Uncovering Bias in Uncovered Interest Parity Tests*

Emilio González-Coya†
Morgan Stanley

Pierre Perron‡
Boston University

September 7, 2024

Abstract

We re-examine the Uncovered Interest Parity (UIP) hypothesis, which posits efficiency in forward foreign exchange and rational expectations. Testing these assumptions involves estimating parameters in a k-step-ahead forecasting model, where forecast errors are expected to be serially correlated up to lags $k-1$. When errors are correlated beyond these lags, OLS is not longer consistent unless the regressors are exogenous. We extend the FGLS procedure developed in Perron and González-Coya (2024) to a setting in which lagged dependent variables are included as regressors. We thus provide a consistent and efficient framework to estimate the parameters of a general k-step-ahead linear forecasting equation. Following the work of Olivari and Perron (2024), we introduce an instrumental variable (IV)-based approach for this problem that requires pre-determined but not necessarily exogenous IVs for consistency. We apply our FGLS procedures to the analysis of the two main specifications to test the UIP. Contrary to most empirical results available in the literature, in particular those based on some OLS regression or GMM, our robust and efficient procedure cannot reject the null hypothesis that the UIP holds. Hence, overall, our results can be viewed as overturning the so-called forward discount anomaly.

Keywords: Feasible Generalized Least-Squares, forward premium anomaly, rational expectations, efficient markets, robust standard errors.

JEL Codes: C22, C31.

*We thank Enrique Sentana and Marco Oilivari for useful discussions.
†Morgan Stanley, 1585 Broadway Avenue, New York, NY 10036 (emilio.gonzalez.coya.sandoval @morganstanley.com).
‡Department of Economics, Boston University, 270 Bay State Rd., Boston, MA, 02215 (perron@bu.edu).

1 Introduction

In this paper, we re-examine the hypothesis of Uncovered Interest Parity (UIP), which in its basic form implies that the (nominal) expected return to speculation in the forward foreign exchange market conditional on available information should be zero. This is an "efficient-markets hypothesis" (EMH) for foreign exchange markets: if all available information is used rationally by risk-neutral agents in determining the spot and forward exchange rates, then the expected rate of return to speculation will be zero and the foreign exchange market is said to be efficient. This is a joint hypothesis since it includes the assumption of rational expectations (REH) and the assumption that the risk premium for the forward rate is zero. In fact, rejection of the UIP hypothesis does not immediately translates into a rejection of the efficiency of the foreign exchange market, which could be due economic agents being risk averse. Still testing whether the UIP hypothesis holds has been and continue to be a topic of considerable interest from both theoretical and empirical perspectives.

Testing the rationality hypothesis and exchange market efficiency is embedded in the general problem of estimating the parameters of a k-step-ahead linear forecasting equation. When the sampling interval is finer than the interval over which forecasts are made (in this case the maturity time of the forward exchanges rates), the forecast error is serially correlated. As noted by Hansen and Hodrick (1980), under rational expectations (REH) the forecast error is serially correlated up to lag $k-1$ and OLS remains consistent but appropriate modifications in the estimation of the asymptotic covariance matrix are needed. However, if the REH is rejected and the forecast error is serially correlated beyond lag $k-1$, OLS is not longer consistent when the regressors are not exogenous; see Perron and González-Coya (2024). Contrary to what is asserted in Hansen and Hodrick (1980), GLS is consistent when the regressors are pre-determined provided the roots of the MA polynomial are inside the unit circle, i.e., MA process is invertible, as shown in Perron and González-Coya (2024). Moreover, GLS remains consistent when the forecast error follows a linear invertible process.

The first contribution of this paper is to provide a consistent and efficient framework to estimate and perform tests of the parameters of a k-step-ahead linear forecasting equation that remains valid whether the REH holds or not. We apply the FGLS procedure developed in Perron and González-Coya (2024) which is consistent using non-exogenous regressors, provided the errors follows an stationary invertible linear process. The second contribution is to extend their FGLS procedure to cases with lagged dependent variables included as regressors. Following the work of Olivari and Perron (2024), we introduce an instrumental variable

(IV)-based approach for this problem that requires pre-determined but not necessarily exogenous IVs for consistency. The third contribution is to apply our FGLS procedures to the two main regressions suggested in the literature to test the UIP. We use 30 years of data for three currencies and we reconsider the framework and regressions used by Fama (1984) and Hansen and Hodrick (1980). We provide extensive simulation experiments to assess the finite sample performance of our FGLS procedure relative to OLS. We show that FGLS achieves important reductions in mean-squared error (MSE) and allow tests with much greater power.

The Fama regression assesses whether the current forward-spot differential, $f_{t,h} - s_t$, is a good predictor of the future change in the spot rate, $s_{t+h} - s_t$. Most results available in the literature suggest a negative estimate of the relevant parameter, which instead should take value one if the UIP holds. This is often referred to as the "forward discount anomaly", wich refers to the widespread empirical finding that the returns on nominal exchange rates is negatively correlated with the lagged forward premium. It implies an appreciating currency for the high interest rate country. It is an "anomaly" as rational expectations would imply the opposite; if all currencies are equally risky, investors would demand higher interest rates on currencies expected to fall in value. Our results, in contrast, indicate positive values, sometimes not significantly different from one. This finding suggests that the "forward discount anomaly" might be a consequence of OLS providing an inconsistent estimate and our FGLS procedure being consistent and efficient under a broader range of possible scenarios.

The regression adopted by Hansen and Hodrick (1980) is to test whether past values of the forward-spot differential help predict the current value $f_{t,h} - s_t$, conditioning on some covariates involving the past forward-spot differentials from some other countries. Under the UIP and EMH, there should be no predictive power as all information contained in the information set at time t should already have been accounted for by the market in setting the forward rates. Here, contrary to most empirical results available in the literature, in particular those based on some OLS regression or GMM, our robust and efficient procedure cannot reject the null hypothesis that the UIP holds. Hence, overall, our results can be viewed as overturning the so-called "forward discount anomaly".

The remainder of this paper is as follows. Section 2 describes the formulations of the UIP hypothesis and reviews previous econometric tests. Section 3 describes the FGLS procedures and Section 4 introduces two instrumental variable (IV) based approaches for a setting in which lagged dependent variables are included as regressors. Section 5 studies the Fama (1984) regression and Section 6 is focused on the Hansen and Hodrick (1980) regression. For both specifications, we provide extensive Monte Carlo experiments to asses the finite sample performance of our FGLS procedures relative to OLS. We also provide empirical results for the sample period 01/1993 to 01/2023. In Section 7 we present a test for the hypothesis that the OLS residuals exhibits serial correlation of order greater than $k-1$ and discuss its empirical implications to understand the various conflicting results. We provide brief concluding remarks in Section 8. A supplement provide additional details.

2 UIP and the Efficient Market Hypothesis

Let $s_t = \ln(S_t)$ and $f_{t,k} = \ln(F_{t,k})$, where S_t and $F_{t,k}$ are the levels of the spot exchange rate and the k-period forward exchange rate at time t. With i_t, the domestic nominal interest rate and i_t^* the corresponding foreign interest rate, the theory of Uncovered Interest Parity (UIP) implies that $E\left(s_{t+k} - s_t|\Phi_t\right) = (i_t - i_t^*)$. Hence, UIP requires the twin assumptions of rational expectations and a constant or zero risk premium. Given the no arbitrage condition, the Covered Interest Parity (CIP) condition implies that $(i_t - i_t^*) = (f_{t,k} - s_t)$ holds as an identity. Hence, the UIP condition is also frequently expressed as $E\left(s_{t+k} - s_t|\Phi_t\right) = (f_{t,k} - s_t)$. Since $s_{t+k} - f_{t,k}$ is an approximate measure of the rate of return to speculation, we can express the efficient-markets hypothesis as $f_{t,k} = E\left(s_{t+k}|\Phi_t\right)$. This implies forecast errors $s_{t+k} - f_{t,k}$ uncorrelated with information available at time t, Φ_t.

2.1 Econometric Tests

As in Hansen and Hodrick (1980) we consider the general problem of estimating the parameters of a k-step-ahead linear forecasting equation, $E\left(y_{t+k}|\Phi_t\right) = x_t'\beta$. Then,

$$y_{t+k} = \alpha + x_t'\beta + u_{t+k}, \tag{1}$$

where rational expectations impose a specific structure on the forecast error $u_{t+k} = y_{t+k} - E\left(y_{t+k}|\Phi_t\right)$. Due to the $k-1$ period overlap in the sequential k-step-ahead forecasts, u_{t+k} has an $MA(k-1)$ representation. Thus

$$E\left(x_t u_{t+k}\right) = E\left(x_t \textstyle\sum_{j=0}^{k-1} \theta_j \varepsilon_{t+k-j}\right) = 0$$

and the OLS estimates of α, β are consistent since the regressors are pre-determined via the rational expectations hypothesis. We shall label this as the “RE case”. Note, however, that we could well be faced with a model of the form

$$y_{t+k} = x_t'\beta + (u_{t+k} + \eta_t),$$

where η_t is some serially correlated process involving innovations dated before period t; e.g., an $AR(1)$ process of the form $\eta_t = \rho\eta_{t-1} + e_t$ for some sequence of i.i.d. innovations $(e_1, ..., e_T)$. In this case, if the regressors are non-exogenous with respect to past values of η_t, OLS is not longer consistent since $E(x_t\eta_t) \neq 0$. We shall label this as the "general case", as it encompasses the "RE case". This motivates the necessity of an estimate that is consistent under the both the "RE case" and the "general case", i.e., allowing the errors to be serially correlated beyond lag $k-1$. Contrary to what is asserted in Hansen and Hodrick (1980), GLS is consistent in both cases, provided the errors $(u_{t+k} + \eta_t)$ can be represented as some invertible linear process, as shown in Perron and González-Coya (2024).

Example 1. *A k-step ahead forecast error process can be serially correlated beyond lags $k-1$ under, e.g., Adaptive Expectations (AE) (see Muth, 1960). An exponentially weighted moving average forecast arises from the following model of expectations adapting to changing conditions. Under AE, it is assumed that the forecast is changed from one period to the next by an amount proportional to the latest observed error,*

$$y_t^e = y_{t-1}^e + \beta\left(y_{t-1} - y_{t-1}^e\right).$$

As shown in Muth (1960), the solution of this difference equation is an exponentially weighted forecast $y_t^e = \beta\sum_{i=1}^{\infty}(1-\beta)^{i-1}y_{t-i}$. The forecast error is thus,

$$u_t = y_t - y_t^e = y_t - \beta\sum_{i=1}^{\infty}(1-\beta)^{i-1}y_{t-i}.$$

In order to characterize the process of the forecast error u_t, we shall impose a functional form on y_t. It is standard in the AE literature, to assume that y_t has a permanent and a transitory component, $y_t = \bar{y}_t + \omega_t$, where the permanent component is defined by $\bar{y}_t = \bar{y}_{t-1} + \varepsilon_t = \sum_{i=1}^{t}\varepsilon_i$ with $\varepsilon_t \sim i.i.d(0, \sigma_\varepsilon^2)$, $\omega_t \sim i.i.d(0, \sigma_\eta^2)$ and ε_t, η_t independent. In this case, the forecast error u_t follows an $AR(1)$ process. The details of this derivation are spelled in the Appendix A.1.

Several tests of the UIP hypothesis have been proposed in the literature. Bilson (1981) and Fama (1984) analyzed regression (1) with $y_t = s_{t+k} - s_t$ and $x_t = f_{t,k} - s_t$ with $k = 1$. In this setting a test of UIP is that $H_0 : \alpha = 0$ and $\beta = 1$. It has been noted by Fama (1984) and many subsequent studies that the estimated slope coefficient β is frequently negative. This is known as the Forward Premium Anomaly: the country with the higher rate of interest has an appreciating currency rather than a depreciating currency; a violation of UIP. Hansen and Hodrick (1980) estimate the model (1) with $y_t = s_{t+k} - f_{t,k}$, the forecast error, and uses lagged dependent variables as regressors; $x_t = (y_{t-k}, y_{t-(k+1)})$ with $k = 13$. The test is

$H_0 : \alpha = 0$ and $\beta_1 = \beta_2 = 0$. A simplified version of this regression, with $x_t = y_{t-k}$ and test $H_0 : \alpha = 0$ and $\beta_1 = 0$, has been studied by Baillie et al. (2023). Hansen and Hodrick (1980) also proposed a test that includes the lagged forecast errors of four other currencies,

$$s^i_{t+k} - f^i_{t,k} = \alpha_i + \beta \left(s^i_{t+k} - f^i_{t,k}\right) + \textstyle\sum_{j\neq i} \alpha_{ij}(s^j_{t+k} - f^j_{t,k}) + u^i_{t+k}, \tag{2}$$

where $s^i_{t+k} - f^i_{t,k}$ is the forecast error for country i and $s^j_{t+k} - f^j_{t,k}$ is the forecast error for country $j \neq i$. We focus on regression (2) as Hansen and Hodrick (1980) concludes that the multicountry test appears to be more powerful.

