Yun-Yeong Kim*

Two-Step Optimal Prediction Under Phillips Triangular Cointegrated System

https://doi.org/10.1515/snde-2024-0107
Received October 1, 2024; accepted May 9, 2025; published online June 26, 2025

Abstract: This study proposes a two-step optimal best linear predictor (OBLP) under Phillips triangular cointegrated system, deduced from a two-step optimal forecasting method, for non-stationary level variables cointegrated with fundamental variables. In the first step, a cointegration equilibrium is estimated. The difference between the cointegration equilibrium and the other predicted variables is optimally forecasted in the second step, with conditional expectations estimated by the lagged fundamental differences and cointegration errors and summed with the cointegration equilibrium. We show that the OBLP has the lowest mean squared forecasting error among linear forecasting methods, such as random walk, cointegration, and augmented error correction models. In the second step, the cointegration error correction model is converted into a vector autoregression model consisting of the cointegration error and the fundamental differences of the variables and is used to estimate conditional expectations. Simulation results comparing the other predictors with the OBLP and forecast results for the US GDP and consumption applying the OBLP support the theoretical predictions of the forecasting efficiency of the OBLP.

Keywords: optimal best linear prediction; Phillips triangular cointegrated system; cointegrated level variable; two-step procedure

JEL Classification: C3

1 Introduction

> And God shall wipe away all tears from their eyes; and there shall be no more death, neither sorrow, nor crying, neither shall there be any more pain: for the former things are passed away. REVELATION 21-1.

The problem of using cointegration information for forecasting has focused on the role of cointegration errors or error correction terms in predicting differences in cointegrated variables. Classical examples in this area include Engle and Yoo (1987), Christoffersen and Diebold (1998), and Elliott (2006). However, when the focus is on predicting the level of a cointegrated variable, the long-run cointegration equilibrium becomes important, which has received little emphasis.[1] An exception is Kim (2023), who addresses this issue in the triangular form of Phillips (1991) for cointegration models.

In particular, Kim (2023) introduced the best linear predictor (BLP) with an asymptotic minimum mean squared forecasting error (MSFE) among the linear predictors of variables in cointegrated systems. Kim (2023)

1 For many economic variables (such as exchange rates, interest rates), it is very important to predict not only the rate of change but also the level itself.

I thank God for knowing everything and leading me to the Navotas Charity Foundation (http://navotas.or.kr/) and confess that I have received all grace through it. However, all remaining errors belong to the author.

***Corresponding author: Yun-Yeong Kim**, Department of International Trade, Dankook University, 126, Jukjeondong, Yongin-si, Gyeonggi-do 448-701, Korea, E-mail: yunyeongkim@dankook.ac.kr

showed that if the autocorrelation coefficient of the cointegration error between the prediction time and predicted targeting time is greater than 1/2, the BLP is deduced from the random walk model. In other cases, the BLP is deduced from the cointegration model. Under this scheme, Kim (2023) suggested a switching predictor that automatically selects a random walk or cointegration model according to the size of the estimated autocorrelation coefficient. He showed that the BLP has a weighted average form of the predictors using the random walk and cointegration models and has the lowest MSFE among the linear form predictors using the variables in cointegrated systems or their lagged variables. Here, there is a difference in $O_p(T^2)$ between the BLP and the other linear-form predictors. Note that the BLPs may differ in $O_p(1)$ depending on the weighting coefficient; however, Kim (2023) did not suggest a weighting coefficient that minimizes MSFE.

Under these circumstances, to improve the forecasting efficiency of the BLP, we propose the optimal best linear predictor (OBLP), which is deduced from a two-step optimal forecasting method for non-stationary level variables cointegrated with the fundamental variables. To do this, the cointegration equilibrium is estimated in the first step. The difference between the cointegration equilibrium and the other predicted variables is optimally forecasted in the second step, with conditional expectations estimated by the lagged fundamental differences and cointegration errors and summed with the cointegration equilibrium. We show that the OBLP has the lowest MSFE among linear forecasting methods, such as random walk, cointegration, and augmented error correction models. In the second step, the cointegration error correction model is converted into a vector autoregression (VAR) model consisting of the cointegration error and the difference in the fundamental variables and is used to estimate conditional expectations.

Note that following Engle and Yoo (1987), Christoffersen and Diebold (1998), and others, Elliott (2006, p. 584, Eq. 11) addressed the problem of optimal forecasting of co-integrated differenced variables in a bivariate VAR(1) model. An OBLP can equivalently be deduced by adding Elliott's (2006) predictor to the forecast baseline-level variable; however, an OBLP has not yet been provided for the general VAR(q) model.

The remainder of this paper is organized as follows. Section 2 introduces the optimal BLP and Section 3 discusses the OBLP estimation. Section 4 provides the Monte Carlo simulation results, and Section 5 presents an application to the prediction of United States' GDP and consumption. Finally, Section 6 concludes the paper.

2 Derivation of the OBLP

First, we assume that the $r \times 1$-vector y_t and the $k \times 1$-variable x_t explaining it are jointly represented by a VAR model; that is, we consider the ℓ ($\equiv r + k$)-dimensional and integrated of order one VAR(p) process of Y_t given by

$$Y_t = \Pi_0 + \Pi_1 Y_{t-1} + \Pi_2 Y_{t-2} + \cdots + \Pi_p Y_{t-p} + \varepsilon_t \tag{1}$$

or

$$\Delta Y_t = \Phi_0 + \Phi Y_{t-1} + \sum_{i=1}^{p-1} \Phi_i \Delta Y_{t-i} + \varepsilon_t \tag{2}$$

where $Y_t \equiv \left(x_t{}', y_t{}'\right)'$, $\overline{\Pi} = \sum_{i=1}^{p} \Pi_i$, $\Phi = \overline{\Pi} - I_\ell$, $\Phi_i = -\sum_{j=i+1}^{p} \Pi_j$ and ε_t is an $\ell \times 1$ vector of an independently and identically distributed (i.i.d henceforth) disturbance term with a finite variance $\Sigma > 0$, where I_ℓ denotes an ℓ-dimensional identity matrix and $\Delta Y_t \equiv Y_t - Y_{t-1}$.

Further, we assume the cointegration of Model (1) (e.g., Johansen 1991) as follows:

Assumption 1. We assume $\Phi = \alpha\beta'$, where α and β are $\ell \times r$ matrices of the full-column rank r where $\beta' \equiv \left(-\gamma', I_r\right)$ of rank r and γ is $k \times r$.

Note that Model (2) may be written as an error correction model (ECM) as

$$\Delta Y_t = \Phi_0 + \alpha z_{t-1} + \sum_{i=1}^{p-1} \Phi_i \Delta Y_{t-i} + \varepsilon_t \tag{3}$$

under Assumption 1, where $z_t = \beta' Y_t$.

Now, we transform Model (3) into a stationary VAR model of the I(0) variables Δx_t and z_t. To obtain this stationary VAR representation, we first define a non-singular square matrix as follows:

$$N \equiv \begin{pmatrix} I_k & 0_{k\times r} \\ -\gamma' & I_r \end{pmatrix}.$$

It should be noted that the lower triangular matrix N transforms the VAR variable $Y_t = (x_t', y_t')'$ into the variable w_t of x_t and cointegration error z_t.

$$N \times Y_t = (x_t', z_t')' \equiv w_t.$$

Following Kim (2012, 2018), we multiply the above matrix N on the left-hand side of Model (1) and modify the VAR coefficients to obtain the following VAR model of the purely stationary variable $\underset{\ell\times 1}{w_{\Delta t}} = (\Delta x_t', z_t')'$:

$$w_{\Delta t} = \psi_0 + \psi_1 w_{\Delta t-1} + \psi_2 w_{\Delta t-2} + \cdots + \psi_p w_{\Delta t-p} + e_t \tag{4}$$

where $e_t = N \times \varepsilon_t$, or a state space form:

$$W_{\Delta t} = \Psi_0 + \Psi W_{\Delta t-1} + \mathrm{e}_t \tag{5}$$

where $\underset{\ell p\times 1}{W_{\Delta t}} = \begin{pmatrix} w_{\Delta t} \\ w_{\Delta t-1} \\ \vdots \\ w_{\Delta t-p+1} \end{pmatrix}$, $\Psi_0 = \begin{pmatrix} \psi_0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}$, $\underset{\ell p\times \ell p}{\Psi} = \begin{pmatrix} \psi_1 & \psi_2 & \cdots & \psi_{p-1} & \psi_p \\ I_\ell & \cdots & & 0 & 0 \\ \vdots & I_\ell & & & \vdots \\ & & \ddots & & \\ 0 & & & I_\ell & 0 \end{pmatrix}$ and $\mathrm{e}_t = \begin{pmatrix} e_t \\ 0 \\ \vdots \\ 0 \end{pmatrix}$.

Note that the columns of ψ_p from the first to k-th are imposed as zero vectors/matrices, following Kim (2012, Theorem 3.2); thus, $\Delta x_{t\text{-}p}$ does not appear in Equation (4).[2]

We define two selection matrices, $M_{a,b} \equiv (I_a, 0_{a\times b})$ and $\overline{M}_{a,b} \equiv (0_{a\times b}, I_a)$. Now, equation (4) can be regarded as a Phillips (1991) triangular representation of a cointegrated system, as follows:

$$\underset{r\times 1}{y_t} = \underset{r\times 1}{\delta} + \underset{r\times 1}{\gamma' x_t} + \underset{r\times 1}{z_t} \tag{6}$$

and

$$\Delta x_t = \mu + u_t \tag{7}$$

for t = 1, 2,…, T, where $\mu = M_{k,r}\psi_0$, $\delta = \overline{M}_{r,k}\psi_0$, $u_t = M_{k,r}\left(\sum_{i=1}^{p}\psi_i w_{\Delta t-i} + e_t\right)$ and $z_t = \overline{M}_{r,k}\left(\sum_{i=1}^{p}\psi_i w_{\Delta t-i} + e_t\right)$.

At time t, we aim to predict the variables y_{it+h} for $1 \le i \le r$ and $h \in Z^+$, where Z^+ denotes a set of positive integers, and $y_t \equiv (y_{1t}, y_{2t}, ..., y_{rt})'$. Let $\left(\widehat{\delta}, \widehat{\gamma}'\right)$ be an OLS (ordinary least square) estimator of (δ, γ').

Furthermore, we assume the following standard regularity conditions, as in Kim (2023):

Assumption 2. *We assume*:

(a) $T^{-1}\sum_{t=1}^{T} z_t = O_p(1)$;

(b) $T^{-2}\sum_{t=1}^{T} x_t = O_p(1)$;

(c) $T^{-3/2}\sum_{t=1}^{T} x_t z_{t+h} = O_p(1)$;

(d) $T^{-3/2}\sum_{t=1}^{T} (x_{t+h} - x_t) x_t' = O_p(1)$;

(e) $T^{-3}\sum_{t=1}^{T} x_t x_t' = O_p(1)$;

2 Campbell and Shiller (1987, Equation 5) used the system (4) without referring to how it is derived from the VAR model of (1) by using the rank deficiency of matrix $\Phi = \alpha\beta'$ in Assumption 1, with the aforementioned zero restriction of the coefficient ψ_p.

(f) $T^{-1}\sum_{t=1}^{T} z_{t+h} z_t \to {}_p E(z_{t+h} z_t)$;

(g) $T^{-1}\sum_{t=1}^{T} u_t = o_p(1)$;

(h) $\begin{bmatrix} T^{1/2}\left(\hat{\delta} - \delta\right) \\ T^{3/2}\left(\hat{\gamma} - \gamma\right) \end{bmatrix} = O_p(1)$;

(i) $E|\varepsilon_t|^4$ *and* $E|w_{\Delta t}|^4 < \infty$.

Remark 1. (i) Assumptions 2 (b)–(e) hold because x_t has a drift term. Please refer to Hamilton (1994, Proposition 17.3). (ii) See Hamilton (1994, 7.2.15), from which Assumption 2 (f) holds with a certain stationarity assumption. ■

To improve forecasting efficiency, we first find the optimal predictor for a variable y_{t+h} that is considered to be dependent in the presence of r cointegrating relationships, as in (6), and use the information set $\Omega_t \equiv \left(Y_1', Y_2', \ldots, Y_t'\right)'$. Then, we suggest the optimal (*scalar*) predictor for y_{it+h}, which belongs to the original y_{t+h} that we want to predict. Since this method uses the system-wide cointegrated error terms simultaneously for prediction, it may have a lower MSFE than finding the optimal predictor for y_{it+h} restrictively. This kind of prediction efficiency improvement is possible if the cointegrated error term of the dependent variable to be predicted is highly correlated with the error terms of the other dependent variables (i.e. y_t except y_{it}).

Next, the MSFE of predictor b (F_t^b) is defined as[3]

$$MSFE^b \equiv T^{-1}\sum_{t=1}^{T}\left(y_{t+h} - F_t^b\right)^2. \tag{8}$$

We then consider the following class of linear predictors (LP) as a baseline for evaluating the optimality of the predictors that we introduce:

$$F_t^{LP} \equiv \theta + \theta_n{}' n_t + s_t, \tag{9}$$

where θ is a $r \times 1$, and θ_n is a $\ell p \times r$, vector/matrix of coefficients, respectively and, for instance, $n_t = \left(Y_t', Y_{t-1}', \ldots, Y_{t-p+1}'\right)'$ is a typical $\ell p \times 1$ I(1) variable selected/generated from a set $\{Y_{t-i}\}_{i=0}^{+\infty}$, and s_t is a $r \times 1$ I(0) variable selected/generated from a set $\{\Delta Y_{t-i}, z_{t-i}\}_{i=0}^{+\infty}$.

We define the best linear predictor (BLP) from Kim (2023, Eq. 2.5) as follows.[4]

$$F_t^{BLP} = \delta + \gamma' x_t + s_t \tag{10}$$

where s_t denotes an $r \times 1$ $O_p(1)$ variable. For instance if $s_t = z_t$, then F_t^{BLP} is a random walk model predictor (an RWP); if $s_t = 0$, then F_t^{BLP} is a cointegration model predictor (a CIP); if $s_t = \lambda\varepsilon_t$, then F_t^{BLP} is one of the predictors of Christoffersen and Diebold (1998, p. 13).[5] Elliott (2006, p. 584, Eq. 11) illustrates a predictor (interpreted as a *level* predictor) in a bivariate VAR(1) model, as follows:

$$\begin{aligned} F_t^{EL} &= y_t + \left(\sum_{i=1}^{h} \rho_c^{i-1}\right)\alpha_2 z_t \\ &= \gamma' x_t + \left\{1 + \left(\sum_{i=1}^{h} \rho_c^{i-1}\right)\alpha_2\right\} z_t \end{aligned} \tag{11}$$

which is a BLP where $\rho_c = \beta\alpha'$ and $\alpha = \left(\alpha_1', \alpha_2'\right)'$.

