf{m}_{t+1}' \end{pmatrix}$$
$$\Rightarrow \begin{pmatrix} \int_0^1 W_{1c}^2\sigma_\varepsilon^2 & \int_0^1 W_{1c}\sigma_\varepsilon^2 & \int_0^1 W_{1c}\mu' \\ \int_0^1 W_{1c}\sigma_\varepsilon^2 & \sigma_\varepsilon^2 & \mu' \\ \int_0^1 W_{1c}\mu & \mu & \mathbf{M} \end{pmatrix} \tag{C5}$$
$$\equiv \mathbf{\Omega}_c$$

and

$$\overline{\mathbf{G}}_n = \begin{pmatrix} -\frac{1}{n^2}\sum_{t=1}^n s_t^2 & -\frac{1}{n^{3/2}}\sum_{t=1}^n s_t & 0 \\ -\frac{1}{n^{3/2}}\sum_{t=1}^n s_t & -1 & 0 \\ \frac{1}{n^{3/2}}\sum_{t=1}^n \frac{\partial}{\partial \alpha}\mathbf{m}_{t+1}(\beta) & 0 & \frac{1}{n}\sum_{t=1}^n \frac{\partial}{\partial \phi'}\mathbf{m}_{t+1}(\beta) \end{pmatrix}$$

$$\Rightarrow \begin{pmatrix} -\int_0^1 W_{1c}^2 & -\int_0^1 W_{1c} & 0 \\ -\int_0^1 W_{1c} & -1 & 0 \\ 0 & 0 & \mathbf{Q} \end{pmatrix} \tag{C6}$$

$$\equiv \mathbf{G}_c$$

Applying these distributional results to (C3), we find

$$\sqrt{n}\delta_n\left(\widehat{\beta} - \beta\right) \Rightarrow \left(\mathbf{G}_c'\mathbf{\Omega}_c^{-1}\mathbf{G}_c\right)^{-1}\mathbf{G}_c'\mathbf{\Omega}_c^{-1}\mathbf{N}_c \tag{C7}$$

The asymptotic distribution of $\widehat{\alpha}$ is obtained by taking the first element of this vector. Let $\mathbf{S}_1 = (10)'$ be a $(p+2)\times 1$ unit vector. Then

$$n\left(\widehat{\alpha} - \alpha\right) \Rightarrow \mathbf{S}_1'\left(\mathbf{G}_c'\mathbf{\Omega}_c^{-1}\mathbf{G}_c\right)^{-1}\mathbf{G}_c'\mathbf{\Omega}_c^{-1}\mathbf{N}_c \tag{C8}$$

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

$$n\,se\left(\widehat{\alpha}\right) = \left(n\mathbf{S}_1'\left(\mathbf{G}_n'\mathbf{\Omega}_n^{-1}\mathbf{G}_n\right)^{-1}\mathbf{S}_1\right)^{1/2}$$

$$= \left(\mathbf{S}_1'\left(\overline{\mathbf{G}}_n'\overline{\mathbf{\Omega}}_n^{-1}\overline{\mathbf{G}}_n\right)^{-1}\mathbf{S}_1\right)^{1/2}$$

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

Thus the t-ratio for α has the asymptotic distribution

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

We state this formally.

Proposition 1. Under (C1)

$$t_n(\alpha) \Rightarrow \frac{\mathbf{S}_1'\left(\mathbf{G}_c'\mathbf{\Omega}_c^{-1}\mathbf{G}_c\right)^{-1}\mathbf{G}_c'\mathbf{\Omega}_c^{-1}\mathbf{N}_c}{\left(\mathbf{S}_1'\left(\mathbf{G}_c'\mathbf{\Omega}_c^{-1}\mathbf{G}_c\right)^{-1}\mathbf{S}_1\right)^{1/2}} \tag{C9}$$

where $\mathbf{N}_c$, $\mathbf{\Omega}_c$, and $\mathbf{G}_c$ are defined in equations (C4), (C5) and (C6).