3 Feasible GLS

We consider the linear regression

$$y_{t+k} = x'_t\beta + u_{t+k}, \tag{3}$$

where the error term follows a short-memory linear process

$$u_{t+k} = C(L)\varepsilon_{t+k} = \textstyle\sum_{j=0}^{\infty} \theta_j \varepsilon_{t+k-j}, \tag{4}$$

where $\varepsilon_t \sim i.i.d.(0, \sigma^2)$. The polynomial $C(L) = \sum_{j=0}^{\infty} c_j L^j$ is assumed to satisfy $c_0 = 1$ (a normalization), $\sum_{j=0}^{\infty} j|c_j| < \infty$ so that the process is short-memory and $C(L)$ is invertible, i.e., we can write $A(L)^{-1}u_{t+k} = \varepsilon_{t+k}$. We use the FGLS procedure developed in Perron and González-Coya (2024) to estimate regression (3). The idea is to consistently approximate u_{t+k} using an autoregression of order k_T,

$$u_{t+k} = \textstyle\sum_{j=1}^{k_T} \rho_j u_{t+k-j} + \varepsilon_{t+k,k_T},$$

with $k_T \to \infty$ and $k_T^3/T \to 0$ as $T \to \infty$, to ensure consistent estimates; see Berk (1974). Replacing equation (3) and re-arranging terms we have,

$$y_{t+k} = \textstyle\sum_{j=1}^{k_T} \rho_j y_{t+k-j} + x'_t\beta - \sum_{j=1}^{k_T} x'_{t-j}\delta_j + \varepsilon_{t+k,k_T}, \tag{5}$$

with $\delta_j = \beta\rho_j$ for $j = 1, \ldots, k_T$. Equation (5) is often called the Durbin regression (see Durbin, 1970). The order of the autoregression, k^*_T, is determined via the minimization of the BIC suggested by Schwarz (1978) for $k_T \in [0, k_{max}]$ where k_{max} is such that $k^3_{max}/T \to 0$ as $T \to \infty$. We use the method suggested by Ng and Perron (2005) to ensure a proper comparison across models with different values of k_T , i.e., using the same effective number of observations. We estimate the Durbin equation (5) via OLS with $k_t = k_T^*$. Using the OLS estimates of ρ_j, $\hat{\rho}_j^D$, we construct the quasi-differenced variables

$$y_t^* = \left(y_t - \sum_{j=1}^{k_T^*} \hat{\rho}_j^D y_{t-j}\right), \quad x_t^* = \left(x_t - \sum_{j=1}^{k_T^*} \hat{\rho}_j^D x_{t-j}\right). \tag{6}$$

The FGLS estimate of β is the OLS estimate of the quasi-differenced regression,

$$y_t^* = x_t^{*\prime}\beta + \varepsilon_{t,k_T^*}, \quad (t = k_T^*, \ldots, T). \tag{7}$$

The resulting estimate $\hat{\beta}_{FGLS}$ will be consistent provided the regressors x_t are pre-determined with respect to values ε_t prior to period $t+k$; see Perron and González-Coya (2024).

3.1 Feasible GLS with Lagged Dependent Variables

Consider regression (3) with $x_t' = (y_{t-h}, w_t')$ for some $h > 0$, where w_t is a vector of $n+1$ pre-determined regressors that includes a constant term. For simplicity we omit the constant term without loss of generality. Write the regression as

$$y_t = \beta y_{t-h} + \alpha' w_t + u_t. \tag{8}$$

If u_t is autocorrelated beyond lags $i \geq h-1$, OLS applied to regression (8) is not consistent as y_{t-h} and u_t are not independent. Wallis (1967) and Malinvaud (1966) studied regression (8) with $h = 1$ where the error terms follows an $AR(1)$ process. Expression for the asymptotic bias of the OLS estimates of α and β are given by Malinvaud (1966) and Griliches (1961). We can write the Durbin regression as follows:

$$\begin{aligned}
y_t &= \sum_{j=1}^{h-1} \rho_j y_{t-j} + \gamma_h y_{t-h} + \sum_{j=h+1}^{k_T^*} \rho_j y_{t-j} \\
&\quad - \left(\sum_{j=1}^{k_T^*-h} \delta_{j+h} y_{t-h-j} + \sum_{j=k_T^*-h+1}^{k_T^*} \delta_{j+h} y_{t-h-j}\right) \\
&\quad + \alpha' w_t - \sum_{j=1}^{k_T^*} \rho_j \alpha' w_{t-j} + \varepsilon_{t,k_T^*} \\
&= \sum_{j=1}^{h-1} \rho_j y_{t-j} + \gamma_h y_{t-h} + \sum_{j=1}^{k_T^*-h} \gamma_{j+h} y_{t-h-j} + \sum_{j=k_T^*-h+1}^{k_T^*} \delta_{j+h} y_{t-h-j} \\
&\quad\quad + \alpha' w_t - \sum_{j=1}^{k_T^*} \psi_j' w_{t-j} + \varepsilon_{t,k_T^*}
\end{aligned} \tag{9}$$

where $\gamma_h = \rho_h + \beta$, $\gamma_{j+h} = \rho_{j+h} - \delta_{j+h}$, for $j = 1, \ldots, k_T^* - h$; $\delta_{j+h} = \rho_{j+h}\beta$, for $j = k_T^* - h + 1, \ldots, k_T^*$; and $\psi_{ij} = \alpha_i \rho_j$ for $j = 1, \ldots, k_T^*$, $i \in 1, \ldots, n$. For the case $h = k_T^* = 1$ and w_t a scalar (i.e., $n = 1$), Wallis (1967) and Malinvaud (1966) (p. 469) propose to estimate regression (9) using OLS and then estimate ρ_1 using $-\psi_j/\alpha$. For the general case $k > 1$ and $k_T^* > h$, the same approach can be applied. Note that we need $k_T^* \geq h$, otherwise the estimates are not consistent. But this condition will be satisfied, at least in large samples, when using the BIC to select the lag order. However, note that when $n > 1$, ρ_j cannot be uniquely identified. We shall propose two estimation methods; one based on the Durbin regression (9) and one based on a first-stage instrumental variable (IV) estimate. These follow similar steps as in the feasible GLS procedure discussed above.

Remark 1. *Estimating the autoregressive coefficients, ρ_j for $j = 1, \ldots, k_T^*$, from regression (9) using the Wallis (1967) and Malinvaud (1966) method does not allow us to uniquely identify ρ_j in the general case with $k > 1$, $k_T^* > k_t$ and $n > 1$. However, we can obtain efficient estimates $\widetilde{\rho}_j^*$, $j = 1, \ldots, k_T^*$ by using a convex combination of the estimates $\widetilde{\rho}_{ij} = -\widetilde{\psi}_{ij}/\widetilde{\alpha}_i$ ($j = 1, \ldots, k_T^*$) for $i \in 1, \ldots, n$. For ease of the exposition, suppose that $n = 2$. Then, we can construct efficient estimates $\widetilde{\rho}_j^* = \lambda\widetilde{\rho}_{1j} + (1-\lambda)\widetilde{\rho}_{2j}$, where the optimal λ that minimize $\mathrm{Var}(\widetilde{\rho}_j^*)$ is (the details are in the Appendix A.2):*

$$\lambda = \frac{Var\left(\widetilde{\rho}_{2j}\right) - Cov\left(\widetilde{\rho}_{1j}, \widetilde{\rho}_{2j}\right)}{Var\left(\widetilde{\rho}_{1j}\right) + Var\left(\widetilde{\rho}_{2j}\right) - 2Cov\left(\widetilde{\rho}_{1j}, \widetilde{\rho}_{2j}\right)}.$$

In practice, we can estimate $Var\left(\widetilde{\rho}_{ij}\right)$ using a first order Taylor expansion,

$$Var\left(\widetilde{\rho}_{ij}\right) = Var\,\mathrm{Var}\left(\widetilde{\psi}_{ij}/\widetilde{\alpha}_i\right) = \frac{\widetilde{\psi}_{ij}^2}{\widetilde{\alpha}_i^2}\left[\frac{Var\left(\widetilde{\psi}_{ij}\right)}{\widetilde{\psi}_{ij}^2} - 2\frac{Cov\left(\widetilde{\psi}_{ij}, \widetilde{\alpha}_i\right)}{\widetilde{\psi}_{ij}\widetilde{\alpha}_i} + \frac{Var\left(\widetilde{\alpha}_i\right)}{\widetilde{\alpha}_i^2}\right].$$

The first order Taylor expansions of $Cov\left(\widetilde{\rho}_{1j}, \widetilde{\rho}_{2j}\right)$ are cumbersome. Note that if the instruments w_t^j are independent of each other, $\mathrm{Cov}\left(\widetilde{\rho}_{1j}, \widetilde{\rho}_{2j}\right)$ will be arbitrarily small in large samples. Hence, we set $Cov\left(\widetilde{\rho}_{1j}, \widetilde{\rho}_{2j}\right) = 0$.

4 Instrumental Variables

An alternative to estimate regression (8) with $n \geq 1$ is to use an instrumental variable procedure. If the regressors w_t are exogenous, then w_{t-h} are valid instruments for y_{t-h} and the two-stage least squares (2SLS) estimates will be consistent. Liviatan (1963) propose to use w_{t-h} as instruments for y_{t-h} and w_t as an instrument for itself, so that the instrument set is $Z_t = \{w_{t-h}, w_t\}$. We can potentially select a larger set of instrumental variables for y_{t-h}. e.g., lags or order $i \geq h$ of w_t. In this case $z_t = \{w_t, w_{t-i}, p > i \geq h\}$, for some $p > h$. Note that p can be adaptively selected using appropriate tests for over-identifying restrictions (see Small, 2007). However, if the regressors w_t are not exogenous and u_t is autocorrelated beyond lags $i \geq h-1$, w_{t-h} are no longer valid instruments. Hence, the IV estimate of regression

(8) using the instrument set z_t will not be consistent. This motivates the use of the GLS-IV procedure suggested by Olivari and Perron (2024). The idea is simple: first transform the model to have serially uncorrelated errors and then estimate the transformed model via IV using the set of transformed instruments. The resulting estimate will be consistent as the instrument set and regressors are pre-determined in the transformed model.

4.1 Estimator Valid with Exogenous Instruments

We first discuss two methods that are valid with exogenous instruments if u_t is autocorrelated beyond lags $h-1$. One is the widely used so-called "optimal GMM" procedure. The other is akin to the GLS procedure discussed above but with the first-step using the GMM estimate to obtain estimate of the residuals and construct the autoregressive filtering.

4.1.1 GMM

Using the set of $n(p-h)$ exogenous instruments z_t, then, as shown in Hansen (1982), the best estimator of $\beta^* = (\beta, \alpha)$ based on the instruments and the moment condition $E(Zu) = 0$ is:

$$\hat{\beta}^*_{GMM} = (X'Z\Omega^{-1}Z'X)^{-1}X'Z\Omega^{-1}Z'y, \tag{10}$$

where the $n(p-h) \times n(p-h)$ matrix Ω is given by $\Omega = \lim_{T\to\infty} T^{-1}E[Z'uu'Z]$. We can write $\Omega = \sum_{s=-\infty}^{\infty} R_v(s)$, where $R_v(s) = E\left[Z^h(t)u_t Z^{h\prime}(t-s)u_{t-s}\right]$ and $Z^h(t) = \{z_t, \ldots, z_{t-h}\}$. Then Ω can thus be consistently estimated by $\hat{\Omega} = \sum_{s=-T}^{T} \lambda(s,m)\hat{R}_v(s)$, where $\hat{R}_v(s) = T^{-1}\sum_{t=1}^{T-|s|} v_t' v_{t+|s|}$, with $v_t = Z^{h\prime}(t)u_t$, $\lambda(s,m)$ is some kernel or weight function and m is the bandwidth; see, e.g., Andrews (1991).