3 For the convenience of the analysis, it is assumed that the sample sizes for the model coefficient and MSFE estimation are all equal to T. Note that the predictor is time t dependent.

4 In Kim (2023, Eq. 2.5), this is given by $s_t = wz_t$, but since w is an arbitrary real number, there is no difference between the BLP definitions in (10) and Kim (2023, Eq. 2.5) except for the vector generalization.

5 However, this cannot be obtained with finite data, as it requires infinite lagged variables to identify the moving average error term ε_t.

However, unrestricted VAR models are generally not based on the BLPs. This is because, for example, a predictor using a VAR(1) model has the form $\pi_1{}'x_t + \pi_2{}'x_{t-1}$ and even if $\pi_1 = \gamma$ holds, the rest of the $\pi_2{}'x_{t-1}$ is not I(0) in general.

Note that the BLPs may differ in $O_p(1)$ depending on the form of the I(0) variable s_t; however, Kim (2023) did not suggest a weighting coefficient that minimizes the MSFE. Therefore, we now suggest the optimal BLP (OBLP) of y_t that minimizes the MSFE among the BLPs. For this purpose, we first exploit the following decomposition:

Proposition 1.

$$y_{t+h} = \delta + \gamma' x_t + E_t(q_{t+h}) + \varepsilon_{t,h} \tag{12}$$

where $q_{t+h} \equiv \gamma'(x_{t+h}\text{-}x_t) + z_{t+h}$, $E_t(q_{t+h}) = K_{0h} + K_{1h}W_{\Delta t}$ *and* $\varepsilon_{t,h} \equiv \sum_{j=1}^{h} \gamma_{\perp j} \sum_{i=1}^{j} \Psi^{j-i} \mathrm{e}_{t+i}$ *with* $\mathrm{K}_{0h} \equiv \sum_{j=1}^{h} \gamma_{\perp j} \left[\sum_{i=0}^{j-1} \Psi^{j} \right] \Psi_0$, $\mathrm{K}_{1h} \equiv \sum_{j=1}^{h} \gamma_{\perp j} \Psi^{j}$ *and*

$$\gamma_{\perp j} = \begin{cases} (\gamma', I_r) M_{\ell,\ell(p-1)} \; if\ j = h \\ (\gamma', 0_{r\times r}) M_{\ell,\ell(p-1)} \; otherwise \end{cases};$$

where $E_t(q_{t+h})$ *denotes a conditional expectation of* q_{t+h} *at a time t.*

However, the conditional expectation $E_t(y_{t+h})$, which is the optimal predictor when y_t is I(0), is not defined in general because there are no finite moments of y_t when y_t is I(1). Therefore, OBLP for y_{t+h} is defined from (12) as the long-run cointegration equilibrium of y_t after adding the conditional expectation of q_{t+h}, which is I(0), as follows:

$$F_t^{OBLP} = \delta + \gamma' x_t + \mathrm{K}_{0h} + \mathrm{K}_{1h} W_{\Delta t}. \tag{13}$$

We now derive the difference in MSFE between the LP and OBLP as follows:

Theorem 1. *Suppose that Assumptions* 1 and 2 *hold. Further suppose that* $\sum_{t=1}^{T} n_t s_t{}' = O_p(T^{3/2})$ *and* $\sum_{t=1}^{T} n_t \varepsilon_{t,h}{}' = O_p(T^{3/2})$*. Then*

$$MSFE^{LP} - MSFE^{OBLP} = \underbrace{(\theta_n - \gamma_n)' \left(T^{-1} \sum_{t=1}^{T} n_t n_t{}' \right) (\theta_n - \gamma_n)}_{O_p(T^2)} + O_p(T)$$

where $\gamma_n = \begin{pmatrix} \gamma \\ 0_{(\ell p-k)\times r} \end{pmatrix}$.

According to Theorem 1, the LP has a larger MSFE than the OBLP owing to a positive definite matrix of size $O_p(T^2)$. Next, the difference in the MSFE between the BLP and OBLP is given as

Corollary 1. *Suppose that* Assumptions 1 *and* 2 *hold and* $T^{-1}\sum_{t=1}^{T} \varepsilon_{t,h} \left[s_t - E_t(q_{t+h}) \right]' \to_p 0.$[6] *Then*

$$MSFE^{BLP} - MSFE^{OBLP} = T^{-1} \sum_{t=1}^{T} \left[s_t - E_t(q_{t+h}) \right] \left[s_t - E_t(q_{t+h}) \right]' + o_p(1),$$

where $T^{-1}\sum_{t=1}^{T} \left[s_t - E_t(q_{t+h}) \right] \left[s_t - E_t(q_{t+h}) \right]' \geq 0.$

6 It may hold from the law of large numbers, for instance, in Hamilton (1994, pp. 193–5) where $\{ [s_t - E_t(q_{t+h})] \varepsilon_{t,h}{}' \}$ is a martingale difference sequence $E_t([s_t - E_t(q_{t+h})] \varepsilon_{t,h}{}') = 0$.

Next, the OBLP for y_{it+h} is given by, which has the minimum MSFE among the BLPs. To demonstrate this, we first define two predictors of y_{it+h}.

$$F_t^{OBLPi} = m_i{}' F_t^{OBLP} \tag{14}$$

and

$$F_t^{BLPi} = m_i{}' \left(\delta + \gamma' x_t\right) + c_t \tag{15}$$

where $m_i = (0_{(i\text{-}1)\times 1}, 1, 0_{(r\text{-}i)\times 1})$ and c_t is an arbitrary I(0) real variable.

For instance, F_t^{BLPi} is a predictor of y_{it+h} using the cointegration error z_{it}[7] in a VAR(1) model that is conformable with y_{it+h}. Note that this is not a predictor deduced from the OBLP of y_{t+h}, as in (13), using all cointegration error vectors z_t.

Accordingly, the optimality of the predictor (14) for y_{it+h} is given by

Theorem 2.

$$MSFE^{BLPi} - MSFE^{OBLPi} \geq o_p(1)$$

for any c_t.

3 Estimation of the OBLP

In this section, we introduce a consistent estimator of the OBLP and demonstrate that the estimated OBLP asymptotically has the same MSFE as the OBLP. To do so, we first rewrite the last three terms on the right-hand side of Equation (12) as follows:

$$q_{t+h} = \mathrm{K}_{0h} + \mathrm{K}_{1h} W_{\Delta t} + \varepsilon_{t,h} \tag{16}$$

because $q_{t+h} = E_t(q_{t+h}) + \varepsilon_{t,h}$ from definition.

Then, we define the estimated OBLP as

$$F_t^{\widehat{OBLP}} = \widehat{\delta} + \widehat{\gamma}' x_t + \widehat{\mathrm{K}}_{0h} + \widehat{\mathrm{K}}_{1h} \widehat{W}_{\Delta t} \tag{17}$$

where the coefficients in (16) are estimated using OLS as follows:

$$\left(\widehat{\mathrm{K}}_{0h}, \widehat{\mathrm{K}}_{1h}\right) \equiv \sum_{t=1}^{T} \widehat{q}_{t+h} \left(1 \quad \widehat{W}_{\Delta t}{}'\right) \left(\sum_{t=1}^{T} \begin{bmatrix} 1 & \widehat{W}_{\Delta t}{}' \\ \widehat{W}_{\Delta t} & \widehat{W}_{\Delta t} \widehat{W}_{\Delta t}{}' \end{bmatrix}\right)^{-1} \tag{18}$$

where $\widehat{z}_t \equiv y_t - \widehat{\delta} - \widehat{\gamma}' x_t$, $\widehat{w}_{\Delta t} = \left(\Delta x_t{}', \widehat{z}_t{}'\right)'$,

$$\underset{r\times(k+1)}{\left(\widehat{\delta}, \widehat{\gamma}'\right)}' \equiv \left(\sum_{t=1}^{T} \underset{(k+1)\times(k+1)}{\begin{bmatrix} 1 & x_t{}' \\ x_t & x_t x_t{}' \end{bmatrix}}\right)^{-1} \sum_{t=1}^{T} \underset{(k+1)\times r}{\begin{bmatrix} y_t{}' \\ x_t y_t{}' \end{bmatrix}},$$

$$\widehat{W}_{\Delta t} = \begin{pmatrix} \widehat{w}_{\Delta t} \\ \widehat{w}_{\Delta t-1} \\ \vdots \\ \widehat{w}_{\Delta t-p+1} \end{pmatrix}.$$

7 These predictors correspond to cases where there are multiple cointegration vectors in the VAR model, but the predictions do not reflect them, and only one cointegration vector is used to build the OBLP.

Now note that

Lemma 1. *Suppose* Assumptions 1 *and* 2 *hold. Then,* $\widehat{\mathrm{K}}_{0h} - \mathrm{K}_{0h} \rightarrow_p 0$ *and* $\widehat{\mathrm{K}}_{1h} - \mathrm{K}_{1h} \rightarrow_p 0$.

We now show that the suggested OBLP estimator has the same MSFE as the OBLP in (15), as follows:

Theorem 3. *Suppose that Assumptions* 1 *and* 2 *hold. Then*

$$MSFE^{\widehat{OBLP}} \rightarrow_p MSFE^{OBLP}$$

where $MSFE^{\widehat{OBLP}} = T^{-1}\sum_{t=1}^{T}\left(y_{t+h} - F_t^{\widehat{OBLP}}\right)\left(y_{t+h} - F_t^{\widehat{OBLP}}\right)'$.

Finally, the OBLP for y_{it+h} is given by

$$F_t^{\widehat{OBLPi}} = m_i'\left[\hat{\delta} + \hat{\gamma}' x_t + \widehat{\mathrm{K}}_{0h} + \widehat{\mathrm{K}}_{1h}\widehat{W}_{\Delta t}\right].$$

The OBLP estimated in this manner can also be shown to have the same predictive efficiency as Theorem 3 based on the consistency of the OLS estimates.

Corollary 2. *Suppose that Assumptions* 1 and 2 *hold. Then*

$$MSFE^{\widehat{OBLPi}} \rightarrow_p MSFE^{OBLPi}$$

where $MSFE^{\widehat{OBLPi}} = T^{-1}\sum_{t=1}^{T}\left(y_{it+h} - F_t^{\widehat{OBLPi}}\right)\left(y_{it+h} - F_t^{\widehat{OBLPi}}\right)'$.

In the example below, we show how the OBLP presented earlier is applied to the VAR(1) model.

Example 1. Under the VAR(1) of Y_t without a constant-term model, the ECM is given by[8]

$$\Delta Y_t = \alpha z_{t-1} + \varepsilon_t \tag{19}$$

where $\alpha = \begin{pmatrix}\alpha_1 \\ \alpha_2\end{pmatrix}$, $z_t = \beta' Y_t \equiv (-\gamma', I_r)\begin{pmatrix}x_t \\ y_t\end{pmatrix}$ and $\varepsilon_t \equiv \begin{pmatrix}\varepsilon_{xt} \\ \varepsilon_{yt}\end{pmatrix}$ conformably.

Then, Equation (19) may also be written in the VAR(1) form of $w_{\Delta t}$:

$$w_{\Delta t} = \Psi w_{\Delta t-1} + e_t \tag{20}$$

where $w_{\Delta t} = \begin{pmatrix}\Delta x_t \\ z_t\end{pmatrix}$, $\Psi = \begin{pmatrix}0 & \alpha_1 \\ 0 & \beta\alpha' + I_r\end{pmatrix}$ and $e_t = \begin{pmatrix}\varepsilon_{xt} \\ \varepsilon_{yt} - \gamma'\varepsilon_{xt}\end{pmatrix} \equiv \begin{pmatrix}e_{1t} \\ e_{2t}\end{pmatrix}$.

Note that the OBLP becomes

$$F_t^{OBLP} = \gamma' x_t + \left\{\left(\sum_{i=1}^{h-1}\gamma\alpha_1\left[\beta\alpha' + I_r\right]^{i-1}\right) + \left(\gamma\alpha_1\left[\beta\alpha' + I_r\right]^{h-1} + \left[\beta\alpha' + I_r\right]^{h}\right)\right\} z_t \tag{21}$$

8 Litterman (1986) proposed a Minnesota prior associated with a VAR(1) structure for Bayesian VAR prediction using a random walk model.

where

$$
\begin{aligned}
E_t(q_{t+h}) &= \sum_{j=1}^{h} \gamma_{\perp j} \Psi^j w_{\Delta t} \\
&= \sum_{j=1}^{h} \gamma_{\perp j} \begin{pmatrix} 0 & \alpha_1 [\beta\alpha' + I_r]^{j-1} \\ 0 & [\beta\alpha' + I_r]^{j} \end{pmatrix} w_{\Delta t} \\
&= (\gamma', I_r) \begin{pmatrix} 0 & \alpha_1 [\beta\alpha' + I_r]^{h-1} \\ 0 & [\beta\alpha' + I_r]^{h} \end{pmatrix} w_{\Delta t} \\
&\quad + \sum_{j=1}^{h-1} (\gamma', 0_{r\times r}) \begin{pmatrix} 0 & \alpha_1 [\beta\alpha' + I_r]^{j-1} \\ 0 & [\beta\alpha' + I_r]^{j} \end{pmatrix} w_{\Delta t} \\
&= \left\{ \gamma' \alpha_1 [\beta\alpha' + I_r]^{h-1} + [\beta\alpha' + I_r]^{h} + \sum_{j=1}^{h-1} \gamma' \alpha_1 [\beta\alpha' + I_r]^{j-1} \right\} z_t
\end{aligned}
\tag{22}
$$

because

$$
\Psi^i = \begin{pmatrix} 0 & \alpha_1 [\beta\alpha' + I_r]^{i-1} \\ 0 & [\beta\alpha' + I_r]^{i} \end{pmatrix}.
$$

Note that if $\alpha_1 = 0$,[9] then (22) becomes

$$
E_t(q_{t+h}) = (\beta\alpha' + I_r)^h z_t, \tag{23}
$$

where $\beta\alpha' = \alpha_2$. In this case, we may get

$$
F_t^{OBLP} = \gamma' x_t + (\beta\alpha' + I_r)^h z_t \tag{24}
$$

from (13).