In the special case that ε_{t+1} and $\mathbf{m}_{t+1}$ are uncorrelated, then $\mu = 0$ and both $\mathbf{\Omega}_c$ and $\mathbf{G}_c$ are block diagonal. Then $\widehat{\alpha}$ is asymptotically independent of $\widehat{\beta}$ 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 $\widehat{\alpha}$ and $\widehat{\beta}$ 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.

APPENDIX 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 $\widehat{c} = n(\widehat{\alpha} - 1)$, which is inconsistent, as shown in equation (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 i.i.d. 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) satisfies the standard conditions for GMM estimation.

Proposition 2. Let A denote the grid bootstrap confidence interval defined in (24). Under Assumption 1, $\mathbb{P}(\alpha \in A) \to 0.95$ as $n \to \infty$.

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 and 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 i.i.d. 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.

APPENDIX D: STOCHASTIC SIMULATIONS

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

APPENDIX 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:

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

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:

$$i_t = \mathbb{E}_t \Delta e_{t+1} + i_t^* + v_t \tag{D1}$$

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

APPENDIX D.2: FIRMS (PRICE SETTING)

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

$$\mathbb{E}_t \Delta p_{t+1} = b \left[\mu - \left(p_t - p_t^* - e_t \right) \right] + \mathbb{E}_t \Delta p_{t+1}^* + \mathbb{E}_t \Delta e_{t+1} \tag{D2}$$

which could be transformed into

$$\begin{aligned} \mathbb{E}_t \Delta p_{t+1} =\ & (1-b) \Delta p_t - (1-b) \Delta p_t^* - (1-b) \Delta e_t + \mathbb{E}_t \Delta p_{t+1}^* + \mathbb{E}_t \Delta e_{t+1} \\ & + \left(p_t^* - \mathbb{E}_{t-1} p_t^* \right) + (e_t - \mathbb{E}_{t-1} e_t) - (p_t - \mathbb{E}_{t-1} p_t) \end{aligned} \tag{D3}$$

APPENDIX D.3: CENTRAL BANK

The domestic central banks follows the Taylor type rule:

$$i_t = (1 - \rho) \left(\gamma_\pi \mathbb{E}_t \Delta \hat{p}_{t+1} + \gamma_x x_t \right) + \rho i_{t-1} + u_t \tag{D4}$$

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

APPENDIX D.4: SYSTEM OF EQUATIONS

From equations (D1), (D2), (D3) 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:

$$\pi_t^e = q_t + \mathbb{E}_t \Delta p_{t+1}^* + s_t^e$$

$$\Delta \hat{p}_t = (1-b)\,\pi_{t-1}^e - (1-b)\,\Delta \hat{e}_t + s_t^e + \left(\Delta \hat{p}_t^* - \pi_{t-1}^{*e}\right) + \left(\Delta \hat{e}_t - s_{t-1}^e\right) - (1-b)\,\Delta \hat{p}_t^* + \mathbb{E}_t \Delta \hat{p}_{t+1}^*$$
(D5)

$$i_t = (1-\rho)\,\gamma_\pi \pi_t^e + \rho i_{t-1} + (1-\rho)\,\gamma_x x_t + u_t$$

$$i_t = s_t^e + i_t^* + v_t$$

$$\pi_t^e \equiv \mathbb{E}_t \Delta p_{t+1}$$

$$\pi_t^{*e} \equiv \mathbb{E}_t \Delta p_{t+1}^*$$

$$s_t^e \equiv \mathbb{E}_t \Delta e_{t+1}$$

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:Q3 to 1998:Q4. The USA is supposed to be the home country. Hence i_t^* and Δp_t^* are the 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 the UK. We use the weight by the FRB for the effective exchange rate. For x_t, we use unemployment rate gap in Boivin (2006).

APPENDIX 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