4.1.2 GMM-GLS-IV

We also consider a FGLS method that does not rely on the Durbin regression (9). Instead, it uses a first-stage GMM estimate to obtain a consistent estimate of the residuals, $\widetilde{u}_t$. We can thus identify the autoregressive parameters when the instruments are exogenous. The GMM-GLS-IV procedure to estimate regression (8) with exogenous instruments w_t is the following: 1) Obtain the GMM estimator of regression (8), given by (10) using the set of instruments $z_t = \{w_t, w_{t-h}, w_{t-h-j}, j = 1, \ldots, h\}$. Compute the residuals, $\widetilde{u}_t = y_t - \widetilde{\alpha}'_{iv} w_t - \widetilde{\beta}_{iv} y_{t-h}$; 2) Select the order k_T^* of the autoregression

$$\widetilde{u}_{t+k} = \textstyle\sum_{j=1}^{k_T} \rho_j \widetilde{u}_{t+k-j} + \varepsilon_{t+k,k_T}, \quad t = 1, ..., T - k_{max}, \tag{11}$$

via the minimization of the BIC for $k_T \in [0, k_{max}]$ where k_{max} is such that $k_{max}^3/T \to 0$ as $T \to \infty$.; 3) Estimate the autoregression (11) with $k_T = k_T^*$ to obtain consistent estimates $\widetilde{\rho}_j$ $(j = 1, \ldots, k^*)$; 4) Use $\widetilde{\rho}_j$ $(j = 1 \ldots, k_T^*)$ to construct the quasi-differenced variables $y_t^* = (y_t - \sum_{j=1}^{k_T^*} \widetilde{\rho}_j y_{t-j})$, $y_{t-h}^* = (y_{t-h} - \sum_{j=1}^{k_T^*} \widetilde{\rho}_j y_{t-h-j})$ and $w_t^* = (w_t - \sum_{j=1}^{k_T^*} \widetilde{\rho}_j w_{t-j})$; 5) The GMM-GLS-IV estimate of β is the IV estimate of the quasi-differenced regression

$$y_t^* = c + \beta y_{t-h}^* + \alpha' w_t^* + \varepsilon_{t,k_T^*}, \tag{12}$$

using the set of quasi-differenced instruments $z_t^* = \{w_t^*, w_{t-h}^*\}$.

4.2 Non-Exogenous Instruments: GLS-IV

We next describe the GLS-IV procedure to estimate regression (8), which is valid with non-exogenous instruments, provided they are pre-determined. The steps are the following: 1) Select the order of the Durbin regression (9), k_T^* via the minimization of BIC for $k_T \in [0, k_{max}]$ where where k_{max} is such that $k_{max}^3/T \to 0$ as $T \to \infty$; 2) Estimate the Durbin regression (9) with the selected value k_T^* using OLS. The estimates the autoregressive coefficients, $\widetilde{\rho}_j^*$, $j = 1, \ldots, k_T^*$ are obtained using the efficient method described in Remark 1; 3) Use $\widetilde{\rho}_j$, $j = 1 \ldots, k_T^*$ to construct the quasi-differenced variables y_t^*, y_{t-k}^* and w_t^*, as in Step 4 for GMM-GLS-IV; 4) The GLS-IV estimate of β is the IV estimate of the quasi-differenced regression (12) using the set of quasi-differenced instruments $z_t^* = \{w_t^*, w_{t-h}^*\}$.

Given that the OLS estimates from the Durbin regression (9) are consistent and that the transformed model (12) has serially uncorrelated errors, the GLS-IV estimates are consistent under the stated conditions. To the best of our knowledge, there is no other consistent estimation method requiring only pre-determined regressors for the general linear regression (8) without restricting the error process u_t. Maximum Likelihood and Non-Linear Least Squares require an a priori known error process. In the simulation experiments reported in Section 6.1.2, we show that the fact that $\widetilde{\rho}_j$ is not uniquely identified when we have more than one regressor $(n > 1)$, does not affect the efficiency of GLS-IV.

5 Fama Regression and the Forward Discount Anomaly

The Fama (1984) regression is

$$s_{t+h} - s_t = \alpha + \beta(f_{t,h} - s_t) + u_{t+h}, \tag{13}$$

with $h = 4$ for 1-month forward rates and $h = 12$ for 3-month forward rates, when using weekly data. Estimates of (13) tell us whether the current forward-spot differential, $f_{t,h} - s_t$, has power to predict the future change in the spot rate, $s_{t+h} - s_t$. Evidence that β is significantly different from zero means that the forward rate observed at t has information about the spot rate to be observed at $t+h$. Under the efficient-market hypothesis, we have $H_0 : \alpha = 0, \beta = 1$. Under the null hypothesis, the log of the forward rate provides an unbiased forecast of the log of the future spot exchange rate. Derivations from $\beta = 1$ are sometimes interpreted as a measure of the variation of the premium in the forward rate.

Let $\bar{\beta}$ be the OLS estimate of β in regression (13). If the estimator is consistent, we have

$$\text{plim}(\bar{\beta}) = \beta = \frac{\text{Cov}\left(f_{t,h} - s_t, s_{t+1} - s_t\right)}{\text{Var}\left(f_{t,h} - s_t\right)}. \tag{14}$$

If expectations are rational, then $s_{t+1} - s_t = E_t\left(s_{t+1}\right) - s_t + \varepsilon_{t+1}$, where $\varepsilon_{t+1} = s_{t+1} - E_t\left(s_{t+1}\right)$ is the forecast error. In this case, $\text{Cov}\left(f_{t,h} - s_t, s_{t+1} - s_t\right) = \text{Cov}\left(f_{t,h} - s_t, E_t\left(s_{t+1}\right) - s_t\right)$. The foreign exchange risk premium when expectations are rational is defined as $rp_t^{re} = f_{t,h} - E_t\left(s_{t+1}\right)$. Under risk neutrality, expected profits from forward market speculation would be zero as agents would drive $f_{t,h}$ into equality with $E_t\left(s_{t+1}\right)$. Write $E_t\left(s_{t+1}\right) - s_t = f_{t,h} - s_t - rp_t^{re}$ and replace into equation (14), so that $\text{plim}_{T\to\infty}(\bar{\beta}) = 1 - \beta^{rp}$, where

$$\beta_{rp} = \frac{\text{Cov}\left(E_t\left(s_{t+1}\right) - s_t, rp_t^{\text{re}}\right) + \text{Var}\left(rp_t^{\text{re}}\right)}{\text{Var}\left(f_{t,h} - s_t\right)}.$$

A negative estimate of β in regression (13) is a robust finding in the literature (see Engel, 1996). This is known as the "forward discount anomaly"; it is a widespread empirical finding that the returns on nominal exchange rates appear to be negatively correlated with the lagged forward premium. Bilson (1981) and Fama (1984) provide evidence that the estimates of β are less than zero. Many subsequent studies have confirmed that finding, for dollar exchange rates and a large number of exchange rates and time periods (see, for example, Bekaert and Hodrick, 1993; Backus et al., 1993; Hai et al., 1997). Froot (1990) notes that the average value of $\bar{\beta}$ over 75 published estimates is -0.88. Only a few of the estimates are greater than zero, and none is greater than 1. The forward discount anomaly implies an appreciating currency for the high interest rate country. It is an "anomaly" as rational expectations would imply the opposite; if all currencies are equally risky, investors would demand higher interest rates on currencies expected to fall in value. The survey by Engel (1996) focuses on the the possibility that $\bar{\beta}_{rp} \neq 0$ among the possible explanations for finding $\bar{\beta} < 0$. Other possible interpretations are that the forward rate is a biased predictor of the future spot rate, and/or that it is evidence of a time-varying risk premium. In this paper, we provide a novel insight on this empirical regularity. We argue that the OLS and GMM estimates of β are not consistent, as the error term in regression (13) is serially correlated beyond lags $k-1$. We find that the FGLS estimates, which are consistent under a broader range of conditions, are significantly non negative but in general smaller than 1.

5.1 Data Set

Daily data were obtained for the spot exchange rates for the U.K. pound (US-UK), Canadian dollar (US-CAD) and Japanese yen (US-JP) as well as the 1-month and 3-month forward exchange rates data for the three currencies. As in Hansen and Hodrick (1980), the data were sampled to form a weekly series constructed by taking observation on Tuesday of each week. If no Tuesday observation was available, we used the Wednesday observations. The source of the forward exchange data for US-UK and US-CAD is Barclays Bank PLC; the source for US-JP is the Bank of Tokyo Mitsubishi. For all the data sets, we use all the information available until 01/17/2023 but the starting date of the time series differ: a) US-UK: starting date of 10/11/1983, 2,050 observations; b) US-CAD: starting date of 12/14/1984, 1,989 observations; c) US-JP: starting date of 09/01/1993, 1,535 observations.

5.2 Simulation Design

In this section, we provide simulation results related to the Fama (1984) regression (13) under the "RE case" or efficient market hypothesis (EMH) , $H_0 : \alpha = 0, \beta = 1$. We consider $h = 4$ for 1-month forward rates and $h = 12$ for 3-month forward rate. As discussed in Section 2, the EMH implies that u_{t+h} has an $MA(h-1)$ representation. We simulate an $MA(h-1)$ process with parameters calibrated to replicate the observed autocorrelation function up to lag $h-1$ of the FGLS residuals of equation (13) using US-UK data. The details are presented in the Appendix A.3. For $h = 4$ we have an $MA(3)$ representation with

$$\theta(L) = \left(1 + 0.59L + 0.82L^2 + 0.61L^3\right). \tag{15}$$

For $h = 12$ we have an $MA(11)$ representation with coefficients

$$(0.76, 0.63, 0.54, 0.66, 0.49, 0.33, 0.24, 0.57, 0.41, 0.12, 0.39).$$

We simulate an error process based on equation (15), i.e., $v_{t+h} = \theta(L)\varepsilon_t$, where $\varepsilon_t \sim i.i.d.$ $N(0, \sigma_\varepsilon^2)$ and σ_ε^2 is estimated using the residuals of an initial Fama FGLS regression (13) for US-UK. The data generating process (DGP) uses $x_t = f_{t,h} - s_t$ observed in the data for US-UK, and the stated simulated error process. We artificially generate y_{t+4} in order to satisfy the null hypothesis $H_0 : \beta = 1$. Thus, the DGP is

$$s_{t+h} - s_t = (f_{t,h} - s_t) + v_{t+h}. \tag{16}$$

We also consider a departure from the "RE case" with errors v_{t+h} following the "gneral case" with serial correlation at lags $i \geq h$. We assume that v_{t+h} follows an $ARMA(1,h)$ process with MA coefficients given by (15) and AR coefficient $\rho = 0.6$. We consider two sampling periods; the complete sample from 10/11/1983 to 01/17/2023 with 2,050 observations and the one spanning 11/01/1989 to 04/01/2021, as considered in Baillie et al. (2023). We perform 5,000 replications. We consider the following estimators of α, β, from regression (13): a) The OLS estimate with HAC standard errors based on the weighting scheme suggested by Andrews (1991) with automatic bandwidth selection using an $ARMA(1,1)$ approximation; b) The Durbin estimate based on the regression (5) with $y_t = s_{t+h} - s_t$ and $x_t = \{1, f_{t,h} - s_t\}$ and k_T^* selected using the BIC. For the complete sample, we consider $k_{max} = 40$ and for the sub-sample we set $k_{max} = 30$; c) The FGLS estimate based on regression (7) with $y_t = s_{t+h} - s_t$, $x_t = \{1, f_{t,h} - s_t\}$ and k_T^* selected using the BIC. The same values of k_{max} are used. We estimate the sample variance of the FGLS estimator using $Var\left(\hat{\gamma}_{FGLS}\right) = (X^{*\prime}X^*)^{-1}\hat{\sigma}^2_{FGLS}$, where X^* is the $T \times 2$ matrix of quasi-differenced regressors (including a constant term) and $\hat{\sigma}^2_{FGLS}$ is the sample variance of the FGLS residuals, $\hat{u}_{FGLS} = y_t - \hat{\gamma}'_{FGLS}x_t$. The confidence intervals at the α nominal level for the jth coefficient are obtained using $\hat{\gamma}_j \pm z_{1-\alpha/2}Var(\hat{\gamma}_{j,FGLS})^{1/2}$, where $z_{1-\alpha/2}$ is the $1 - \alpha/2$ quantile of the normal distribution. We set $\alpha = 0.05$ so that 95% nominal level confidence intervals are applied.