Note that if $r = 1$ and $|\beta\alpha' + 1| < 1$, then the OBLP (24) approaches the RWP if h is small, while if h is large it approaches the CIP because $(\beta\alpha' + I_r)^h$ is close to zero. Thus, the VAR(1) model under the restriction $\alpha_1 = 0$ can be viewed as a generalized model that may approximate the RWP or the CIP depending on h. ■

4 Monte Carlo Simulation Results

We conducted a Monte Carlo experiment[10] to verify the small-sample properties of the proposed predictors. The basic simulation model used has the following triangular form:[11]

$$
\underset{2\times 1}{y_t} = \delta + \gamma' \underset{2\times 1}{x_t} + z_t, \tag{25}
$$

$$
x_t = \mu + x_{t-1} + u_t, \tag{26}
$$

9 This implies that x_t is exogenous and z_t does not Granger cause Δx_{t+i}.

10 GAUSS20 was used for the simulations, and the codes are available on request.

11 Equations (25)–(27) can be expressed in a state space form as follows (and this is exploited in the simulation):

$$
\begin{aligned}
y_t &= \delta + (\gamma', I_2)(x_t{}', z_t{}')' \\
\begin{pmatrix} x_t \\ z_t \end{pmatrix} &= \begin{pmatrix} \mu \\ 0 \end{pmatrix} + \begin{pmatrix} 0_{2\times 2} & 0 \\ 0 & \Psi \end{pmatrix} \begin{pmatrix} x_{t-1} \\ z_{t-1} \end{pmatrix} + \begin{pmatrix} u_t \\ \varepsilon_t \end{pmatrix}.
\end{aligned}
$$

and

$$z_t = \Psi z_{t-1} + \varepsilon_t \tag{27}$$

for $t = 1, 2, \ldots, 100$ and $(u_t, \varepsilon_t)' \sim \text{IIDN}(0, \text{I}_4)$, where $y_t = (y_{1t}, y_{2t})'$. The parameters are set as follows: $\gamma' = \begin{pmatrix} 0.5 & 0.1 \\ 0.1 & 0.5 \end{pmatrix}$, $\Psi = \begin{pmatrix} \lambda & \rho \\ \rho & \lambda \end{pmatrix}$, $\lambda = 0.5$ or 0.7, $\rho = 0.1, \ldots, 0.4$, $\delta = \begin{pmatrix} 0.1 \\ 0.1 \end{pmatrix}$ and $\mu = \begin{pmatrix} 0.1 \\ 0.1 \end{pmatrix}$ or $\begin{pmatrix} 0.5 \\ 0.5 \end{pmatrix}$, respectively. Here, ρ denotes the correlation of the different co-integration errors.

Then, y_{1t+h} is predicted at time t using 100 samples, where $h = 1, 2, \ldots, 20$. The MSFEs are calculated for the five predictors; RWP, CIP, OBLP, restricted OBLP using only the cointegration error of the predicted variable y_{1t} (OBLP1) and ECM predictor (ECMP).[12] We also add a deterministic trend to the predictors to obtain

(i) RWP: $(1, 0)' \times (y_t + h\gamma'\mu)$
(ii) CIP: $(1, 0)' \times (\delta + \gamma' x_t + h\gamma'\mu)$
(iii) ECMP: $(0, 0, 1, 0)' \left\{ Y_t + E\left[(Y_{t+h} - Y_t) z_t\right] E\left[z_t z_t'\right]^{-1} z_t \right\} + (1, 0)' \times h\gamma'\mu$

where $Y_t = (x_t', y_t')'$.

Subsequently, the MSFEs are computed as the mean of the samples from 10,000 repetitions of the aforementioned experiments.

The simulation results in Appendix A confirm the theoretical expectation; that is, when the prediction period h is small, the OBLP has the best MSFE among the predictors in terms of prediction stability and efficiency. In Appendix B, the ratio of the relative size of the MSFE of predictor b to that of the OBLP is calculated as

$$\frac{MSFE^b - MSFE^{OBLP}}{MSFE^{OBLP}}$$

and plotted it on a graph with the forecast horizon (1–20) on the x-axis.

From the calculations, we obtained the following results. First, the variation in the cointegration matrix γ does not significantly change the forecast results. Second, an increase in ρ leads to an increase in the forecasting efficiency of OBLP compared with that of OBLP1, which seems to be because the increased forecasting efficiency of OBLP uses additional cointegration errors from other equations in the forecast.

Third, the CIP shows a much lower forecast efficiency than the OBLP for short-term forecasts but slightly better efficiency than the OBLP as the forecast horizon increases. However, as the error correction process slows (i.e. as λ increases), the decrease in the MSFE of the CIP relative to that of the OBLP is delayed as the forecast period increases.

Finally, the RWP is inferior to the OBLP over the period, but converges to the OBLP as the forecast horizon increases, whereas the ECMP is less efficient than the OBLP as the forecast horizon increases. This phenomenon is further exacerbated as μ, which represents the magnitude of the deterministic trend, increases.

---

12 From the ECMs, as in Elliott (2006), the following equation is derived by iterative substitution.

$$Y_{t+h} - Y_t = \alpha \left[\sum_{j=0}^{h-1} (\beta'\alpha + I_2)^j \right] z_t + \widetilde{\varepsilon}_{t+h}$$

where $\widetilde{\varepsilon}_{t+h}$ consists of the error terms after time t. Therefore, the ECMP is given as in (iii).

$$F_t^{ECM} = (0, 0, 1, 0)' \left\{ Y_t + E\left[(Y_{t+h} - Y_t) z_t\right] E\left[z_t z_t'\right]^{-1} z_t \right\}.$$

Table 1: Unit root test Results.

Included terms	None	Intercept		Trend and intercept	
	ADF	ADF	ERS	ADF	ERS
GDP	1.000	0.000	5,675.509	0.0088	264.7067
Consumption	1.000	0.000	5,840.787	0.0441	265.2141
Net exports	0.7157	0.7219	22.51895	0.2347	7.019256
Interest rate term spread	0.5398	0.7992	6.203761	0.2180	3.012781
Government expenditure	1.0000	0.2782	1,026.363	0.0637	24.87423

1) P-value for the null hypothesis: The variable has a unit root, and the lag length is selected using the Schwarz criterion. 2) The critical values for the 1 % level are 1.99 (when an intercept is included in the test equation) and 3.96 (when the trend and intercept are included in the test equation) according to Elliott et al. (1996, Table 1). Autoregressive spectral ordinary least squares (OLS) was used as an estimation method.

5 Application to the United States GDP and Consumption Prediction

In this section, we conduct out-of-sample forecasts for US GDP and consumption using the predictors suggested in Section 4. We compare the forecast performance with the MSFE calculated using h-period (1, 2,…, 20)-ahead estimated forecast errors. The data used have a quarterly frequency that extends from Q3 1976 to Q3 2023.[13] Therefore, the analysis of the out-of-sample predictive performance of the proposed model consists of forecasting US GDP and consumption for each quarter from Q1 2018 to Q3 2023 using data from Q3 1976 to Q4 2017.

The data source is the United States Federal Reserve Board at St. Louis FRED. The cointegration fundamentals initially considered for US GDP and consumption are interest rate term spread, net exports, and government expenditure. All variables, except interest rate term spread and net exports, are log-transformed.

Before proceeding, we conduct augmented Dickey–Fuller (ADF) and Elliott–Rothenberg–Stock (ERS) point optimal tests to check the unit root of the variables considered. Table 1 presents the unit root test results. The ADF test results show that the null hypothesis (i.e. that the variable has a unit root) is not rejected at the 1 % significance level when the test equation does not include a trend or intercept term. The results of the ERS test show that the null hypothesis is not rejected at the 1 % significance level when the test equation includes a trend or an intercept term.

Therefore, although this is somewhat restrictive for GDP and consumption in the added trend or intercept term cases, we assume that all variables have unit roots and proceed with the following analysis:

We then conduct Johansen cointegration tests using a VAR model to check whether a cointegration vector exists in the VAR model. We set the lag length of the VAR model to 1, based on the most parsimonious Schwarz information criterion. The Johansen test results, shown in Table 2, indicate that the trace and maximum eigenvalue tests jointly indicate one cointegrating equation at the level of 0.05.

Next, we estimate the forecasting model presented in Section 4 and calculate the forecast errors, as shown in Appendix C. In Figure 1, the ratio of the relative size of the MSFE of predictor b to the OBLP is calculated as follows:

$$\frac{MSFE^{b} - MSFE^{OBLP}}{MSFE^{OBLP}}$$

and plotted it on a graph with the forecast horizon (1–20) on the x-axis. The prediction errors for GDP in Figure 1 show that OBLP1 outperforms the other predictors for most forecast horizons when evaluated based on its forecasting efficiency and stability. The CIP is the strongest in long-run forecasting, but it shows a very large absolute value of forecast error in the short-term horizon.

13 This is the maximum period for which data are available. See the end of this section for a detailed description of these variables.

Table 2: Cointegration rank test results

Trace				
Hypothesized No. of CE(s)	**Eigenvalue**	**Trace Statistic**	**0.05 Critical value**	**Prob.**[b]
None[a]	0.554488	185.9750	60.06141	0.0000
At most 1	0.083672	34.77985	40.17493	0.1572
At most 2	0.056804	18.43968	24.27596	0.2279
At most 3	0.029847	7.503785	12.32090	0.2778
At most 4	0.009778	1.837474	4.129906	0.2062
Maximum eigenvalue				
Hypothesized No. of CE(s)	**Eigenvalue**	**Max-Eigen Statistic**	**0.05 Critical value**	**Prob.**[b]
None[a]	0.554488	151.1952	30.43961	0.0000
At most 1	0.083672	16.34017	24.15921	0.3936
At most 2	0.056804	10.93589	17.79730	0.3917
At most 3	0.029847	5.666310	11.22480	0.3892
At most 4	0.009778	1.837474	4.129906	0.2062

Max-eigenvalue test indicates one cointegrating equation at the 0.05 level. [a]denotes rejection of the hypothesis at the 0.05 level. [b]MacKinnon et al. (1999) p-values.

The forecasting and estimation results are generally consistent with the macroeconomic theory. First, OBLP1 has the best forecasting results when using the interest rate term spread, government expenditure, and net exports as co-integrating explanatory variables for GDP (or consumption).[14] For example, adding M1 and consumption to the GDP forecast or excluding government expenditure leads to worse forecasting results.[15] This is likely because M1 has a low direct correlation with GDP, or it may be because consumption already has redundant information about GDP forecasting that interest rate term spread, government expenditure and net exports already have, and therefore does not contribute much to the forecast.

However, changing the interest rate term spreads from the 10-year treasury constant maturity minus the 2-year treasury constant maturity to the 10-year treasury constant maturity minus the federal fund rate (FFR) seems to reduce the forecasting efficiency of the OBLP because the 2-year treasury constant maturity interest rate reflects the investment securities market conditions more closely than the FFR.

We also find that the single cointegration vector OBLP1 yields better forecasting results than the OBLP, with the two cointegration vectors of GDP and consumption as dependent variables. This reflects the lower correlation between errors in the cointegration of GDP and consumption, suggesting that the error correction process mechanisms for GDP and consumption are different. In addition, the forecasts of the CIP and OBLP class models tend to converge as the forecast horizon h increases.

14 In economic theory, these variables represent exogenous monetary (interest rate term spread and M1) and fiscal (government expenditure) policies and foreign shocks (net exports), respectively. In particular, see Estrella and Hardouvelis (1991), Estrella and Mishkin (1996, 1998), and Kishor and Koenig (2010) on the predictability of economic downturns from the interest rate term spread.

15 These additional estimates are not reported in the text.

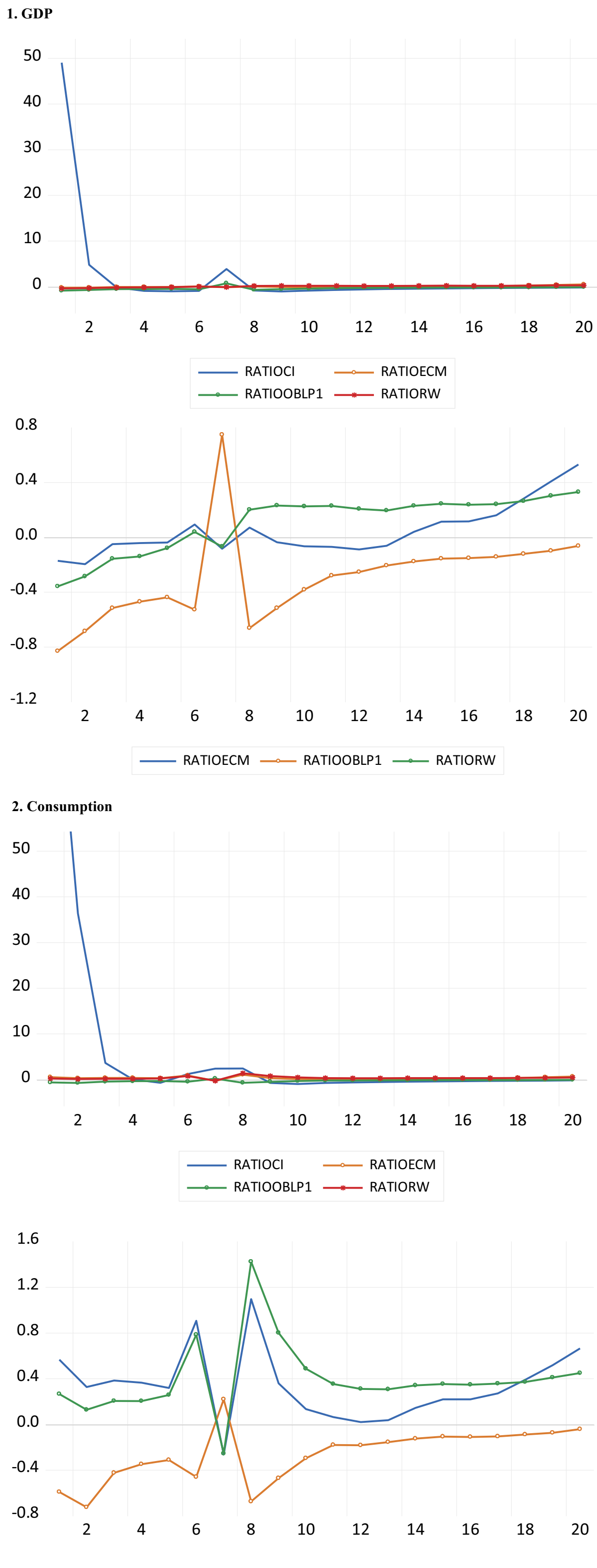


Figure 1: MSFE ratio relative to OBLP. Note: The ratio of the relative size of the MSFE of a predictor b to the OBLP is calculated as $\frac{MSFE^{b}-MSFE^{OBLP}}{MSFE^{OBLP}}$.