The simulation results are presented in Table 1 for $h = 4$ and Table 2 for $h = 12$. In line with the theory, the mean squared error (MSE) of OLS is small when the error term follows an $MA(h)$ process and deteriorates when the error is serially correlated at lags $i \geq h$. Clearly, FGLS outperforms OLS and Durbin in all cases. Even under the "RE case", the MSE of OLS is on average 3.4 times larger than of FGLS, while the MSE of Durbin is close to that of OLS case. When the errors follow an $ARMA(1,h)$ process, OLS is not longer efficient and its MSE is on average 27 times that of FGLS. As expected, the Durbin estimate remains consistent in this case, but its MSE is on average 3 times that of FGLS. FGLS also has the smallest variance. The variance of Durbin is on average 3 times the variance of FGLS for $h = 4$ and 2 times the variance of FGLS for $h = 12$. The coverage rates of the confidence interval for FGLS are near the nominal 90% and have shortest lengths. OLS with HAC standard errors exhibits substantial size distortions with $ARMA(1,h)$ errors, especially when $h = 12$ case. The OLS based HAC standard errors provides confidence intervals close to the nominal level in some cases, at the expense of a very large variance. For the $ARMA(1,h)$ case with $k = 4$ (12) the variance of OLS is 30 (25) times the variance of FGLS. This results are in line with those from Perron and González-Coya (2024).

For the simulations related to power, we use US-UK observed data $x_t = f_{t,h} - s_t$ for the period 11/01/1989-04/01/2021 with 1-month forward rates, $h = 4$. The DGP is

$$s_{t+4} - s_t = \beta(f_{t,4} - s_t) + v_{t+4}, \tag{17}$$

where $\beta \in \{0, 0.25, 0.5, 0.75, 1, 1.25, 1.5, 1.75, 2\}$. The error process v_{t+4} is simulated as before; under the "RE case" v_{t+4} follows an $MA(3)$ process, and under the "general case" v_{t+4} follows an $ARMA(1,3)$ process. Figures 1 and 2 present plots of the empirical rejection frequencies for t-test of $H_0 : \beta = 1$ with nominal size 0.05 for the OLS estimates with HAC-$ARMA(1,1)$ standard errors, the FGLS estimates based on regression (7) and the Durbin estimates based on the regression (5). For the latter two, k_T^* is selected via BIC with $k_{max} = 30$. Figure 1 pertains to "RE case" with the errors an $MA(3)$ process, while Figure 2 pertains to the "general case" with errors following an $ARMA(1,3)$ process.

Note that in Figure 1, "RE case" with $MA(3)$ errors, OLS and Durbin has very small rejection frequencies even when β is far from the null value 0, even though they are consistent, which can be attributed to their lack of efficiency in finite samples. As shown in Figure 2, the power functions are similar in the case with $ARMA(1,3)$ errors, with the exception that the power function of OLS decreases and flattens with an almost constant 10% rejection frequency for all values, despite having a more liberal size.

5.3 Empirical Results

We present the estimation results of regression (13) using the three estimates considered before: OLS, Durbin estimates based on the regression (5) and FGLS based on regression (7). We use data for three currencies, US-UK, US-CAD and US-JP and we consider two sampling periods; the first one is the complete sample and the second one spans from 11/01/1989 to 04/01/2021, the sampling period considered in Baillie et al. (2023). For the Durbin and FGLS estimates we set $k_{max} = 40$ for the full sample and $k_{max} = 30$ for the sub-sample period. Tables 3 and 4 present the estimation results using 1-month forward rates ($h = 4$) and the 3-month forward rates ($h = 12$), respectively. In line with the empirical regularity in the literature, the OLS estimates of β are not significantly positive in all cases. In contrast, the Durbin and FGLS estimates are positive in most cases and significantly non negative in some. In some cases, such as the US-JP exchange rate for 3-month forward rates, we observe a negative significant OLS estimate and a positive significant FGLS estimate. Recall that the Durbin and FGLS estimates are consistent even when the error follows a general linear process. We thus interpret the large differences between the OLS and the FGLS estimates as evidence of OLS being inconsistent. This finding suggests that the forward discount anomaly might be a consequence of OLS providing an inconsistent estimate of β.

6 Hansen and Hodrick Regression

We now turn to an estimation problem that shares some of the main features, though with added complexities. Our aim is to efficiently estimate Hansen and Hodrick (1980) regression,

$$y_{t+h} = \alpha + \beta y_t + \sum_{j \neq i} \alpha_j w_t^j + u_{t+h}, \tag{18}$$

where $y_{t+h} = s_{t+h}^i - f_{t,h}^i$ and $w_t^j = s_t^j - f_{t-h,h}^j$ for $j \neq i$. Note that if α_j is significantly different from zero, then w_{t-h}^j is correlated with the regression variable y_t and can be potentially used as an instrument. If the lagged forecast error of country $j \neq i$, w_{t-h}^j, is exogenous i.e., uncorrelated with the residuals u_s for all t and s then GMM and GMM-GLS-IV are consistent. If w_{t-h}^j is only pre-determined, the OLS and GMM (and thereby GMM-GLS-IV) estimates are not consistent in general, but GLS-IV will remain consistent. For GMM and the first-step for GMM-GLS-IV we consider the set of instrumental variables $Z = \{w_t^j, w_{t-h}^j, w_{t-h-l}^j, j \neq i, l = 1, \ldots, h\}$. We use weekly data for 1-month forward rates and 3-month forward rates. In the former case, $h = 4$ whereas in the second $h = 12$. We consider the same data set as in Section 5.1. We first present simulations tailored to this problem to shed light on the properties of the various estimators under a range of plausible scenarios.

6.1 Simulation Results

We present two sets of Monte Carlo experiments. The DGP is inspired by the Hansen and Hodrick (1980) regression. We consider 1-month forward rates so that $h = 4$,

$$y_{t+4} = \alpha_0 + \beta y_t + \sum_{j \neq i} \alpha_j w_t^j + u_{t+4}, \tag{19}$$

where $y_{t+4} = s_{t+4}^i - f_{t,4}^i$, $i = UK$ and $w_t^j = s_t^j - f_{t-4,4}^j$, $j = \{CAD, JP\}$. In the first set of Monte Carlo experiments, the simulation design is based on observed weekly spot and forward exchange rates for US-UK, US-CAD and US-JP. We use actual forecast errors w_t^j for US-CAD and US-JP. By construction, the regressors w_t^j will be exogenous. In Section 6.1.2

we present the second set of Monte Carlo experiments, where the regressors w_t^j are jointly simulated with the error process u_t, so they are serially correlated and non-exogenous.

We start with the case of exogenous instruments. We consider the DGP (19) under the null hypothesis $H_0 : \alpha_0 = 0, \beta = 0$. We set $\alpha_1 = \alpha_2 = 1$, so that w_t^j, $j = \{CAD, JP\}$ can be used as instruments. We use actual observed data for w_t^j, $j = \{CAD, JP\}$ and we generate y_{t+4} according with the DGP (19) imposing the null hypothesis. As discussed in Section 2, the "RE case" implies that u_{t+4} has an $MA(3)$ representation. We simulate an $MA(3)$ process with parameters calibrated to replicate the observed autocorrelation function up to lag 3 of the GLS-IV residuals from equation (18) using the real data set, following the same procedure as described in Appendix A.3. The resulting parameters are,

$$\theta(L) = \left(1 + 0.82L + 0.74L^2 + 0.47L^3\right). \tag{20}$$

We simulate an error process based on equation (20), $u_{t+h} = \theta(L)\varepsilon_t$, where $\varepsilon_t \sim i.i.dN(0, \sigma_\varepsilon^2)$ and σ_ε^2 is estimated using the GLS-IV residuals from the initial regression (19). We use the weekly spot exchange rates for US-UK, US-CAD, US-JP for the period between period November 2010 to April 2020 (492 observations). We perform 5,000 replications. We consider the following estimates: a) OLS applied to regression (19). We use HAC standard errors with the Quadratic Spectral weighting scheme of Andrews (1991) with automatic bandwidth selection using an $ARMA(1,1)$ approximation; b) GMM: the estimate $\widetilde{\gamma}$ (10) from regression (19) using the optimal weighting matrix as described in Section 4.1.1. We use the set of instruments $z_t = \{w_t^j, w_{t-h}^j, w_{t-h-l}^j, j = \{CAD, JP\}, l = 1, \ldots, h\}$; c) GLS-IV: the estimate from the procedure described in Section 4.2. For the Durbin regression (9), we set $k_{max} = 30$. The autoregressive coefficients ρ_j are estimated using the efficient method described in Remark 1. For the IV estimates, we use the set of quasi-differenced instruments $z_t^* = \{w_t^{*j}, w_{t-h}^{*j}, j = \{CAD, JP\}\}$. The variance estimate is:

$$\text{Avar}\left(\hat{\gamma}_{GLS-IV}\right) = (X^{*\prime}Z^*(Z^{*\prime}Z^*)^{-1}Z^{*\prime}X^*)^{-1}\hat{\sigma}^2_{GLS-IV},$$

where X^* is the $T \times 4$ matrix of quasi-differenced regressors (including a constant term) and $\hat{\sigma}^2_{GLS-IV}$ is the sample variance of the GLS-IV residuals $\hat{u}_{GLS-IV} = y_t - \hat{\beta}_{GLS-IV}y_{t-1} - \hat{\alpha}'_{GLS-IV}w_t$; d) GMM-GLS-IV: the estimate from the procedure described in Section 4.1.2. Step 1 uses the GMM estimate $\widetilde{\gamma}$ described above. For the autoregression (11), we set $k_{max} = 30$. For the IV estimate, we use the set of quasi-differenced instruments $z_t^* = \{w_t^{*j}, w_{t-h}^{*j}, j = \{CAD, JP\}\}$. The variance estimate is computed as,

$$\text{Avar}\left(\hat{\gamma}_{GMM-GLS-IV}\right) = (X^{*\prime}Z^*(Z^{*\prime}Z^*)^{-1}Z^{*\prime}X^*)^{-1}\hat{\sigma}^2_{GMM-GLS-IV}.$$

where X^* is the $T \times 4$ matrix of quasi-differenced regressors (including a constant term) and $\hat{\sigma}^2_{GMM-GLS-IV}$ is the variance of the GMM-GLS-IV residuals,

$$\hat{u}_{GMM-GLS-IV} = y_t - \hat{\beta}_{GMM-GLS-IV} y_{t-1} - \hat{\alpha}'_{GMM-GLS-IV} w_t.$$

6.1.1 The Case with Exogenous Instruments

We start with the case with exogenous instruments and first assess the finite sample size of the estimators. The simulation results are presented in the first panel of Table 5, which report the MSE, bias, variance, coverage rate and average length of confidence intervals for the parameter β. Under the "RE case", the error term follows an $MA(3)$ process so that the OLS estimates are consistent. The GMM estimate has a variance that is half of the OLS variance but with a coverage rate below the nominal level. The GMM-GLS-IV procedure using the GMM as a first step estimate achieves an important reduction in MSE while maintaining coverage rates near the nominal level. The finite sample performance of GLS-IV and GMM-GLS-IV are similar. Both have the smallest MSE and yield confidence intervals with coverage rates near the nominal level and the shortest length, unlike GMM.

We now consider simulations under the "general case". We use the DGP (19). However, we consider a departure of the efficient market hypothesis in which the error term u_{t+4} is serially correlated at lags $i \geq 4$. In particular, we assume that u_{t+4} follows an $ARMA(1,3)$ process with MA coefficients given by (20) and AR coefficient $\rho = 0.6$. The forward rate is generated in the same way as before. We consider the same family of estimators. The results are presented in the second panel of Table 5. In this case, as the error process is correlated beyond lag 3, OLS is not longer consistent. This translates in an important increase in MSE and bias. The confidence intervals are meaningless, in that they have huge size distortions, with a coverage rate smaller than 20%. Since the instruments are exogenous, the GMM estimate remains consistent in this case but is not efficient. The MSE of GMM is 8 times that of FGLS. The GMM-GLS-IV estimate provides marked improvements over GMM and achieves important reduction in MSE along with confidence intervals having coverage rates near the nominal level. The finite sample performance of the GLS-IV estimate is similar to GMM-GLS-IV. Both estimates have the smallest MSE and variance with confidence intervals coverage rates near the nominal level and the shortest length.

We next consider the power of the various tests. We consider the DGP (19) for a grid values of β around the null hypothesis $H_0 : \beta = 0$. In particular, we consider $\beta \in \{-0.3, -0.2, -0.1, 0, 0.1, 0.2, 0.3\}$. We set $\alpha = (0, 1, 1)$, so that w_t^j, $j = \{CAD, JP\}$ can be used as instruments. We use the actual observed data for w_t^j, $j = \{CAD, JP\}$ and we generate y_{t+4} according with the DGP (19). The error process u_{t+4} is simulated as before: under the "RE case", u_{t+4} follows an $MA(3)$ process, and under the "general case" u_{t+4} follows an $ARMA(1,3)$ process with the same parameter configurations used earlier. In Figures 3 and 4, we plot the empirical rejection frequencies of nominal $\alpha = 0.05$ t-test of $H_0 : \beta = 0$, for the same set of estimates considered in the previous section. Figure 3 pertains to the "RE case", while Figure 4 present results for the "general case". Note that for the "RE case" with $MA(3)$ errors case, OLS has good power but is slightly outperformed by the FGLS-based estimates. Remarkably, GMM presents power distortions as it rejects the null hypothesis over 20%. In the "general case" with $ARMA(1,3)$ errors, Figure 4 shows that the power of GMM-GLS-IV and GLS-IV remains the same while that of OLS exhibits huge power distortions skewed at negative values of β. The GMM estimate also has large size distortions and skewed power functions.

In summary, for the case with exogenous instruments, GMM-GLS-IV and GLS-IV clearly have better properties. The performance of OLS is nearly as good in the "RE case" but completely breaks down in the "general case". Hence, GMM-GLS-IV and GLS-IV are clearly the more robust method of estimation and testing.

6.1.2 Simulations with Non-exogenous Instruments

We now consider the case with non-exogenous instruments and first assess the finite sample size of the estimators followed by some power comparisons. The DGP is based on regression of Hansen and Hodrick (1980), considering 1-month forward rates ($h = 4$),

$$y_{t+4} = \alpha_0 + \beta y_t + \textstyle\sum_{j \neq i} \alpha_j w_t^j + u_{t+4}, \tag{21}$$

where $y_{t+4} = s_{t+4}^i - f_{t,4}^i$ and $w_t^j = s_t^j - f_{t-4,4}^j$, $j \neq i$. In this case, the set of instruments is simulated so that they are serially correlated and non exogenous. Accordingly, we set

$$w_t^j = \rho_w w_{t-1}^j + v_t^j + \gamma \varepsilon_{t-1},$$

for $j = 1, 2$, with $v_t^j \sim i.i.d.N(0,1)$ independent of $\varepsilon_t \sim i.i.d.N(0,1)$. Note that ε_t is shock affecting u_t and thus, the instrument w_t^j is not exogenous whenever $\gamma \neq 0$. We set $\gamma = 0.3$ and $\rho_w = 0.5$. The process y_{t+4} is generated according to the DGP (21) under the null hypothesis $\alpha_0 = 0, \beta = 0$. We set $\alpha_1 = \alpha_2 = 1$ so that w_t^j $(j = 1, 2)$ can be used as instruments. Under the "RE case", the error u_{t+4} follows a $MA(3)$ process. Under the "general case", u_{t+4} follows the same $ARMA(1,3)$ process as in Section 6.1. We consider the same set of estimates: OLS+HAC, GMM, GLS-IV and GMM-GLS-IV. The sample size is $T = 300$ and 5,000 replications are used.

The simulation results for the "RE case" are presented in the first panel of Table 6. In this case, the OLS estimates are consistent requiring only pre-determined regressors. It is, however, the less efficient estimate and the coverage rates of the associated confidence intervals are below the nominal level. The bad performance of GMM is again corrected when using the GMM-GLS-IV estimate; it achieves the smallest MSE and the coverage rate is better than that of the GLS-IV estimate. The GMM-GLS-IV and GLS-IV estimates have confidence intervals with coverage rates slightly below to the nominal level (92% and 87%, respectively). This results are in line with the Monte Carlo simulation results for the non exogenous regressors and/or instruments cases in Perron and González-Coya (2024) and Olivari and Perron (2024).

The simulation results for the "general case" with $ARMA(1,3)$ errors are presented in the second panel of Table 6. Since the serial correlation in the errors extends beyond lag 3, OLS and GMM are not longer consistent. This is reflected in large MSE, bias and variance. The size distortions are exacerbated with coverage rates of the confidence intervals below 65% (26% for OLS). The finite sample performance of the GLS-IV estimate is in line with the fact that it is consistent. Surprisingly, despite the fact that the first step GMM estimates of the GMM-GLS-IV procedure are not consistent, the resulting GMM-GLS-IV estimate has smaller MSE than the GLS-IV estimate. These results suggest that the FGLS-based procedures are very robust to the first-stage estimates of the autocorrelation coefficients. The MSE of GLS-IV is almost 12 times smaller than the MSE of OLS and it achieves with a variance that is on average 5 times smaller than the variance of the other estimates. The coverage rates of the confidence intervals of the GMM-GLS-IV and GLS-IV estimates are near the nominal level with the smallest length.

We now consider the power analysis. We consider the DGP (21) for a grid values of β around the null hypothesis $H_0 : \beta = 0$. In particular, we consider $\beta \in \{-0.3, -0.2, -0.1, 0, 0.1, 0.2, 0.3\}$. We set $\alpha = (0, 1, 1)$, so that w_t^j, $j = 1, 2$ can be used as instruments. We generate y_{t+4} according with the DGP (21). Under the "RE case", u_{t+4} follows an $MA(3)$ process, and under the "general case" with u_{t+4} an $ARMA(1,3)$ process. In Figures 5 and 6, we plot the empirical rejection frequencies of nominal $\alpha = 0.05$ t-test of $H_0 : \beta = 0$, for the same set of estimates considered in the previous section. Figure 5 pertains to the "RE case", while Figure 6 pertains to the "general case with $ARMA(1,3)$ errors.

From the results in Figure 5 for $MA(3)$ errors, all the tests have similar power functions, though somewhat lower for OLS and GMM when β is positive. The FGLS-based procedures exhibit the same power function, with a rejection frequency of the null hypothesis that is slightly higher than the 5% nominal level. For GMM the null rejection frequency is higher than 15%. The results in Figure 6 pertaining to the $ARMA(1, 3)$ errors case show that the power functions of the FGLS-based procedures exhibit the same behavior as in the $MA(3)$ errors case. On the other hand, OLS and GMM are now subject to important size distortions. The power function of those estimates is biased towards negative values of β. Note that OLS rejects the null hypothesis $H_0 : \beta = 0$ almost 10% of the times when $\beta = -0.2$, while it rejects H_0 with a frequency higher than 75% when $\beta = 0$. Hence, the only reliable test are those obtained using the FGLS-based procedures.

6.2 Empirical Results

We now report estimates of the model (18) using weekly spot exchange rates for the U.K. pound (US-UK), Canadian dollar (US-CAD) and Japanese yen (US-JP). We use the complete US-JP sample that spans from June 1995 to January 2023, with 1,442 observations. We consider the OLS+HAC, GMM, GMM-GLS-IV and GLS-IV estimates. For the FGLS-based methods we set $k_{max} = 40$. The estimation results for 1-month forward rates ($h = 4$) are presented in Table 7, while those for 3-month forward rates ($h = 12$) are in Table 8.

In Section 6.1.2 we provided evidence that the GMM-GLS-IV and GLS-IV estimates are consistent requiring only pre-determined instruments w_t. We shall thus focus on the FGLS-based estimates. First note that both estimates are similar across all currencies and forecast horizons. For 1-month and 3-month forward rates, the FGLS-based estimates cannot reject the null hypothesis $H_0 : \alpha_0 = \beta = \alpha_1 = \alpha_2 = 0$ for any of the currencies. In contrast, OLS rejects H_0 for US-JP, for 1-month and 3-month forward rates. There is no consensus in the literature about the rejection of this null hypothesis. For 3-month forward rates ($h = 12$), Hansen and Hodrick (1980) using OLS rejects the null hypothesis for US-CAD and two other currencies (Deutsche mark and Swiss franc) for data between 1975 and 1979.

Note the relatively high standard errors of the GMM estimates for all the currencies. This suggests that the instruments w_t^j may be non exogenous. On the other hand, despite the fact that the first stage GMM estimate in the GMM-GLS-IV procedure is very noisy, the resulting GMM-GLS-IV estimate is much more efficient, and close to the GLS-IV estimate. It can be argued that while the GLS-IV estimate is valid with only pre-determined instruments, it might still be subject to a “weak instrument” problem; see, e.g., Andrews et al., 2019. We provide statistical evidence to argue that this is not the case. In particular, we consider the test of Staiger and Stock (1997) for weak instruments and the Wu-Hausman exogeneity test (Hausman, 1978 and Wu, 1973). The detailed implementation for the GLS-IV estimate are outlined in Appendix A.4. Note that standard weak instruments tests are valid for GLS-IV as the 2SLS regression (12) has uncorrelated errors. Table 9 reports the weak instruments test, the F_{SS} test statistic, see Section A.4.1, and the exogeneity test F_{WH}, see (A.5), together with the corresponding p-values. The null hypothesis for weak instruments is rejected for all currencies and $h = 4, 12$ at the 1% level of significance. The Wu-Hausman rejects the null hypothesis of exogeneity for US-UK and US-JP at least at the 5% level of significance for $h = 4, 12$ in all cases. Overall, these results indicate that we can be confident about the estimates and tests obtained using the GLS-IV procedure when applied to the Hansen and Hodrick (1980) regression (18).

7 Uncovering the OLS Bias

As shown in Section 2.1, if the error term in regression (1) is serially correlated at lags $q > k - 1$, the OLS estimator is not longer consistent. In this section we briefly present the results from applying the Cumby and Huizinga (1992) test (CH-test) for autocorrelation at lags $q > h - 1$ to both the Fama (1984) regression (13) and the Hansen and Hodrick (1980) regression (18). This test is well suited for our purpose since the null hypothesis is that the error process is a moving average of known order $q = h - 1 > 0$ against the general alternative that the autocorrelations are nonzero at lags greater than q. The CH-Test is a Wald test of the null hypothesis that the regression error is uncorrelated with itself at lags $q + 1$ through $q + s$. A general formulation for two-stage least squares and two-step two-stage least squares is presented in Cumby and Huizinga (1992). We require the errors u_t to be unconditionally homoskedastic. We refer to the paper by Cumby and Huizinga (1992) for the details about the implementation of the test, which require the estimate of several quantities. We simply note that for the Fama regression (13) we use FGLS residuals,

$$\hat{u}_{t+h,FLGS} = y_{t+h} - \hat{\alpha}_{FGLS} - \hat{\beta}_{FGLS} x_t,$$

where $y_{t+h} = s_{t+h} - s_t$ and $x_t = f_{t,h} - s_t$, whereas for the Hansen and Hodrick (1980) regression (18) we use GLS-IV residuals,

$$\hat{u}_{t+h,GLS-IV} = y_{t+h} - \hat{\alpha}_{0GLS-IV} - \hat{\beta}_{GLS-IV} y_t - \hat{\alpha}_{1GLS-IV} w_t^1 - \hat{\alpha}_{2GLS-IV} w_t^2,$$

where $y_{t+h} = s_{t+h}^i - f_t^i$ and $w_t^j = s_t^j - f_{t-h}^j$ for $j \neq i$. In Table 10, we provide the results of the CH-Test statistics, labelled $l_{q,s}$, of the null hypothesis that the regression error in the Fama regression (13) with 3-month forward rates is uncorrelated with itself at lags $q+1$ to $q+s$ with $q = h-1 = 11$ and $s = \{12, 15, 20\}$. For both sample periods, we obtain a rejection of the null hypothesis for every s and conclude that the error term in regression (13) is serially correlated at lags $q > 11$. Table 11 provide similar results for the Hansen and Hodrick (1980) regression (18). The specifications are similar except that we set $s = \{5, 10, 50\}$. Again, for both sample periods, we have statistical evidence to reject the null hypothesis for every s. Hence, again here the error term in regression (18) is also serially correlated at lags $q > 11$.

8 Concluding Remarks

We re-examined the statistical evidence about the hypothesis of Uncovered Interest Parity, which is a joint hypothesis of efficiency in the forward foreign exchange markets and rational expectations. Testing rationality hypothesis and exchange market efficiency is embedded in the general problem of estimating the parameters of a h-step-ahead linear forecasting equation. Under the null hypothesis, the forecast errors are serially correlated up to lags $h-1$ and OLS is consistent. However, if the errors are serially correlated beyond lags $h-1$, OLS is not longer consistent. This observation motivates using FGLS-based methods that are robust to the structure of the error process. We apply the FGLS procedure developed in Perron and González-Coya (2024) and we extend it to a setting with lagged dependent variables included as regressors. The resulting instrumental variables-based procedure, GLS-IV, is consistent requiring only pre-determined IVs. Using these FGLS methods, we study the main two UIP specifications in the literature: the Fama (1984) and Hansen and Hodrick (1980) regressions. We provide novel insights about the forward premium anomaly. Applying the consistent FGLS method we show that the estimates of β for three currencies are always non negative at the 1% significance level. A result that is contrary to the general finding that the OLS estimates are negative. Hence, the so-called "forward discount anomaly" is not as severe as previously thought. We also show statistically significant discrepancies between the OLS and GLS-IV estimates in the Hansen and Hodrick (1980) regression. We rationalize these discrepancies by showing that the regression residuals are in fact serially correlated beyond lags k, and thus OLS is not consistent, while the FGLS methods remain consistent. This point to the usefulness of adopting our more robust FGLS procedure. Not only is consistent and efficient under a wider range of contexts but, as we have shown, can deliver estimates that are different and point to a different assessment of the empirical facts.