6 Conclusions

This study proposes the OBLP, which is deduced from a two-step optimal forecasting method for non-stationary-level variables cointegrated with fundamental variables. For this, the cointegration equilibrium is estimated in the first step. The difference between the cointegration equilibrium and other predicted variables is optimally forecasted in the second step, with conditional expectations estimated by the lagged fundamental differences and cointegration errors, and summed with the cointegration equilibrium. We show that the OBLP has the lowest MSFE among the linear forecasting methods of random walk, cointegration, and augmented error correction models. In the second step, the cointegration error correction model is converted into a VAR model consisting of the cointegration error and the difference in the fundamental variables and is used to estimate conditional expectations. In the simulation results, we compare the other predictors with the OBLP, and our forecast results for the US GDP and consumption applying the OBLP support the theoretical predictions of the forecasting efficiency of the OBLP.

Finally, it would be interesting to apply the predictions using the OBLP to other macroeconomic variables, such as interest rates, stock prices, and exchange rates, in an empirical analysis.

Proof of Theorems

Proposition 1: Note that

$$y_{t+h} = \delta + \gamma' x_t + q_{t+h}, \tag{28}$$

from (6) where we may write

$$\begin{aligned} q_{t+h} &= \gamma' \left(x_{t+h} - x_t \right) + z_{t+h} \\ &= \gamma' \sum_{j=1}^{h} \Delta x_{t+i} + z_{t+h} \\ &= \gamma' \Delta x_{t+h} + z_{t+h} + \gamma' \sum_{j=1}^{h-1} \Delta x_{t+j} \\ &= (\gamma', I_r) w_{\Delta t+h} + (\gamma', 0_{r\times r}) \sum_{j=1}^{h-1} w_{\Delta t+j} \\ &= (\gamma', I_r) M_{\ell,\ell(p-1)} W_{\Delta t+h} + (\gamma', 0_{r\times r}) M_{\ell,\ell(p-1)} \sum_{j=1}^{h-1} W_{\Delta t+j} \\ &= \sum_{j=1}^{h} \gamma_{\perp j} W_{\Delta t+j} = \mathrm{K}_{0h} + \mathrm{K}_{1h} W_{\Delta t} + \varepsilon_{t,h} \end{aligned} \tag{29}$$

where

$$w_{\Delta t+j} = M_{\ell,\ell(p-1)} W_{\Delta t+j}$$

and

$$W_{\Delta t+j} = \left[\sum_{i=0}^{j-1} \Psi^j \right] \Psi_0 + \Psi^j W_{\Delta t} + \sum_{i=1}^{j} \Psi^{j-i} \mathrm{e}_{t+i}$$

from a repetitive substitution in (5). So, the claimed result holds from (28) and (29). ■

Theorem 1: Note that

$$y_{t+h} - F_t^{LP} = \left(y_{t+h} - F_t^{OBLP}\right) + \left(F_t^{OBLP} - F_t^{LP}\right) \tag{30}$$

where

$$F_t^{OBLP} - F_t^{LP} = \delta - \theta + (\gamma_n - \theta_n)' n_t + E_t(q_{t+h}) - s_t, \tag{31}$$

and $\gamma' x_t \equiv \gamma_n' n_t$ from definitions (9) and (29).

Therefore, note that

$$\begin{aligned} MSFE^{LP} &\equiv T^{-1}\sum_{t=1}^{T}\left(y_{t+h} - F_t^{LP}\right)\left(y_{t+h} - F_t^{LP}\right)' \\ &= T^{-1}\sum_{t=1}^{T}\left(y_{t+h} - F_t^{OBLP}\right)\left(y_{t+h} - F_t^{OBLP}\right)' + T^{-1}\sum_{t=1}^{T}\left(F_t^{OBLP} - F_t^{LP}\right)\left(F_t^{OBLP} - F_t^{LP}\right)' \\ &\quad + T^{-1}\sum_{t=1}^{T}\left(y_{t+h} - F_t^{OBLP}\right)\left(F_t^{OBLP} - F_t^{LP}\right)' + T^{-1}\sum_{t=1}^{T}\left(F_t^{OBLP} - F_t^{LP}\right)\left(y_{t+h} - F_t^{OBLP}\right)' \\ &= MSFE^{OBLP} + O_p(T^2) \end{aligned} \tag{32}$$

because, from (31),

$$\begin{aligned} &T^{-1}\sum_{t=1}^{T}\left(F_t^{OBLP} - F_t^{LP}\right)\left(F_t^{OBLP} - F_t^{LP}\right)' \\ &= \underbrace{(\delta-\theta)(\delta-\theta)'}_{O_p(1)} + (\gamma_n-\theta_n)'\underbrace{T^{-1}\sum_{t=1}^{T} n_t n_t'}_{O_p(T^2)}(\gamma_n-\theta_n) + \underbrace{T^{-1}\sum_{t=1}^{T}\left[E_t(q_{t+h}) - s_t\right]\left[E_t(q_{t+h}) - s_t\right]'}_{O_p(1)} \\ &\quad + \underbrace{(\delta-\theta)T^{-1}\sum_{t=1}^{T} n_t'(\gamma_n-\theta_n)}_{O_p(T)} + \underbrace{(\gamma_n-\theta_n)'T^{-1}\sum_{t=1}^{T} n_t(\delta-\theta)'}_{O_p(T)} \\ &\quad + \underbrace{(\delta-\theta)T^{-1}\sum_{t=1}^{T}\left[E_t(q_{t+h}) - s_t\right]'}_{O_p(1)} + \underbrace{T^{-1}\sum_{t=1}^{T}\left[E_t(q_{t+h}) - s_t\right](\delta-\theta)'}_{O_p(1)} \\ &\quad + \underbrace{(\gamma_n-\theta_n)'T^{-1}\sum_{t=1}^{T} n_t\left[E_t(q_{t+h}) - s_t\right]'}_{O_p(T^{1/2})} + \underbrace{T^{-1}\sum_{t=1}^{T}\left[E_t(q_{t+h}) - s_t\right]n_t'(\gamma_n-\theta_n)}_{O_p(T^{1/2})} \end{aligned} \tag{33}$$

and

$$T^{-1}\sum_{t=1}^{T}\left(y_{t+h} - F_t^{OBLP}\right)\left(F_t^{OBLP} - F_t^{LP}\right)' = \underbrace{T^{-1}\sum_{t=1}^{T}\varepsilon_{t,h}\left[\theta - \delta + (\theta_n - \gamma_n)' n_t + s_t - E_t(q_{t+h})\right]'}_{O_p(T^{1/2})}$$

from Assumption 2 and $y_{t+h} - F_t^{OBLP} = \varepsilon_{t,h}$ using (12). Thus, the claimed result from Equation (32) holds. ■

Corollary 1: Note that

$$MSFE^{BLP} - MSFE^{OBLP} = T^{-1}\sum_{t=1}^{T}\left[s_t - E_t(q_{t+h})\right]\left[s_t - E_t(q_{t+h})\right]' + \underbrace{T^{-1}\sum_{t=1}^{T}\varepsilon_{t,h}\left[s_t - E_t(q_{t+h})\right]'}_{o_p(1)} + \underbrace{T^{-1}\sum_{t=1}^{T}\left[s_t - E_t(q_{t+h})\right]\varepsilon_{t,h}'}_{o_p(1)},$$

from assumption because

$$MSFE^{BLP} \equiv T^{-1}\sum_{t=1}^{T}\left(y_{t+h} - F_t^{OBLP} + F_t^{OBLP} - F_t^{BLP}\right)\left(y_{t+h} - F_t^{OBLP} + F_t^{OBLP} - F_t^{BLP}\right)'$$

and

$$F_t^{OBLP} - F_t^{BLP} = E_t(q_{t+h}) - s_t$$

from (12). Thus, the claimed results hold. ■

Theorem 2: Note that

$$m_i'\left(MSFE^{BLP} - MSFE^{OBLP}\right)m_i \geq o_p(1) \tag{34}$$

from Corollary 1 and $T^{-1}\sum_{t=1}^{T}\left[s_t - E_t(q_{t+h})\right]\left[s_t - E_t(q_{t+h})\right]' \geq 0$. Now (34) implies that

$$MSFE^{BLPi} - MSFE^{OBLPi} \geq o_p(1) \tag{35}$$

where

$$\begin{aligned} m_i' MSFE^{BLP} m_i &= T^{-1}\sum_{t=1}^{T}\left[y_{it+h} - m_i'(\delta + \gamma' x_t) - m_i' s_t\right]\left[y_{it+h} - m_i'(\delta + \gamma' x_t) - m_i' s_t\right]' \\ &= MSFE^{BLPi} \end{aligned}$$

and

$$\begin{aligned} m_i' MSFE^{OBLP} m_i &= T^{-1}\sum_{t=1}^{T}\left[y_{it+h} - m_i' F_t^{OBLP}\right]\left[y_{it+h} - m_i' F_t^{OBLP}\right]' \\ &= MSFE^{OBLPi} \end{aligned}$$

because $m_i' s_t = c_t$ while both s_t and c_t are all arbitrary. ■

Lemma 1: Before proceeding, note that

$$\widehat{q}_{t+h} = \mathrm{K}_{0h} + \mathrm{K}_{1h} W_{\Delta t} + \varepsilon_{t,h} + \widehat{q}_{t+h} - q_{t+h} \tag{36}$$

from (16) where $\widehat{q}_{t+h} \equiv \widehat{\gamma}'(x_{t+h} - x_t) + \widehat{z}_{t+h}$ and

$$\begin{aligned} \widehat{q}_{t+h} - q_{t+h} &= (\widehat{\gamma} - \gamma)'(x_{t+h} - x_t) + \delta - \widehat{\delta} + (\gamma - \widehat{\gamma})' x_{t+h} \\ &= \delta - \widehat{\delta} + (\gamma - \widehat{\gamma})' x_t. \end{aligned} \tag{37}$$

Then, the claimed result holds as

$$\left(\widehat{\mathrm{K}}_{0h}, \widehat{\mathrm{K}}_{1h}\right) = \left(\mathrm{K}_{0h}, \mathrm{K}_{1h}\right)\sum_{t=1}^{T}\begin{pmatrix}1 \\ W_{\Delta t}\end{pmatrix}\begin{pmatrix}1 & \widehat{W}_{\Delta t}'\end{pmatrix}\left(\sum_{t=1}^{T}\begin{bmatrix}1 & \widehat{W}_{\Delta t}' \\ \widehat{W}_{\Delta t} & \widehat{W}_{\Delta t}\widehat{W}_{\Delta t}'\end{bmatrix}\right)^{-1}$$

$$+\sum_{t=1}^{T}\varepsilon_{t,h}\left(1 \quad \widehat{W}_{\Delta t}{}'\right)\left(\sum_{t=1}^{T}\begin{bmatrix}1 & \widehat{W}_{\Delta t}{}' \\ \widehat{W}_{\Delta t} & \widehat{W}_{\Delta t}\widehat{W}_{\Delta t}{}'\end{bmatrix}\right)^{-1}$$

$$+\sum_{t=1}^{T}\left(\widehat{q}_{t+h}-q_{t+h}\right)\left(1 \quad \widehat{W}_{\Delta t}{}'\right)\left(\sum_{t=1}^{T}\begin{bmatrix}1 & \widehat{W}_{\Delta t}{}' \\ \widehat{W}_{\Delta t} & \widehat{W}_{\Delta t}\widehat{W}_{\Delta t}{}'\end{bmatrix}\right)^{-1}$$

[from (18) and (36)]

$$= \left(\mathrm{K}_{0h}, \mathrm{K}_{1h}\right)$$

$$+\left(\mathrm{K}_{0h}, \mathrm{K}_{1h}\right)\left\{\left(\underbrace{T^{-1}\sum_{t=1}^{T}\begin{bmatrix}1 & \widehat{W}_{\Delta t}{}' \\ W_{\Delta t} & W_{\Delta t}\widehat{W}_{\Delta t}{}'\end{bmatrix}}_{O_p(1)}\right)\left(T^{-1}\sum_{t=1}^{T}\begin{bmatrix}1 & \widehat{W}_{\Delta t}{}' \\ \widehat{W}_{\Delta t} & \widehat{W}_{\Delta t}\widehat{W}_{\Delta t}{}'\end{bmatrix}\right)^{-1} - I_{\ell p+1}\right\}$$

$$+\left[\underbrace{T^{-1}\sum_{t=1}^{T}\sum_{j=1}^{h}\gamma_{\perp j}\sum_{i=1}^{j}\Psi^{j-i}\mathrm{e}_{t+i}\left(1 \quad \widehat{W}_{\Delta t}{}'\right)}_{o_p(1)}\right]\left(\underbrace{T^{-1}\sum_{t=1}^{T}\begin{bmatrix}1 & \widehat{W}_{\Delta t}{}' \\ \widehat{W}_{\Delta t} & \widehat{W}_{\Delta t}\widehat{W}_{\Delta t}{}'\end{bmatrix}}_{O_p(1)}\right)^{-1}$$

$$+\left(\underbrace{\delta-\widehat{\delta}}_{o_p(1)}, \underbrace{T\left(\gamma-\widehat{\gamma}\right)'}_{o_p(1)}\right)\left(\underbrace{\begin{bmatrix}T^{-1} & 0 \\ 0 & T^{-2}I_k\end{bmatrix}\sum_{t=1}^{T}\begin{bmatrix}1 & \widehat{W}_{\Delta t}{}' \\ x_t & x_t\widehat{W}_{\Delta t}{}'\end{bmatrix}}_{O_p(1)}\right)\times\left(\underbrace{T^{-1}\sum_{t=1}^{T}\begin{bmatrix}1 & \widehat{W}_{\Delta t}{}' \\ \widehat{W}_{\Delta t} & \widehat{W}_{\Delta t}\widehat{W}_{\Delta t}{}'\end{bmatrix}}_{O_p(1)}\right)^{-1}$$