References

Andrews, D. W. K. (1991). “Heteroskedasticity and autocorrelation consistent covariance matrix estimation.” *Econometrica*, 59, 817–858.

Andrews, I., Stock, J. H., and Sun, L. (2019). “Weak instruments in instrumental variables regression: Theory and practice.” *Annual Review of Economics*, 11, 727–753.

Backus, D. K., Gregory, A. W., and Telmer, C. I. (1993). “Accounting for forward rates in markets for foreign currency.” *The Journal of Finance*, 48, 1887–1908.

Baillie, R. T., Diebold, F. X., Kapetanios, G., and Kim, K. H. (2023). “A new test for market efficiency and uncovered interest parity.” *Journal of International Money and Finance*, 130, 102765.

Bekaert, G. and Hodrick, R. J. (1993). “On biases in the measurement of foreign exchange risk premiums.” *Journal of International Money and Finance*, 12, 115–138.

Berk, K. N. (1974). “Consistent autoregressive spectral estimates.” *The Annals of Statistics*, 2, 489–502.

Bilson, J. F. O. (1981). “The "‘speculative efficiency"’ hypothesis.” *The Journal of Business*, 54, 435–451.

Cumby, R. E. and Huizinga, J. (1992). “Testing the autocorrelation structure of disturbances in ordinary least squares and instrumental variables regressions.” *Econometrica*, 60, 185–195.

Durbin, J. (1970). “Testing for serial correlation in least-squares regression when some of the regressors are lagged dependent variables.” *Econometrica*, 38, 410–421.

Engel, C. (1996). “The forward discount anomaly and the risk premium: A survey of recent evidence.” *Journal of Empirical Finance*, 3, 123–192.

Fama, E. F. (1984). “Forward and spot exchange rates.” *Journal of Monetary Economics*, 14, 319–338.

Froot, K. A. (1990). “Short rates and expected asset returns.” *National Bureau of Economic Research Cambridge, Mass., USA*.

Griliches, Z. (1961). “A note on serial correlation bias in estimates of distributed lags.” *Econometrica*, 29, 65–73.

Hai, W., Mark, N. C., and Wu, Y. (1997). “Understanding spot and forward exchange rate regressions.” *Journal of Applied Econometrics*, 12, 715–734.

Hansen, L. P. (1982). "Large sample properties of generalized method of moments estimators." *Econometrica*, 50, 1029–1054.

Hansen, L. P. and Hodrick, R. J. (1980). "Forward exchange rates as optimal predictors of future spot rates: An econometric analysis." *Journal of Political Economy*, 88, 829–853.

Hausman, J. A. (1978). "Specification tests in econometrics." *Econometrica*, 46, 1251–1271.

Liviatan, N. (1963). "Consistent estimation of distributed lags." *International Economic Review*, 4, 44–52.

Malinvaud, E. (1966). *Statistical Methods of Econometrics.* Amsterdam: North-Holland.

Muth, J. F. (1960). "Optimal properties of exponentially weighted forecasts." *Journal of the American Statistical Association*, 55, 299–306.

Ng, S. and Perron, P. (2005). "A note on the selection of time series models." *Oxford Bulletin of Economics and statistics*, 67, 115–134.

Olivari, M. and Perron, P. (2024). "GLS-IV in time series regressions." *Working Paper, Department of Economics, Boston University.*

Perron, P. and González-Coya (2024). "Feasible GLS for time series regression." *Working Paper, Department of Economics, Boston University.*

Schwarz, G. (1978). "Estimating the dimension of a model." *The Annals of Statistics*, 6, 461–464.

Small, D. S. (2007). "Sensitivity analysis for instrumental variables regression with overidentifying restrictions." *Journal of the American Statistical Association*, 102, 1049–1058.

Staiger, D. and Stock, J. H. (1997). "Instrumental variables regression with weak instruments." *Econometrica*, 65, 557–586.

Wallis, K. F. (1967). "Lagged dependent variables and serially correlated errors: A reappraisal of three-pass least squares." *The Review of Economics and Statistics*, 49, 555–567.

Wu, D. M. (1973). "Alternative tests of independence between stochastic regressors and disturbances." *Econometrica*, 41, 733–750.

Table 1: Simulation results under EMH, 1-month forward rate. RE implies MA(3) errors; General implies ARMA(1,3) errors.

			MSE	Bias	Variance	Coverage	Length
11/83-01/23	RE	OLS	0.8152	0.7304	0.7598	0.94	3.3952
		FGLS	0.1801	0.3388	0.1919	0.96	1.7167
		Durbin	0.4950	0.5553	0.5057	0.95	2.7888
	General	OLS	4.9239	1.7974	4.4738	0.93	8.2112
		FGLS	0.1445	0.3068	0.1525	0.96	1.5311
		Durbin	0.4927	0.0035	0.4955	0.95	2.7607
01/89-04/21	RE	OLS	1.0063	0.7929	1.0019	0.94	3.8879
		FGLS	0.3108	0.4422	0.3347	0.95	2.2667
		Durbin	0.9838	0.7883	1.0471	0.96	4.0135
	General	OLS	6.2271	1.9730	5.8364	0.92	9.3351
		FGLS	0.2736	0.4212	0.2926	0.96	2.1200
		Durbin	0.9924	0.7929	1.0502	0.95	4.0196

We use US-UK observed data $x_t = f_{t,k} - s_t$, $T_1 = 2,050$, $T_2 = 1,600$.
For OLS we use HAC standard errors as described in the text.

Table 2: Simulation results under EMH, 3-month forward rate. RE implies MA(11) errors; General implies ARMA(1,11) errors.

			MSE	Bias	Variance	Coverage	Length
11/83-01/23	RE	OLS	1.6898	1.0541	1.5036	0.92	4.7404
		FGLS	0.7098	0.6761	0.7261	0.95	3.3387
		Durbin	1.3941	0.9358	1.4148	0.95	4.6647
	General	OLS	18.5744	3.4972	15.3260	0.91	15.0322
		FGLS	0.6019	0.6283	0.6332	0.96	3.1202
		Durbin	1.3964	0.9370	1.4193	0.95	4.6720
01/89-04/21	RE	OLS	1.8895	1.0869	1.6558	0.91	4.9454
		FGLS	0.9995	0.7817	0.9834	0.94	3.8838
		Durbin	1.9843	1.1502	2.0973	0.96	5.6801
	General	OLS	20.6180	3.5933	16.2122	0.88	15.3295
		FGLS	0.8825	0.7451	0.9027	0.95	3.7256
		Durbin	1.9782	1.1467	2.0995	0.96	5.6832

We use US-UK observed data $x_t = f_{t,k} - s_t$, $T_1 = 2,050$, $T_2 = 1,600$.
For OLS we use HAC standard errors as described in the text.

Table 3: Fama (1984) model estimation results, 1-month forward rate (SE in parenthesis).

		α			β		
		OLS	*Durbin*	FGLS	OLS	*Durbin*	FGLS
US-UK	11/1983-01/2023	-0.00167	-0.00031	-0.00006	-1.08291	-0.09390	0.08457
		(0.001)	(4.1E-04)	(3.3E-04)	(0.316)	(0.380)	(0.191)
	10/1989-04/2021	-0.00241	-0.00067	-0.00027	-1.32248	0.44659	-0.28232
		(0.002)	(0.001)	(0.001)	(0.359)	(0.781)	(0.464)
US-CA	12/1984-01/2023	-0.00015	-0.00000	0.00004	-0.24314	0.14125	0.20060
		(0.001)	(2.5E-04)	(2.1E-04)	(0.333)	(0.376)	(0.212)
	10/1989-04/2021	-0.00005	0.00000	0.00005	0.04296	0.49843	0.19765
		(0.002)	(0.001)	(0.001)	(0.384)	(0.652)	(0.375)
US-JP	09/1993-01/2023	-0.00165	-0.00018	0.00005	-0.08744	0.26966	0.17845
		(0.002)	(0.001)	(3.3E-04)	(0.110)	(0.145)	(0.081)
	09/1993-04/2021	-0.00101	-0.00027	0.00011	-0.13273	0.14601	0.31225
		(0.003)	(0.001)	(0.001)	(0.110)	(0.282)	(0.152)

Table 4: Fama (1984) model estimation results, 3-month forward rate (SE in parenthesis).

		α			β		
		OLS	*Durbin*	FGLS	OLS	*Durbin*	FGLS
US-UK	11/1983-01/2023	-0.00439	-0.00035	0.00001	-0.98614	0.64187	0.29696
		(0.004)	(4.1E-04)	(3.3E-04)	(0.218)	(0.382)	(0.278)
	10/1989-04/2021	-0.00011	-0.00011	0.00018	0.34508	1.25038	1.15847
		(0.004)	(0.001)	(0.001)	(0.238)	(0.474)	(0.332)
US-CA	12/1984-01/2023	-0.00036	0.00004	0.00006	-0.21093	0.51211	0.31376
		(0.003)	(2.5E-04)	(2.1E-04)	(0.218)	(0.343)	(0.261)
	10/1989-04/2021	-0.00040	0.00002	-0.00002	0.19149	0.55918	0.31524
		(0.003)	(0.001)	(0.001)	(0.252)	(0.399)	(0.301)
US-JP	09/1993-01/2023	-0.00355	-0.00017	-0.00015	-0.10423	0.45446	0.30124
		(0.005)	(0.001)	(3.3E-04)	(0.138)	(0.152)	(0.099)
	09/1993-04/2021	-0.00158	-0.00003	-0.00002	-0.26883	0.49440	0.32649
		(0.005)	(0.001)	(0.001)	(0.139)	(0.157)	(0.101)

Table 5: Simulation results with exogenous instruments. RE implies MA(3) errors; General implies ARMA(1,3) errors.

		MSE	Bias	Variance	Coverage	Length
RE	OLS	0.19	3.43	0.17	0.93	0.16
	GMM	0.21	3.66	0.08	0.77	0.11
	GMM-GLS-IV	0.14	2.96	0.13	0.94	0.14
	GLS-IV	0.15	2.99	0.12	0.93	0.13
General	OLS	3.66	18.05	0.37	0.19	0.24
	GMM	0.79	6.98	0.31	0.75	0.22
	GMM-GLS-IV	0.09	2.39	0.11	0.96	0.12
	GLS-IV	0.10	2.42	0.11	0.95	0.12

Weekly data for US-CAD, US-JP for the period November 2010 to April 2020 ($T = 492$). For OLS we use HAC standard errors as described in the text.

Table 6: Simulation results with non exogenous instruments. RE implies MA(3) errors; General implies ARMA(1,3) errors.

		MSE	Bias	Variance	Coverage	Length
RE	OLS	0.32	4.63	0.25	0.89	0.19
	GMM	0.37	4.84	0.19	0.82	0.17
	GMM-GLS-IV	0.17	3.28	0.13	0.92	0.14
	GLS-IV	0.23	3.74	0.13	0.87	0.14
General	OLS	3.96	18.46	0.47	0.26	0.26
	GMM	1.67	10.63	0.56	0.65	0.26
	GMM-GLS-IV	0.19	3.47	0.15	0.91	0.15
	GLS-IV	0.26	3.71	0.15	0.89	0.14

Weekly data for US-CAD, US-JP for the period November 2010 to April 2020 ($T = 492$). For OLS we use HAC standard errors as described in the text.

Table 7: Hansen and Hodrick (1980) 1-month forward model estimation results. (SE in parenthesis).