[from (37) and definition in (12)]

$$= \left(\mathrm{K}_{0h}, \mathrm{K}_{1h}\right) + o_p(1)$$

from Assumption 2 because

$$\left(T^{-1}\sum_{t=1}^{T}\begin{bmatrix}1 & W_{\Delta t}{}' \\ W_{\Delta t} & W_{\Delta t}\widehat{W}_{\Delta t}{}'\end{bmatrix}\right)\left(T^{-1}\sum_{t=1}^{T}\begin{bmatrix}1 & \widehat{W}_{\Delta t}{}' \\ \widehat{W}_{\Delta t} & \widehat{W}_{\Delta t}\widehat{W}_{\Delta t}{}'\end{bmatrix}\right)^{-1} \to_p \quad I_{\ell p+1}$$

using the following facts:

$$T^{-1}\sum_{t=1}^{T}\left(\widehat{W}_{\Delta t}-W_{\Delta t}\right) \tag{38}$$

$$= T^{-1}\sum_{t=1}^{T}\begin{pmatrix}\left[\underset{k\times 1}{0}{}', \left(\underset{k\times r}{\gamma-\widehat{\gamma}}\right)'\left(\underset{r\times 1}{\delta-\widehat{\delta}}+\left(\underset{k\times r}{\gamma-\widehat{\gamma}}\right)'\underset{k\times 1}{x_t}\right)'\right]' \\ \left[0', \left(\gamma-\widehat{\gamma}\right)'\left(\delta-\widehat{\delta}+\left(\gamma-\widehat{\gamma}\right)'x_{t-1}\right)'\right]' \\ \vdots \\ \left[0', \left(\gamma-\widehat{\gamma}\right)'\left(\delta-\widehat{\delta}+\left(\gamma-\widehat{\gamma}\right)'x_{t-p+1}\right)'\right]'\end{pmatrix} \to_p 0,$$

from $w_{\Delta t} = \left(\Delta x_t{}', z_t{}'\right)'$ and $\widehat{z}_t - z_t = \delta - \widehat{\delta} + \left(\gamma - \widehat{\gamma}\right)' x_t$ using Assumption 2 (b) - (i);

$$T^{-1}\sum_{t=1}^{T}\left(\widehat{W}_{\Delta t} - W_{\Delta t}\right)\widehat{W}_{\Delta t}{}' \tag{39}$$

$$= T^{-1}\sum_{t=1}^{T}\underbrace{\begin{pmatrix} \left[0_{k\times1}{}', \left(\widehat{z}_t - z_t\right)'\right]' \\ \left[0_{k\times1}{}', \left(\widehat{z}_{t-1} - z_{t-1}\right)'\right]' \\ \vdots \\ \left[0_{k\times1}{}', \left(\widehat{z}_{t-p+1} - z_{t-p+1}\right)'\right]' \end{pmatrix}}_{\ell p\times1}\underbrace{\left(\Delta x_t{}', \widehat{z}_t{}', \quad \Delta x_{t-1}{}', \widehat{z}_{t-1}{}', \quad \cdots \quad \Delta x_{t-p+1}{}', \widehat{z}_{t-p+1}{}'\right)}_{1\times\ell p} = o_p(1),$$

because

$$T^{-1}\sum_{t=1}^{T}\left[0', \left(\widehat{z}_t - z_t\right)'\right]'\left[\Delta x_t{}', \widehat{z}_t{}'\right]$$
$$= T^{-1}\sum_{t=1}^{T}\begin{pmatrix} 0 \\ \delta - \widehat{\delta} + \left(\gamma - \widehat{\gamma}\right)' x_t \end{pmatrix}\left(\Delta x_t{}', \left[z_t + \delta - \widehat{\delta} + \left(\gamma - \widehat{\gamma}\right)' x_t\right]'\right) = o_p(1)$$

from

$$\underbrace{\left(\gamma - \widehat{\gamma}\right)}_{O_p(T^{-3/2})}\underbrace{T^{-1}\sum_{t=1}^{T}\Delta x_t{}'}_{O_p(1)} + \underbrace{\left(\gamma - \widehat{\gamma}\right)'}_{O_p(T^{-3/2})}\underbrace{T^{-1}\sum_{t=1}^{T}x_t\Delta x_t{}'}_{O_p(T^{1/2})} = o_p(1)$$

and

$$\left(\delta - \widehat{\delta}\right)\underbrace{T^{-1}\sum_{t=1}^{T}z_t{}'}_{O_p(1)} + \underbrace{\left(\delta - \widehat{\delta}\right)\left(\delta - \widehat{\delta}\right)'}_{o_p(1)}$$
$$+ \underbrace{\left(\delta - \widehat{\delta}\right)T^{-1}\sum_{t=1}^{T}x_t{}'\left(\gamma - \widehat{\gamma}\right)}_{o_p(1)} + \underbrace{\left(\gamma - \widehat{\gamma}\right)'}_{O_p(T^{-3/2})}\underbrace{T^{-1}\sum_{t=1}^{T}x_t z_t{}'}_{O_p(T^{1/2})}$$
$$+ \underbrace{\left(\gamma - \widehat{\gamma}\right)' T^{-1}\sum_{t=1}^{T}x_t\left(\delta - \widehat{\delta}\right)'}_{o_p(1)} + \underbrace{\left(\gamma - \widehat{\gamma}\right)'}_{O_p(T^{-3/2})}\underbrace{T^{-1}\sum_{t=1}^{T}x_t x_t{}'}_{O_p(T^{-2})}\underbrace{\left(\gamma - \widehat{\gamma}\right)}_{O_p(T^{-3/2})} = o_p(1)$$

using Assumption 2;[16]

$$T^{-1}\sum_{t=1}^{T}\mathrm{e}_{t+i}\left(1 \quad \widehat{W}_{\Delta t}{}'\right) \to_p 0 \tag{40}$$

from following facts (i) and (ii);

(i) $T^{-1}\sum_{t=1}^{T}\mathrm{e}_{t+i}\left(1 \quad \left(W_{\Delta t} - \widehat{W}_{\Delta t}\right)'\right) \to_p 0,$

using (38) and Assumption 2(c) where $\Psi_i = 0$ for all i, and from a law of large numbers where ε_t is an i.i.d process with a finite variance $\Sigma > 0$, and

16 The same result applies to other time difference terms in a similar way.

(ii) $T^{-1}\sum_{t=1}^{T} \mathrm{e}_{t+i}\begin{pmatrix}1 & W_{\Delta t}{'}\end{pmatrix} \rightarrow_p 0$

from the law of large numbers in White (2001, Exercise 3.77) because $\left\{\mathrm{e}_{t+i}\begin{pmatrix}1 & W_{\Delta t}{'}\end{pmatrix}\right\}$ is a martingale difference sequence from

$$E_t\left[\mathrm{e}_{t+i}\begin{pmatrix}1 & W_{\Delta t}{'}\end{pmatrix}\right] = 0 \text{ and } E\left[\mathrm{e}_{t+i}\begin{pmatrix}1 & W_{\Delta t}{'}\end{pmatrix}\right] = 0, \tag{41}$$

where ε_t is an i.i.d process with a finite variance $\Sigma > 0$.

$$E\left|\mathrm{e}_{t+i}W_{\Delta t}{'}\right|^2 < \left(E\left|\mathrm{e}_{t+i}\right|^4\right)^{1/2}\left(E\left|W_{\Delta t}{'}\right|^4\right)^{1/2} < \infty \tag{42}$$

from $E\left|\varepsilon_t\right|^4$ and $E\left|w_{\Delta t}\right|^4 < \infty$[17] under assumption using Cauchy Schwarz inequality. ■

Theorem 3: Note that

$$\begin{aligned} y_{t+h} - \widehat{F_t^{OBLP}} &= y_{t+h} - F_t^{OBLP} + F_t^{OBLP} - \widehat{F_t^{OBLP}} \\ &= \varepsilon_{t,h} + F_t^{OBLP} - \widehat{F_t^{OBLP}} \end{aligned} \tag{43}$$

where $F_t^{OBLP} = \delta + \gamma' x_t + \mathrm{K}_{0h} + \mathrm{K}_{1h} W_{\Delta t}$ from (12).

Further, note that

$$\begin{aligned} F_t^{OBLP} - \widehat{F_t^{OBLP}} &= \delta - \hat{\delta} + \mathrm{K}_{0h} - \hat{\mathrm{K}}_{0h} + (\gamma - \hat{\gamma})' x_t + \mathrm{K}_{1h}\left(W_{\Delta t} - \widehat{W}_{\Delta t}\right) + \left(\mathrm{K}_{1h} - \hat{\mathrm{K}}_{1h}\right)\widehat{W}_{\Delta t} \\ &= \delta - \hat{\delta} + \mathrm{K}_{0h} - \hat{\mathrm{K}}_{0h} + (\gamma - \hat{\gamma})' x_t + \mathrm{K}_{1h}\begin{pmatrix} 0 \\ z_t - \hat{z}_t \end{pmatrix} + \left(\mathrm{K}_{1h} - \hat{\mathrm{K}}_{1h}\right)\widehat{W}_{\Delta t} \\ &= \delta - \hat{\delta} + \mathrm{K}_{0h} - \hat{\mathrm{K}}_{0h} + (\gamma - \hat{\gamma})' x_t \\ &\quad + \mathrm{K}_{1h,2}\left(\delta - \hat{\delta}\right) + \mathrm{K}_{1h,2}(\gamma - \hat{\gamma})' x_t + \left(\mathrm{K}_{1h} - \hat{\mathrm{K}}_{1h}\right)\widehat{W}_{\Delta t} \\ &= \hat{A} + \left(I_r + \mathrm{K}_{1h,2}\right)(\gamma - \hat{\gamma})' x_t + \left(\mathrm{K}_{1h} - \hat{\mathrm{K}}_{1h}\right)\widehat{W}_{\Delta t} \end{aligned} \tag{44}$$

from following definition;

$$\mathrm{K}_{1h}\begin{pmatrix} 0 \\ z_t - \hat{z}_t \end{pmatrix} \equiv \mathrm{K}_{1h,2}\left(\delta - \hat{\delta}\right) + \mathrm{K}_{1h,2}(\gamma - \hat{\gamma})' x_t$$

for the third equality using $z_t - \hat{z}_t = \left(\delta - \hat{\delta}\right) + (\gamma - \hat{\gamma})' x_t$, where $\hat{A} \equiv \left(I_r + \mathrm{K}_{1h,2}\right)\left(\delta - \hat{\delta}\right) + \mathrm{K}_{0h} - \hat{\mathrm{K}}_{0h}$.

From (43), we may write

$$\begin{aligned} MSFE^{OBLP} &= MSFE^{OBLP} + T^{-1}\sum_{t=1}^{T}\left(F_t^{OBLP} - \widehat{F_t^{OBLP}}\right)\left(F_t^{OBLP} - \widehat{F_t^{OBLP}}\right)' \\ &\quad + T^{-1}\sum_{t=1}^{T}\varepsilon_{t,h}\left(F_t^{OBLP} - \widehat{F_t^{OBLP}}\right)' + T^{-1}\sum_{t=1}^{T}\left(F_t^{OBLP} - \widehat{F_t^{OBLP}}\right)\varepsilon_{t,h}{'}. \end{aligned} \tag{45}$$

Now the claimed result holds, as

$$\widehat{MSFE^{OBLP}} - MSFE^{OBLP} = o_p(1) \tag{46}$$

because

17 In this study, absolute values and inequalities are meant to hold for each element of the corresponding matrices.

$$T^{-1}\sum_{t=1}^{T}\left(F_t^{OBLP}-F_t^{\widehat{OBLP}}\right)\left(F_t^{OBLP}-F_t^{\widehat{OBLP}}\right)'$$

$$= \underbrace{\widehat{A}\widehat{A}'}_{o_p(1)} + \left(I_r+\mathrm{K}_{1h,2}\right)\underbrace{\left(\gamma-\widehat{\gamma}\right)'}_{O_p(T^{-3/2})}\left(\underbrace{T^{-1}\sum_{t=1}^{T}x_tx_t'}_{O_p(T^2)}\right)\underbrace{\left(\gamma-\widehat{\gamma}\right)}_{O_p(T^{-3/2})}\left(I_r+\mathrm{K}_{1h,2}\right)'$$

$$+\underbrace{\left(\mathrm{K}_{1h}-\widehat{\mathrm{K}}_{1h}\right)}_{o_p(1)}\left(\underbrace{T^{-1}\sum_{t=1}^{T}\widehat{W}_{\Delta t}\widehat{W}_{\Delta t}'}_{O_p(1)}\right)\underbrace{\left(\mathrm{K}_{1h}-\widehat{\mathrm{K}}_{1h}\right)'}_{o_p(1)} + \left(I_r+\mathrm{K}_{1h,2}\right)\underbrace{\left(\gamma-\widehat{\gamma}\right)'}_{O_p(T^{-3/2})}\underbrace{T^{-1}\sum_{t=1}^{T}x_t}_{O_p(T)}\underbrace{\widehat{A}'}_{o_p(1)}$$

$$+\underbrace{\widehat{A}}_{o_p(1)}\underbrace{T^{-1}\sum_{t=1}^{T}x_t'}_{O_p(T)}\underbrace{\left(\gamma-\widehat{\gamma}\right)}_{O_p(T^{-3/2})}\left(I_r+\mathrm{K}_{1h,2}\right)' + \underbrace{\left(\mathrm{K}_{1h}-\widehat{\mathrm{K}}_{1h}\right)}_{o_p(1)}\underbrace{T^{-1}\sum_{t=1}^{T}\widehat{W}_{\Delta t}}_{O_p(1)}\underbrace{\widehat{A}'}_{o_p(1)}$$

$$+\underbrace{\widehat{A}}_{o_p(1)}\underbrace{T^{-1}\sum_{t=1}^{T}\widehat{W}_{\Delta t}'}_{O_p(1)}\underbrace{\left(\mathrm{K}_{1h}-\widehat{\mathrm{K}}_{1h}\right)'}_{o_p(1)} + \left(I_r+\mathrm{K}_{1h,2}\right)\underbrace{\left(\gamma-\widehat{\gamma}\right)'}_{O_p(T^{-3/2})}\underbrace{T^{-1}\sum_{t=1}^{T}x_t\widehat{W}_{\Delta t}'}_{O_p(T^{1/2})}\underbrace{\left(\mathrm{K}_{1h}-\widehat{\mathrm{K}}_{1h}\right)'}_{o_p(1)}$$

$$+\underbrace{T^{-1}\sum_{t=1}^{T}\widehat{W}_{\Delta t}x_t'}_{O_p(T^{1/2})}\underbrace{\left(\gamma-\widehat{\gamma}\right)}_{O_p(T^{-3/2})}\left(I_r+\mathrm{K}_{1h,2}\right)' = o_p(1)$$

from (39), (44) and Lemma 1; and

$$T^{-1}\sum_{t=1}^{T}\left(F_t^{OBLP}-F_t^{\widehat{OBLP}}\right)\varepsilon_{t,h}' = T^{-1}\sum_{t=1}^{T}\left(\widehat{A}+\left(I_r+\mathrm{K}_{1h,2}\right)\left(\gamma-\widehat{\gamma}\right)'x_t+\left(\mathrm{K}_{1h}-\widehat{\mathrm{K}}_{1h}\right)\widehat{W}_{\Delta t}\right)\varepsilon_{t,h}'$$

$$=\underbrace{\widehat{A}}_{o_p(1)}\underbrace{T^{-1}\sum_{t=1}^{T}\varepsilon_{t,h}'}_{O_p(1)} + \left(I_r+\mathrm{K}_{1h,2}\right)\underbrace{\left(\gamma-\widehat{\gamma}\right)'}_{O_p(T^{-3/2})}\underbrace{T^{-1}\sum_{t=1}^{T}x_t\varepsilon_{t,h}'}_{O_p(T^{1/2})}$$

$$+\underbrace{\left(\mathrm{K}_{1h}-\widehat{\mathrm{K}}_{1h}\right)}_{o_p(1)}\underbrace{T^{-1}\sum_{t=1}^{T}\widehat{W}_{\Delta t}\varepsilon_{t,h}'}_{O_p(1)} = o_p(1)$$

from (39), (40), and (41), Assumption 2(c), and Lemma 1. ∎

Corollary 2: Note that

$$m_i'MSFE_t^{\widehat{OBLP}}m_i \to_p m_i'MSFE_t^{OBLP}m_i$$

from Theorem 3. So, the claimed result holds. ∎

Appendix A

Simulation Results of MSFEs

	$\lambda = 0.5, \rho = 0.1, \mu = 0.1$					$\lambda = 0.5, \rho = 0.2, \mu = 0.1$				
h	**RWP**	**CIP**	**OBLP**	**OBLP1**	**ECMP**	**RWP**	**CIP**	**OBLP**	**OBLP1**	**ECMP**
1	1.594	1.713	1.358	1.359	1.363	1.635	1.848	1.348	1.391	1.351
2	2.524	2.078	1.988	1.981	2.002	2.56	2.184	1.994	2.022	2.005
3	3.129	2.392	2.384	2.376	2.409	3.221	2.518	2.452	2.469	2.473
4	3.635	2.696	2.726	2.712	2.764	3.706	2.839	2.84	2.829	2.875
5	3.915	2.926	2.979	2.955	3.026	4.08	3.147	3.186	3.167	3.236
6	4.311	3.244	3.311	3.276	3.375	4.502	3.494	3.548	3.523	3.614
7	4.63	3.513	3.564	3.533	3.638	4.915	3.815	3.868	3.85	3.935
8	5.07	3.882	3.928	3.903	4.004	5.203	4.108	4.163	4.148	4.246
9	5.341	4.124	4.174	4.15	4.258	5.542	4.407	4.464	4.451	4.561
10	5.627	4.43	4.492	4.463	4.579	5.819	4.678	4.734	4.721	4.843
11	5.994	4.779	4.819	4.805	4.933	6.189	5.041	5.107	5.093	5.224
12	6.291	5.122	5.168	5.15	5.286	6.459	5.304	5.368	5.351	5.506
13	6.636	5.44	5.486	5.465	5.611	6.829	5.665	5.737	5.721	5.898
14	6.954	5.751	5.787	5.772	5.92	7.17	5.995	6.07	6.044	6.235
15	7.246	6.037	6.08	6.065	6.234	7.44	6.304	6.368	6.345	6.551
16	7.63	6.476	6.515	6.503	6.69	7.974	6.81	6.875	6.85	7.058
17	8.03	6.859	6.901	6.893	7.116	8.246	7.108	7.162	7.141	7.359
18	8.313	7.16	7.208	7.197	7.444	8.561	7.389	7.451	7.432	7.64
19	8.625	7.531	7.575	7.561	7.811	8.877	7.697	7.757	7.731	7.981
20	8.901	7.835	7.876	7.861	8.152	9.228	8.035	8.091	8.062	8.33

	$\lambda = 0.5, \rho = 0.3, \mu = 0.1$					$\lambda = 0.5, \rho = 0.4, \mu = 0.1$				
h	**RWP**	**CIP**	**OBLP**	**OBLP1**	**ECMP**	**RWP**	**CIP**	**OBLP**	**OBLP1**	**ECMP**
1	1.651	2.173	1.342	1.471	1.348	1.704	3.172	1.357	1.622	1.36
2	2.526	2.519	2.013	2.151	2.023	2.518	3.673	2.154	2.429	2.15
3	3.212	2.95	2.64	2.732	2.659	3.22	4.136	2.923	3.156	2.923
4	3.766	3.265	3.105	3.156	3.138	3.967	4.634	3.71	3.884	3.714
5	4.301	3.626	3.573	3.587	3.629	4.626	5.08	4.41	4.535	4.416
6	4.762	3.962	3.967	3.966	4.039	5.156	5.404	4.955	5.025	4.969
7	5.312	4.36	4.404	4.386	4.484	5.751	5.803	5.527	5.564	5.561
8	5.778	4.719	4.796	4.769	4.893	6.399	6.254	6.125	6.128	6.174
9	6.189	5.023	5.101	5.078	5.214	6.984	6.642	6.635	6.608	6.714
10	6.518	5.321	5.418	5.388	5.535	7.465	6.947	7.018	6.981	7.088
11	6.842	5.622	5.724	5.688	5.862	8.014	7.375	7.508	7.46	7.612
12	7.319	6.06	6.159	6.129	6.3	8.578	7.91	8.093	8.044	8.205
13	7.724	6.409	6.51	6.489	6.65	8.969	8.295	8.545	8.461	8.657
14	8.043	6.747	6.849	6.831	7.006	9.378	8.647	8.92	8.833	9.056
15	8.374	7.112	7.218	7.193	7.381	9.928	9.11	9.402	9.317	9.55
16	8.764	7.414	7.52	7.491	7.689	10.274	9.397	9.699	9.616	9.9
17	9.206	7.878	8.005	7.976	8.188	10.854	9.904	10.236	10.139	10.435
18	9.693	8.346	8.471	8.446	8.664	11.295	10.311	10.637	10.545	10.869
19	10.054	8.69	8.809	8.779	9.016	11.813	10.775	11.109	11.016	11.378
20	10.542	9.157	9.27	9.241	9.52	12.396	11.261	11.582	11.493	11.904

	$\lambda = 0.5$, $\rho = 0.1$, $\mu = 0.5$					$\lambda = 0.5$, $\rho = 0.2$, $\mu = 0.5$				
h	**RWP**	**CIP**	**OBLP**	**OBLP1**	**ECMP**	**RWP**	**CIP**	**OBLP**	**OBLP1**	**ECMP**
1	1.643	1.727	1.381	1.385	1.387	1.61	1.847	1.328	1.384	1.329
2	2.606	2.101	2.02	2.016	2.036	2.478	2.205	1.975	2.026	1.987
3	3.084	2.341	2.353	2.333	2.376	3.191	2.598	2.522	2.55	2.554
4	3.549	2.601	2.635	2.622	2.67	3.554	2.84	2.827	2.837	2.875
5	3.965	2.924	2.957	2.941	3.006	4.035	3.154	3.177	3.173	3.24
6	4.244	3.226	3.265	3.248	3.324	4.452	3.499	3.55	3.537	3.63
7	4.565	3.498	3.548	3.525	3.608	4.824	3.826	3.896	3.871	3.987
8	4.936	3.818	3.869	3.846	3.937	5.155	4.072	4.145	4.111	4.279
9	5.305	4.19	4.239	4.221	4.318	5.566	4.449	4.527	4.496	4.704
10	5.576	4.509	4.558	4.54	4.646	5.944	4.799	4.863	4.842	5.074
11	5.938	4.84	4.899	4.875	4.982	6.171	5.003	5.068	5.044	5.351
12	6.247	5.175	5.23	5.208	5.328	6.403	5.281	5.363	5.337	5.688
13	6.691	5.567	5.612	5.598	5.723	6.826	5.705	5.776	5.753	6.154
14	7.095	5.955	6.007	5.981	6.132	7.243	6.074	6.152	6.122	6.662
15	7.409	6.242	6.277	6.261	6.435	7.662	6.479	6.557	6.53	7.178
16	7.661	6.539	6.575	6.562	6.735	7.991	6.791	6.869	6.842	7.593
17	8.03	6.847	6.887	6.878	7.051	8.273	7.086	7.157	7.131	7.947
18	8.437	7.24	7.274	7.27	7.448	8.671	7.495	7.566	7.541	8.509
19	8.846	7.668	7.726	7.709	7.925	8.945	7.775	7.848	7.822	8.963
20	9.203	8.048	8.111	8.089	8.36	9.375	8.251	8.325	8.293	9.621

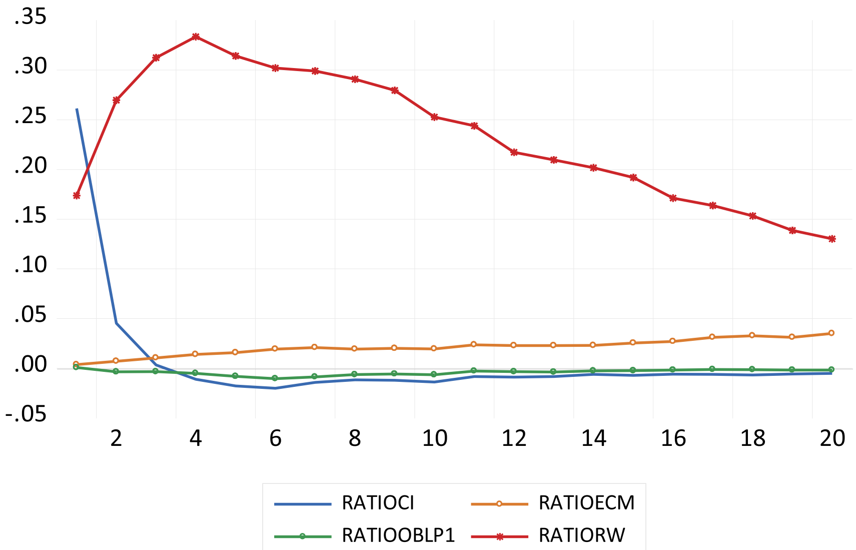
	$\lambda = 0.5$, $\rho = 0.3$, $\mu = 0.5$					$\lambda = 0.5$, $\rho = 0.4$, $\mu = 0.5$				
h	**RWP**	**CIP**	**OBLP**	**OBLP1**	**ECMP**	**RWP**	**CIP**	**OBLP**	**OBLP1**	**ECMP**
1	1.653	2.196	1.35	1.488	1.356	1.711	3.202	1.368	1.635	1.373
2	2.564	2.603	2.086	2.219	2.108	2.56	3.684	2.156	2.458	2.165
3	3.277	3.019	2.707	2.806	2.728	3.266	4.243	2.973	3.226	2.991
4	3.77	3.367	3.198	3.246	3.238	3.947	4.619	3.684	3.872	3.73
5	4.296	3.653	3.604	3.614	3.674	4.625	5.077	4.411	4.537	4.485
6	4.749	3.954	3.973	3.964	4.058	5.17	5.575	5.102	5.185	5.184
7	5.229	4.378	4.435	4.41	4.545	5.77	6.013	5.722	5.756	5.858
8	5.645	4.745	4.839	4.799	4.972	6.355	6.433	6.298	6.307	6.511
9	6.046	5.085	5.214	5.163	5.374	6.925	6.889	6.913	6.874	7.174
10	6.383	5.37	5.511	5.466	5.736	7.331	7.245	7.366	7.306	7.707
11	6.744	5.715	5.868	5.822	6.147	7.887	7.724	7.941	7.861	8.374
12	7.166	6.028	6.175	6.127	6.543	8.388	8.205	8.503	8.412	9.066
13	7.494	6.338	6.479	6.434	6.941	8.91	8.622	8.981	8.874	9.72
14	7.945	6.734	6.878	6.832	7.445	9.471	9.07	9.466	9.345	10.337
15	8.264	7.058	7.187	7.144	7.926	10.086	9.623	10.041	9.92	11.037
16	8.625	7.414	7.538	7.499	8.365	10.429	9.947	10.376	10.261	11.525
17	8.908	7.694	7.8	7.768	8.722	10.887	10.371	10.818	10.701	12.222
18	9.273	8.047	8.137	8.125	9.218	11.392	10.845	11.294	11.182	12.832
19	9.595	8.405	8.503	8.478	9.756	11.712	11.165	11.622	11.507	13.362
20	10.057	8.876	8.993	8.957	10.469	12.091	11.557	12.029	11.917	13.986