	α_0				β			
	OLS	GMM	GMM-GLS-IV	GLS-IV	OLS	GMM	GMM-GLS-IV	GLS-IV
US-UK	-0.000	0.001	0.000	0.000	-0.087	0.044	1.436	0.392
	(0.001)	(0.002)	(0.001)	(0.001)	(0.040)	(3.875)	(1.439)	(0.278)
US-CAD	-0.000	0.000	-0.000	-0.000	-0.035	0.247	0.040	1.223
	(0.001)	(0.001)	(0.001)	(0.001)	(0.051)	(1.012)	(0.631)	(0.959)
US-JP	0.003	0.004	0.001	0.007	0.120	0.024	-0.039	-0.534
	(0.002)	(0.022)	(0.001)	(0.005)	(0.054)	(6.181)	(0.598)	(1.060)
	α_1				α_2			
US-UK	0.143	0.029	-0.442	-0.081	0.016	0.025	-0.112	-0.037
	(0.059)	(1.650)	(0.435)	(0.135)	(0.045)	(0.599)	(0.133)	(0.039)
US-CAD	0.012	-0.096	0.053	-0.455	0.031	0.050	0.023	0.034
	(0.046)	(0.280)	(0.162)	(0.358)	(0.037)	(0.074)	(0.024)	(0.031)
US-JP	-0.056	-0.042	-0.045	0.050	0.015	0.013	-0.027	0.028
	(0.063)	(1.262)	(0.092)	(0.187)	(0.058)	(0.421)	(0.044)	(0.045)

For US-UK, α_1 is the coefficient for US-CAD and α_2 is the coefficient for US-JP;
for US-CAD, α_1 is the coefficient for US-UK and α_2 is the coefficient for US-JP;
for US-JP, α_1 is the coefficient for US-UK and α_2 is the coefficient for US-CAD.

Table 8: Hansen and Hodrick (1980) 3-month forward model estimation results. (SE in parenthesis).

	α_0				β			
	OLS	GMM	GMM-GLS-IV	GLS-IV	OLS	GMM	GMM-GLS-IV	GLS-IV
US-UK	-0.001	0.000	-0.000	-0.001	0.091	0.937	0.536	0.376
	(0.004)	(0.004)	(0.000)	(0.001)	(0.077)	(2.486)	(0.424)	(0.212)
US-CAD	0.000	0.000	-0.000	-0.000	0.005	0.103	1.063	0.594
	(0.003)	(0.004)	(0.001)	(0.001)	(0.071)	(1.379)	(0.604)	(0.206)
US-JP	-0.001	-0.002	0.000	0.002	0.104	0.112	1.090	2.349
	(0.004)	(0.006)	(0.001)	(0.004)	(0.062)	(0.703)	(0.919)	(1.372)
	α_1				α_2			
US-UK	0.055	-0.378	-0.102	-0.104	-0.026	-0.146	-0.067	-0.054
	(0.090)	(1.028)	(0.134)	(0.123)	(0.059)	(0.408)	(0.043)	(0.027)
US-CAD	0.026	-0.044	-0.290	-0.239	0.097	0.088	0.017	0.085
	(0.070)	(0.420)	(0.158)	(0.097)	(0.046)	(0.130)	(0.037)	(0.023)
US-JP	-0.051	-0.038	-0.193	-0.342	0.064	0.111	-0.074	-0.062
	(0.106)	(0.254)	(0.124)	(0.199)	(0.108)	(0.219)	(0.059)	(0.128)

For US-UK, α_1 is the coefficient for US-CAD and α_2 is the coefficient for US-JP;
for US-CAD, α_1 is the coefficient for US-UK and α_2 is the coefficient for US-JP;
for US-JP, α_1 is the coefficient for US-UK and α_2 is the coefficient for US-CAD.

Table 9: Weak instruments (first row) and Wu-Hausman (second row) tests for GLS-IV.

		df_1	df_2	Statistic	p-value
1-month	US-UK	72	1309	13.02	2E-16
		1	1379	7.91	0.00499
	US-CAD	78	1299	7.785	2E-16
		1	1375	4.022	0.044
	US-JP	48	1349	17.90	2E-16
		1	1395	30.18	4.67E-08
3-month	US-UK	75	1288	22.75	2E-16
		1	1361	56.27	1.13E-13
	US-CAD	75	1288	24.689	2E-16
		1	1361	10.67	0.001
	US-JP	39	1348	18.745	2E-16
		1	1385	9.066	0.00265

Table 10: Cumby and Huizinga (1992) test of the null hypothesis that the Fama regression error is uncorrelated with itself at lags $q+1$ to $q+s$, $q=11$.

		US-UK	US-CA	US-JP
10/1984-01/2023	12	196.499	321.282	223.436
	15	70.894	48.850	61.324
	20	57.016	47.526	44.952
10/1989-04/2021	12	152.768	280.233	642.996
	15	34.919	48.315	29.323
	20	29.680	75.373	54.568

Table 11: Cumby and Huizinga (1992) test of the null hypothesis that the Hansen-Hodrick regression error is uncorrelated with itself at lags $q+1$ to $q+s$, $q=11$.

		US-UK	US-CA	US-JP
10/1984-01/2023	5	1087.218	203.5084	646.771
	10	1134.685	219.5272	624.1312
	50	525.2874	330.3852	1115.21
10/1989-04/2021	5	48811.75	43522.55	61825.54
	10	114112.9	424805.4	65447.43
	50	93889.12	377855.8	37567.25

Figure 1: Empirical Rejection Frequencies of Nominal 5% t-Test of $H_0 : \beta = 1$. Fama regression, RE case with MA(3) errors.

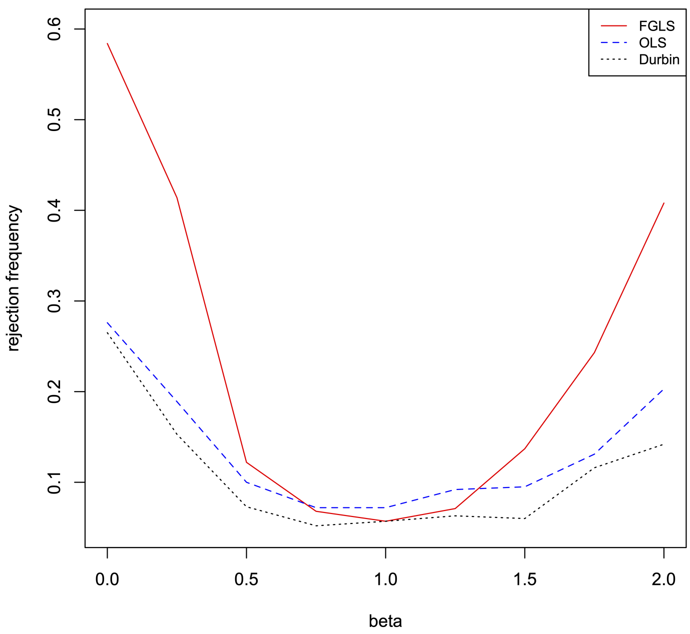


Figure 2: Empirical Rejection Frequencies of Nominal 5% t-Test of $H_0 : \beta = 1$. Fama regression, General case with ARMA(1,3) errors.

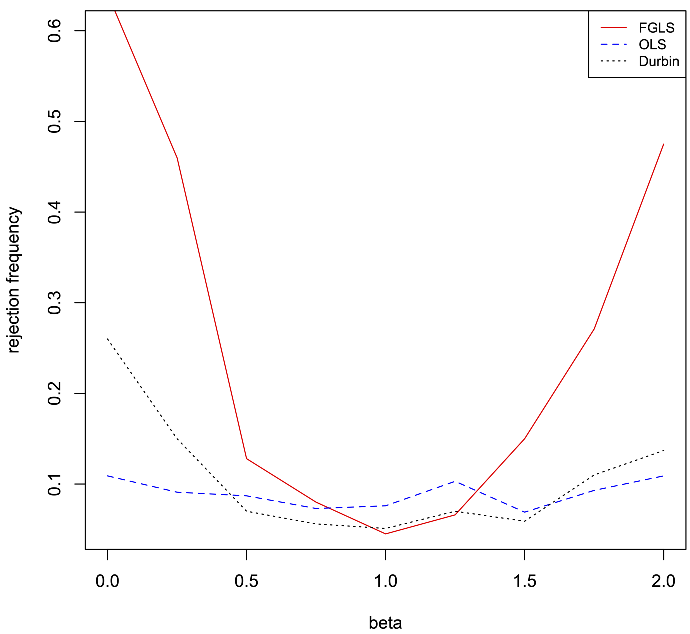


Figure 3: Empirical Rejection Frequencies of Nominal 5% t-Test of $H_0 : \beta = 0$. Hansen and Hodrick regression, RE case with MA(3) errors.

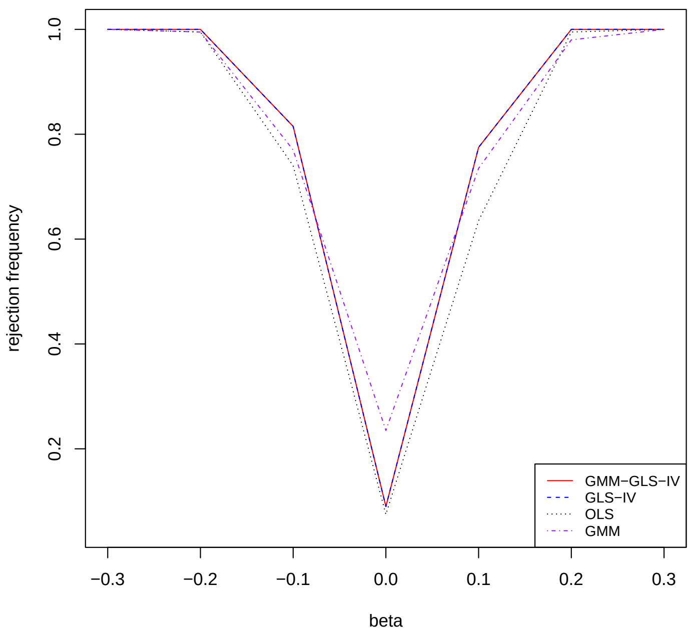


Figure 4: Empirical Rejection Frequencies of Nominal 5% t-Test of $H_0 : \beta = 0$. Hansen and Hodrick regression, General case with ARMA(1,3) errors.

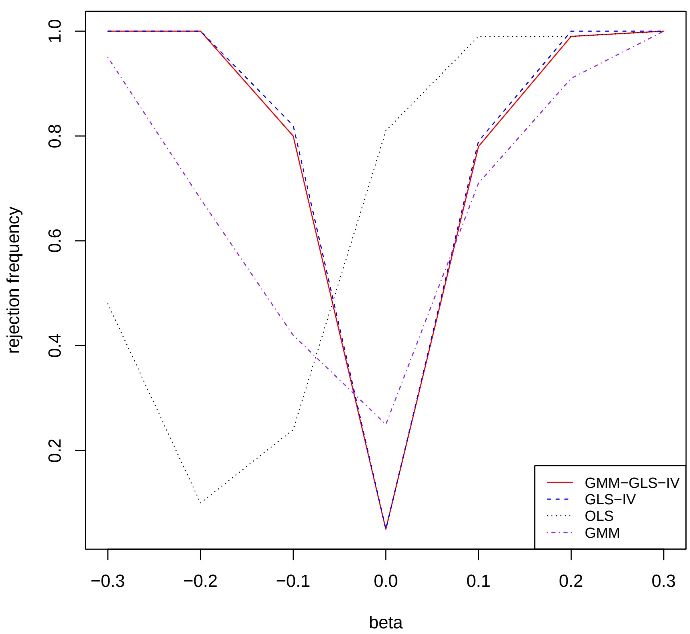


Figure 5: Empirical Rejection Frequencies of Nominal 5% t-Test of $H_0 : \beta = 0$. Non-exogenous IVs, RE case with MA(3) errors.