	$\lambda = 0.7$, $\rho = 0.1$, $\mu = 0.1$					$\lambda = 0.7$, $\rho = 0.2$, $\mu = 0.1$				
h	**RWP**	**CIP**	**OBLP**	**OBLP1**	**ECMP**	**RWP**	**CIP**	**OBLP**	**OBLP1**	**ECMP**
1	1.426	2.396	1.31	1.318	1.314	1.461	3.336	1.354	1.442	1.358
2	2.575	2.827	2.237	2.25	2.257	2.545	3.853	2.311	2.493	2.315
3	3.493	3.265	2.952	2.964	2.989	3.425	4.296	3.145	3.329	3.165
4	4.212	3.7	3.569	3.557	3.618	4.186	4.776	3.904	4.071	3.931
5	4.77	4.015	3.99	3.972	4.054	4.905	5.229	4.586	4.735	4.624
6	5.285	4.363	4.409	4.387	4.482	5.561	5.658	5.2	5.325	5.242
7	5.773	4.719	4.796	4.773	4.879	6.093	6.054	5.762	5.85	5.814
8	6.224	5.018	5.13	5.095	5.234	6.67	6.452	6.333	6.363	6.412
9	6.663	5.397	5.541	5.494	5.656	7.223	6.872	6.864	6.864	6.962
10	7.005	5.734	5.896	5.843	6.038	7.646	7.159	7.238	7.215	7.365
11	7.393	6.064	6.233	6.178	6.385	8.151	7.566	7.735	7.681	7.872
12	7.809	6.427	6.595	6.541	6.777	8.589	7.966	8.215	8.131	8.391
13	8.247	6.829	6.992	6.936	7.196	9.101	8.416	8.707	8.611	8.9
14	8.602	7.133	7.299	7.241	7.515	9.668	8.867	9.197	9.096	9.402
15	8.959	7.495	7.651	7.594	7.893	10.254	9.345	9.673	9.583	9.885
16	9.354	7.926	8.106	8.044	8.373	10.639	9.635	9.976	9.878	10.231
17	9.677	8.271	8.451	8.388	8.747	11.14	10.055	10.419	10.31	10.726
18	10.117	8.684	8.869	8.797	9.182	11.586	10.466	10.859	10.737	11.196
19	10.399	8.943	9.112	9.05	9.467	12.076	10.956	11.365	11.241	11.747
20	10.793	9.334	9.473	9.429	9.854	12.528	11.366	11.786	11.658	12.168

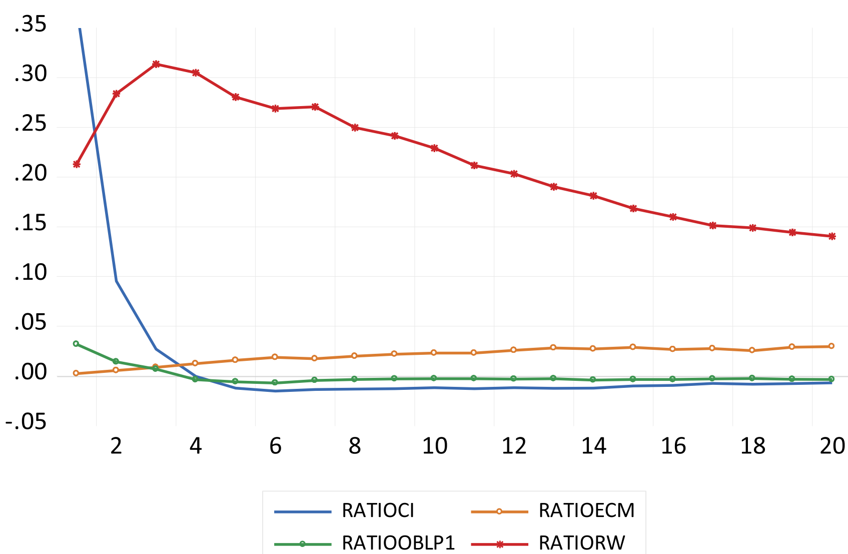
	$\lambda = 0.7$, $\rho = 0.1$, $\mu = 0.5$					$\lambda = 0.7$, $\rho = 0.2$, $\mu = 0.5$				
h	**RWP**	**CIP**	**OBLP**	**OBLP1**	**ECMP**	**RWP**	**CIP**	**OBLP**	**OBLP1**	**ECMP**
1	1.467	2.525	1.371	1.385	1.377	1.419	2.435	1.325	1.334	1.331
2	2.592	2.965	2.33	2.336	2.348	2.507	2.906	2.272	2.285	2.285
3	3.463	3.359	3.029	3.027	3.072	3.351	3.303	2.981	2.976	3.004
4	4.135	3.756	3.587	3.582	3.657	4.049	3.678	3.543	3.529	3.579
5	4.687	4.067	4.023	4.007	4.117	4.766	4.149	4.13	4.111	4.194
6	5.303	4.476	4.509	4.481	4.635	5.215	4.481	4.539	4.509	4.648
7	5.816	4.82	4.897	4.87	5.093	5.668	4.813	4.928	4.89	5.074
8	6.305	5.206	5.322	5.284	5.574	6.05	5.144	5.27	5.233	5.479
9	6.755	5.583	5.721	5.686	6.029	6.585	5.533	5.686	5.637	5.967
10	7.1	5.869	6.031	5.98	6.395	7.06	5.965	6.131	6.078	6.503
11	7.418	6.175	6.361	6.303	6.785	7.346	6.287	6.464	6.404	6.9
12	7.82	6.549	6.738	6.682	7.259	7.708	6.652	6.817	6.764	7.346
13	8.047	6.783	6.972	6.925	7.581	8.119	6.951	7.103	7.056	7.728
14	8.338	7.051	7.256	7.196	8.001	8.446	7.267	7.417	7.369	8.167
15	8.839	7.457	7.652	7.598	8.56	8.764	7.557	7.705	7.664	8.604
16	9.309	7.92	8.111	8.059	9.184	9.124	7.91	8.087	8.035	9.129
17	9.715	8.319	8.505	8.449	9.764	9.496	8.261	8.439	8.383	9.653
18	10.118	8.761	8.95	8.893	10.448	9.892	8.659	8.829	8.78	10.226
19	10.434	9.086	9.259	9.207	10.968	10.192	8.92	9.102	9.042	10.731
20	10.713	9.386	9.556	9.496	11.556	10.519	9.244	9.435	9.364	11.307

Appendix B

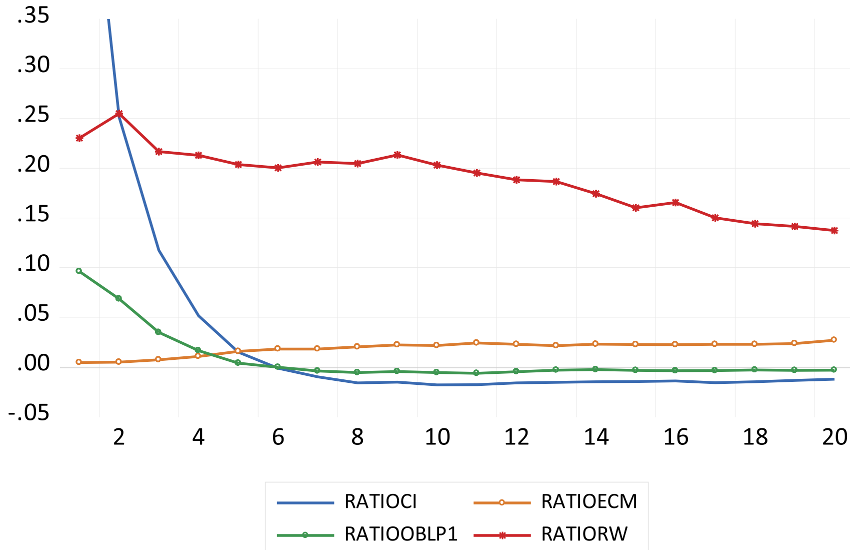
Graphs of MSFE Ratio Comparison



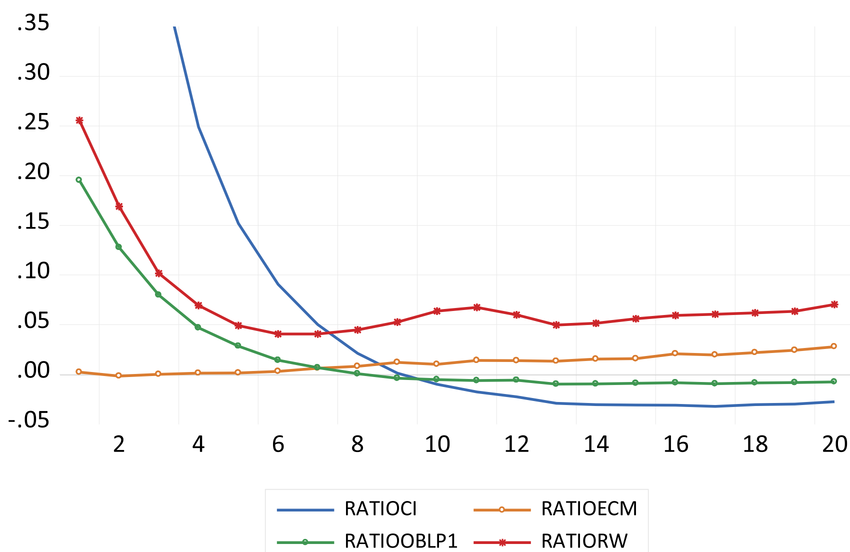
$\rho = 0.1$ and $\mu = 0.1$



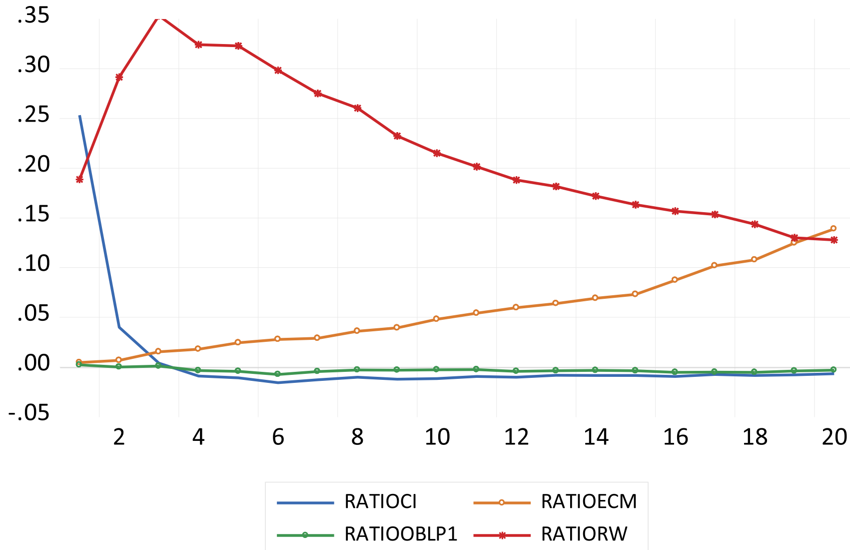
$\rho = 0.2$ and $\mu = 0.1$



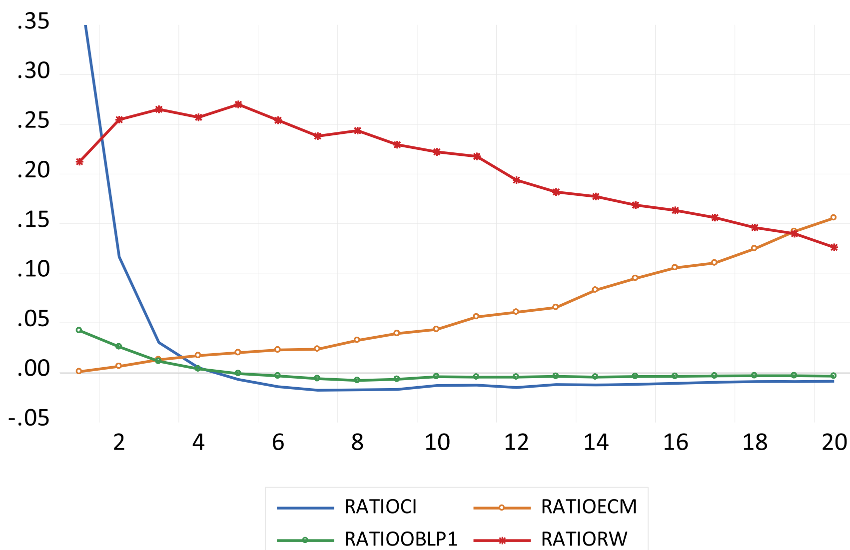
$\rho = 0.3$ and $\mu = 0.1$



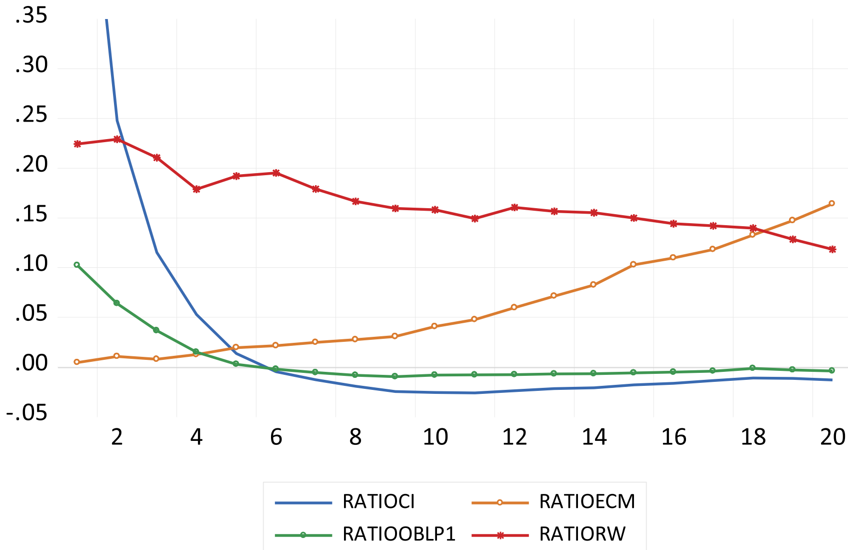
$\rho = 0.4$ and $\mu = 0.1$



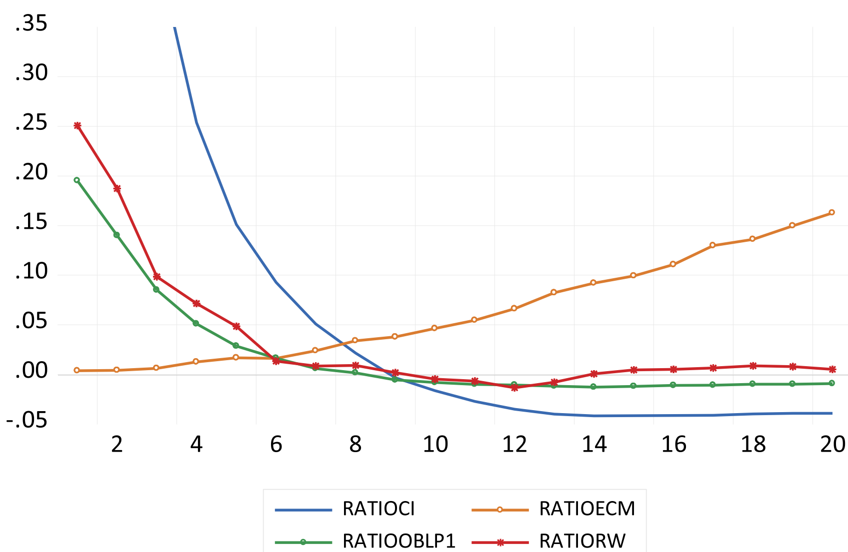
$\rho = 0.1$ and $\mu = 0.5$



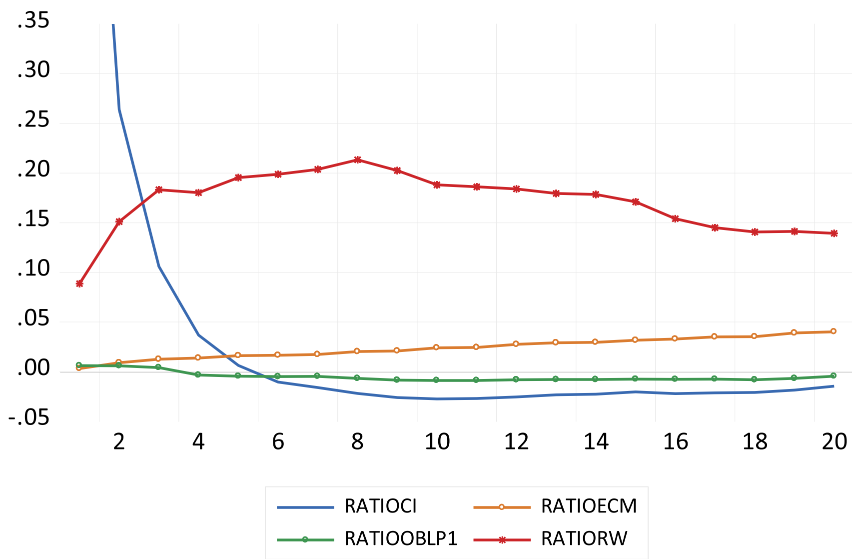
$\rho = 0.2$ and $\mu = 0.5$



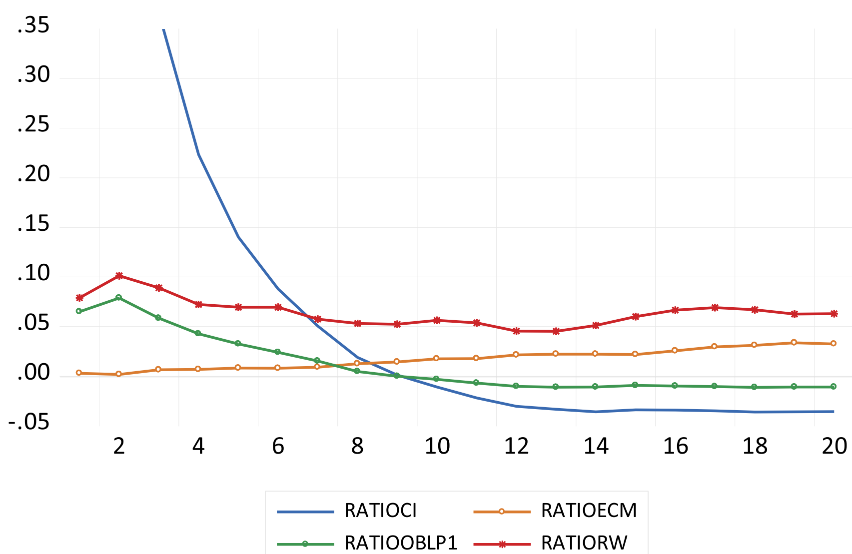
$\rho = 0.3$ and $\mu = 0.5$



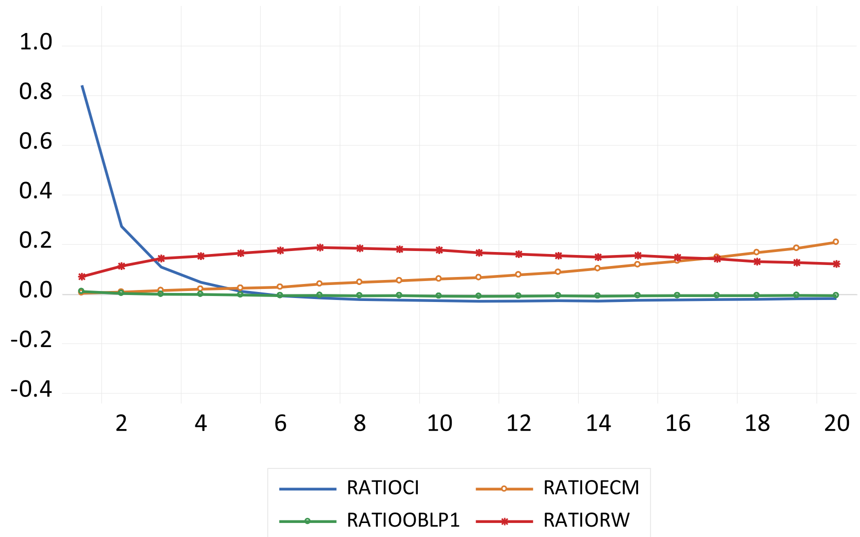
$\rho = 0.4$ and $\mu = 0.5$



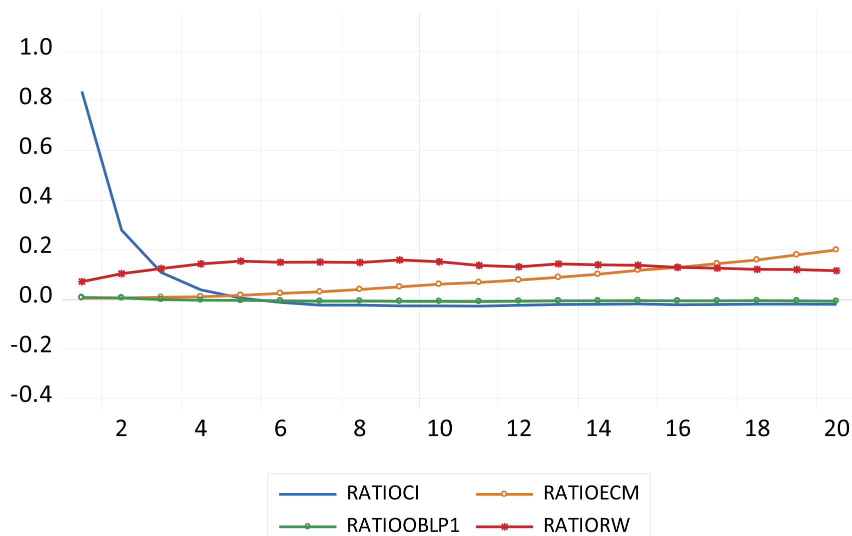
$\rho = 0.1$ and $\mu = 0.1$



$\rho = 0.2$ and $\mu = 0.1$



$\rho = 0.1$ and $\mu = 0.5$



$\rho = 0.2$ and $\mu = 0.5$

Appendix C

Comparison of Prediction Errors of the Predictors for the US GDP and Consumption

	GDP				
h	RWP	CIP	OBLP	OBLP1	ECMP
1	3.26E-05	0.002542909	5.08E-05	8.62E-06	4.21E-05
2	0.000212488	0.001727403	0.000297154	9.37E-05	0.000238805
3	0.000772134	0.00080382	0.00091498	0.000442591	0.000869382
4	0.001777196	0.000195502	0.00206539	0.001095312	0.001979646
5	0.002673287	1.97E-05	0.002900983	0.001630236	0.00279066
6	0.001825035	0.000180058	0.001757161	0.000829912	0.001920786
7	0.001893291	0.009930316	0.002031432	0.003550588	0.001862645
8	0.001600361	0.000260322	0.001333017	0.000450299	0.001426997
9	0.003279124	1.26E-06	0.00266211	0.001287562	0.002565353
10	0.006899988	0.000725072	0.005630794	0.00348158	0.005258846
11	0.012821841	0.003259789	0.010435198	0.007528494	0.009704139
12	0.018498792	0.006379417	0.015333165	0.0114428	0.01397476
13	0.02892438	0.012980655	0.024204309	0.019219619	0.022716846
14	0.03427973	0.016643283	0.027874189	0.022960305	0.028990529
15	0.04223897	0.022315043	0.033952655	0.028682749	0.03781501
16	0.049728285	0.027842055	0.040176917	0.034099237	0.044824833
17	0.057020857	0.033361531	0.045941975	0.03940116	0.053326852
18	0.064524805	0.03915583	0.051032039	0.044872106	0.065428491
19	0.069314251	0.042905731	0.053188094	0.047978423	0.074835721
20	0.08025364	0.051597906	0.060331916	0.056535873	0.092365352
	Consumption				
h	RWP	CIP	OBLP	OBLP1	ECMP
1	4.77E-05	0.00398025	3.76E-05	1.54E-05	5.90E-05
2	0.000107305	0.003556216	9.51E-05	2.61E-05	0.000126229
3	0.000557568	0.0021511	0.000462535	0.000266604	0.00064064
4	0.001307121	0.001145057	0.001084896	0.000707384	0.001481522
5	0.002157744	0.000554193	0.00171602	0.00118166	0.002263322
6	0.00109426	0.001362586	0.0006129	0.000331722	0.001168679
7	0.003764306	0.017251972	0.005047532	0.006157309	0.003721043
8	0.001018033	0.001450559	0.000419959	0.000136812	0.000880767
9	0.002544972	0.000382011	0.001412778	0.000748579	0.001921366
10	0.006929691	0.000175614	0.004658435	0.0032792	0.005284781
11	0.016977752	0.003636799	0.012544745	0.010286453	0.013357517
12	0.022811909	0.006568016	0.017395848	0.014238673	0.017753095
13	0.031429434	0.011511286	0.02403677	0.020307374	0.024942795
14	0.038337336	0.01582725	0.028578247	0.025032869	0.032731111
15	0.047577724	0.021942604	0.035141204	0.031422496	0.042874999
16	0.054502229	0.026720578	0.040448571	0.0360104	0.049362836
17	0.060781599	0.031168638	0.044794872	0.040078619	0.056965715
18	0.070767806	0.038427513	0.051584102	0.047050077	0.071714052
19	0.075168454	0.041687772	0.053291281	0.049389318	0.080913991
20	0.083066152	0.047619605	0.057283969	0.054908134	0.095380875

Data Description

GDP:	Gross Domestic Product, Billions of Dollars, Quarterly, Seasonally Adjusted Annual Rate.
Consumption:	Personal Consumption Expenditures, Billions of Dollars, Quarterly, Seasonally Adjusted Annual Rate.
Net exports:	Net Exports of Goods and Services, Billions of Dollars, Quarterly, Seasonally Adjusted Annual Rate.
M1:	M1 for the United States, National Currency, Quarterly, Seasonally Adjusted.
Term Spread:	10-Year Treasury Constant Maturity Minus 2-Year Treasury Constant (or Federal Fund Rate). Maturity (%), Quarterly, Not Seasonally Adjusted.
Government Expenditure:	Federal Government: Current Expenditures, Billions of Dollars, Quarterly, Seasonally Adjusted Annual Rate.

References

Campbell, John Y., and Robert J. Shiller. 1987. "Cointegration and Tests of Present Value Models." *The Journal of Political Economy* 95 (5): 1062–88.

Christoffersen, Peter F., and Francis X. Diebold. 1998. "Cointegration and Long-Horizon Forecasting." *Journal of Business & Economic Statistics* 16 (4): 450–8.

Elliott, Graham. 2006. "Forecasting with Trending Data." In *Handbook of Economic Forecasting*, Vol. 1, edited by G. Elliott, C. W. J. Granger, and A. Timmermann, 555–604. Elsevier.

Elliott, Graham, Thomas J. Rothenberg, and James H. Stock. 1996. "Efficient Tests for an Autoregressive Unit Root." *Econometrica* 64 (4): 813–6.

Engle, Rovert F., and Byung Sam Yoo. 1987. "Forecasting and Testing in Co-integrated Systems." *Journal of Econometrics* 35 (1): 143–59.

Estrella, Arturo, and Gikas A. Hardouvelis. 1991. "The Term Structure as a Predictor of Real Economic Activity." *Journal of Finance* 46 (2): 555–76.

Estrella, Arturo, and Frederic S. Mishkin. 1996. "The Yield Curve as a Predictor of US Recessions." *Current Issues in Economics and Finance* 2 (7). FRB of New York.

Estrella, Arturo, and Frederic S. Mishkin. 1998. "Predicting US Recessions: Financial Variables as Leading Indicators." *Review of Economics and Statistics* 80 (1): 45–61.

Hamilton, James D. 1994. *Time Series Analysis*. Princeton University Press.

Johansen, Søren. 1991. "Estimation and Hypothesis Testing of Co-integration Vectors in Gaussian Vector Autoregressive Models." *Econometrica* 59 (6): 1551–80.

Kim, Yun-Yeong. 2012. "Stationary Vector Autoregressive Representation of Error Correction Models." *Theoretical Economics Letters* 2 (2): 152–6.

Kim, Yun-Yeong. 2018. "Does Monetary Policy Affect Long-Run Expectation of Non-stationary Real Interest Rates?" *Applied Economics* 50 (12): 1342–61.

Kim, Yun-Yeong. 2023. "Best Linear Prediction in Cointegrated Systems." *Advances in Econometrics, Essays in Honor of Joon Y. Park: Econometric Theory* 45A: 367–91.

Kishor, Kundan N., and Evan F. Koenig. 2010. "Yield-spreads as Predictors of Economic Activity: A Real Time VAR Analysis." FRB of Dallas Working Paper. https://doi.org/10.24149/wp1008.

Litterman, Robert B. 1986. "Forecasting with Bayesian Vector Autoregressions — Five Years of Experience." *Journal of Business and Economic Statistics* 4 (1): 25–38.

MacKinnon, James G., Alfred A. Haug, and Leo Michelis. 1999. "Numerical Distribution Functions of Likelihood Ratio Tests for Cointegration." *Journal of Applied Econometrics* 14 (5): 563–77, https://doi.org/10.1002/(sici)1099-1255(199909/10)14:5<563::aid-jae530>3.3.co;2-i.

Phillips, P. C. B. 1991. "Optimal Inference in Cointegrated Systems." *Econometrica* 59 (2): 283–306.

White, H. 2001. *Asymptotic Theory for Econometricians (Revised Edition)*. Cambridge, MA: Academic Press.

Supplementary Material: This article contains supplementary material (https://doi.org/10.1515/snde-2024-0107).