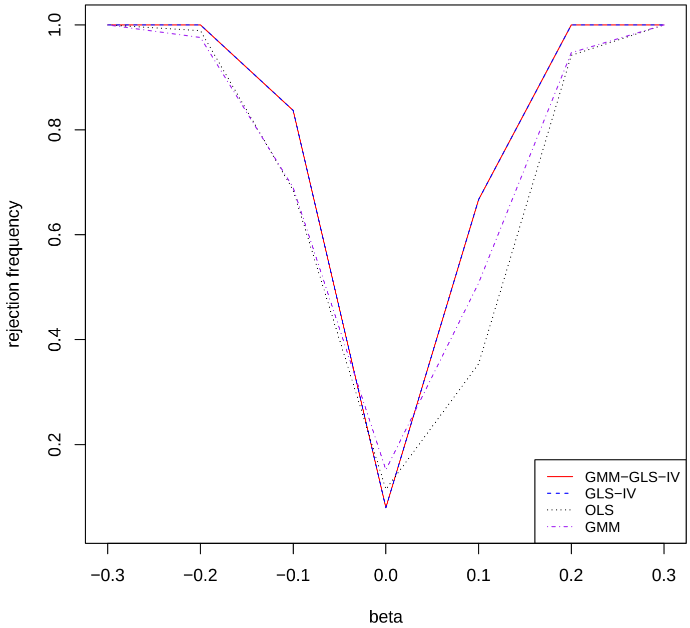
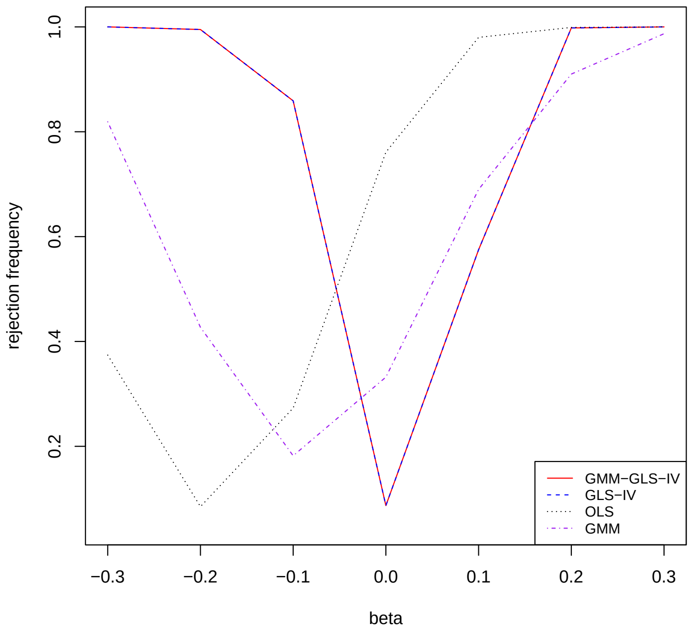


Figure 6: Empirical Rejection Frequencies of Nominal 5% t-Test of $H_0 : \beta = 0$. Non-exogenous IVs, General case with ARMA(1,3) errors.



Appendix

A.1 Proof of Remark 2

Assume that y_t has a permanent and a transitory component, $y_t = \bar{y}_t + \eta_t$, where the permanent component is defined by $\bar{y}_t = \bar{y}_{t-1} + \eta_t = \sum_{i=1}^{t} \varepsilon_i$, where $\varepsilon_i \sim i.i.d(0, \sigma_\varepsilon^2)$, $\eta_i \sim i.i.d(0, \sigma_\eta^2)$ and ε_i, η_i are assumed to be independent. Then we can write

$$u_t = \left(\sum_{i=1}^{t} \varepsilon_i + \eta_t\right) - \beta \sum_{i=1}^{\infty} (1-\beta)^{i-1} \left(\sum_{j=1}^{t-i} \varepsilon_j + \eta_{t-i}\right).$$

Hence,

$$\begin{aligned} u_t - (1-\beta)u_{t-1} &= \left(\sum_{i=1}^{t} \varepsilon_i + \eta_t\right) - (1-\beta)\left(\sum_{i=1}^{t-1} \varepsilon_i + \eta_{t-1}\right) - \beta \left(\sum_{i=1}^{t-1} \varepsilon_i + \eta_{t-1}\right) \\ &= \varepsilon_t + \eta_t - \eta_{t-1}. \end{aligned}$$

Thus, we can write the forecast error as $u_t = (1-\beta)u_{t-1} + v_t$, where $v_t = \varepsilon_t + \eta_t - \eta_{t-1}$. Note that ε_t and η_t can be allowed to be correlated, so that v_t is invertible in general.

A.2 Efficient Estimate of Autoregressive Coefficients for GLS-IV

Suppose that $n = 2$. The efficient linear combination $\widetilde{\rho}_j^* = \lambda \widetilde{\rho}_{1j} + (1-\lambda)\widetilde{\rho}_{2j}$ is obtained when λ minimizes $\mathrm{Var}(\widetilde{\rho}_j^*)$. Hence, the the minimization problem is $\min_{\lambda_1,\lambda_2} \mathrm{Var}\left(\lambda_1 \widetilde{\rho}_{1j} + \lambda_2 \widetilde{\rho}_{2j}\right)$, subject to $\lambda_1 + \lambda_2 = 1$.The first-order conditions of this problem are:

$$2\lambda_1 \mathrm{Var}\left(\widetilde{\rho}_{1j}\right) + 2\lambda_2 \mathrm{Cov}\left(\widetilde{\rho}_{1j}, \widetilde{\rho}_{2j}\right) = \gamma 2\lambda_2 \mathrm{Var}\left(\widetilde{\rho}_{2j}\right) + 2\lambda_2 \mathrm{Cov}\left(\widetilde{\rho}_{1j}, \widetilde{\rho}_{2j}\right) = \gamma \lambda_1 + \lambda_2 = 1,$$

where γ is the Lagrangian multiplier. The solution is

$$\lambda_1 = \frac{\mathrm{Var}\left(\widetilde{\rho}_{2j}\right) - \mathrm{Cov}\left(\widetilde{\rho}_{1j}, \widetilde{\rho}_{2j}\right)}{\mathrm{Var}\left(\widetilde{\rho}_{1j}\right) + \mathrm{Var}\left(\widetilde{\rho}_{2j}\right) - 2\mathrm{Cov}\left(\widetilde{\rho}_{1j}, \widetilde{\rho}_{2j}\right)}.$$

A.3 Simulation Design

We simulate an $MA(h-1)$ process with parameters calibrated to replicate the observed autocorrelation function up to lag $h-1$ of the FGLS residuals of regression (13) using US-UK data for the Fama regression (Section 5) and using GLS-IV residuals of regression (18) for the Hansen and Hodrick (1980) regression (Section 6). We obtain the parameters $\theta_1, \ldots, \theta_{k-1}$ by solving the non-linear system of $h-1$ equations given by the $MA(h-1)$ autocorrelation functions (ACF) for $j = 1, \ldots, h-1$. Let v_t be an $MA(q)$ process with $q = h-1$, then the variance of v_t is

$$\gamma_0 = \left(1 + \theta_1^2 + \theta_1^2 + \cdots + \theta_q^2\right) \sigma_\varepsilon^2.$$

The autocovariance function of v_t is

$$\gamma_j = \begin{cases} \left(\theta_j + \theta_{j+1}\theta_1 + \theta_{j+2}\theta_2 + \cdots + \theta_q \theta_{q-j}\right) \sigma_\varepsilon^2, & \text{for } j = 1, \ldots, q, \\ 0, & \text{for } j > q. \end{cases}$$

Thus, the autocorrelation function of v_t is

$$\psi_j = \begin{cases} \frac{\theta_j + \theta_{j+1}\theta_1 + \theta_{j+2}\theta_2 + \cdots + \theta_q \theta_{q-j}}{1 + \theta_1^2 + \theta_1^2 + \cdots + \theta_q^2}, & \text{for } j = 1, \ldots, q, \\ 0, & \text{for } j > q. \end{cases}$$

Using the observed first q autocorrelations of the residuals of the regression, ACF_j ($j = 1, \ldots, q$) we define a system of q non-linear equations with q unknown variables $\hat{\theta}_1, \ldots, \hat{\theta}_q$:

$$\begin{aligned} \frac{\hat{\theta}_1 + \hat{\theta}_2\hat{\theta}_1 + \hat{\theta}_3\hat{\theta}_2 + \cdots + \hat{\theta}_q\hat{\theta}_{q-1}}{1 + \hat{\theta}_1^2 + \hat{\theta}_1^2 + \cdots + \hat{\theta}_q^2} &= ACF_1 \\ \frac{\hat{\theta}_2 + \hat{\theta}_3\hat{\theta}_1 + \hat{\theta}_3\hat{\theta}_2 + \cdots + \hat{\theta}_q\hat{\theta}_{q-2}}{1 + \hat{\theta}_1^2 + \hat{\theta}_1^2 + \cdots + \hat{\theta}_q^2} &= ACF_2 \\ &\vdots \\ \frac{\hat{\theta}_q}{1 + \hat{\theta}_1^2 + \hat{\theta}_1^2 + \cdots + \hat{\theta}_q^2} &= ACF_q \end{aligned}$$

A.4 Diagnostic Tests for GLS-IV

Rewrite the model using the quasi-differenced variables y_t^*, y_{t-k}^* and w_t^* as

$$y = Y\beta + W\gamma + \varepsilon \tag{A.1}$$

$$Y = Z\Pi + W\Phi + V \tag{A.2}$$

where (A.1) is the structural equation of interest, $y = (y_{k+1}^*, \ldots, y_T^*)$ is a $(T-k) \times 1$ vector and $Y = [y_{-k}^*]$ is a $(T-k) \times 1$ vector with the lagged dependent variable (the only endogenous variable in the model). (A.2) is the reduced form equation for Y, W is the $(T-k) \times K_1$ matrix of exogenous regressors with row t, $W_t = [1, w_t^{1*}, w_t^{2*}]$ and $K_1 = 3$ (i.e. W includes a constant term). Z is the $(T-k) \times K_2$ matrix of quasi-differenced instruments with row t, $Z_t = \{w_t^{j*}, w_{t-k}^{j*}, j \neq i\}$ and $K_2 = 4$. ε and V are, respectively, a $(T-k) \times 1$ vector and a $(T-k) \times 2$ matrix of error terms. Note that the quasi-differenced regression (A.1) has serially uncorrelated errors, with covariance matrix Σ. Assume that $E[\varepsilon_t^2] = \sigma_{\varepsilon\varepsilon}$, $E[V_t\varepsilon_t] = \Sigma_{V\varepsilon}$, and $E[V_tV_t'] = \Sigma_{VV}$. Let $\bar{Z} = [X, Z]$, it is assumed throughout that $E[\bar{Z}_t\,(u_tV_t')] = 0$.

A.4.1 Staiger and Stock (1997) Weak Instruments Test

We are interested in testing $\Pi = 0$ in the regression (A.2). Π shall be modeled as local to zero, so that the F statistic is $O_p(1)$. Staiger and Stock (1997) make the assumption that $\Pi = \Pi_T = c/\sqrt{T-k}$ where c is a fixed $K_2 \times 1$ vector. Before proceeding we provide some additional definitions and notation. Let $Q = E[\bar{Z}_t\bar{Z}_t']$, partitioned so that $E[W_tW_t'] = Q_{WW}$, $E[W_tZ_t'] = Q_{WZ}$, and $E[Z_tZ_t'] = Q_{ZZ}$. Also let $\rho = \Sigma_{VV}^{-1/2\prime}\Sigma_{V\varepsilon}\sigma_{\varepsilon\varepsilon}^{-1/2}$. Let $P_R = R\,(R'R)^{-1}\,R'$ and $M_R = I - P_R$ where R is a general $a \times b$ matrix with $a \geq b$, and let "$\perp$" denote the residuals from the projection on W, so $Z^{\perp} = M_W Z$, $Y^{\perp} = M_W Y$, etc. Let $\bar{W} = [Y, W]$ and $\bar{Y} = [y, Y]$ and let I_k denote the k-dimensional identity matrix. Staiger and Stock (1997) assume that the following limits hold jointly: 1) $(\varepsilon'\varepsilon/T, V'\varepsilon/T, V'V/T) \xrightarrow{p} (\sigma_{\varepsilon\varepsilon}, \Sigma_{V\varepsilon}, \Sigma_{VV})$; 2) $T^{-1}\bar{W}'\bar{W} \xrightarrow{p} Q$; 3)

$$\left(T^{-1/2}W'\varepsilon, T^{-1/2}Z'\varepsilon, T^{-1/2}W'V, T^{-1/2}Z'V\right) \Rightarrow \left(\Psi_{W\varepsilon}, \Psi_{Z\varepsilon}, \Psi_{WV}, \Psi_{ZV}\right),$$

where $\Psi \equiv \Psi'_{W\varepsilon}, \Psi'_{Z\varepsilon}, \operatorname{vec}(\Psi_{WV})'$, with $\Psi_{WV} \sim N(0, \Sigma \otimes Q)$. Define $\lambda = \Omega^{1/2} C \Sigma_{VV}^{-1/2}$, where $\Omega = Q_{ZZ} - Q_{ZX} Q_{XX}^{-1} Q_{XZ}$,

$$z_u = \Omega^{-1/2\prime}\left(\Psi_{Zu} - Q_{ZX}Q_{XX}^{-1}\Psi_{Xu}\right)\sigma_{uu}^{-1/2}$$

and $z_V = \Omega^{-1/2}\left(\Psi_{ZV} - \ Q_{ZX}Q_{XX}^{-1}\Psi_{XV}\right)\Sigma_{VV}^{-1/2}$. The random variable $\left[z_u' \operatorname{vec}(z_V)'\right]'$ is distributed $N(0, \bar{\Sigma} \otimes \ I_{K_2})$, where $\bar{\Sigma}$ is the $(n+1)\times(n+1)$ matrix with $\bar{\Sigma}_{11} = 1, \bar{\Sigma}_{22} = I_n, \bar{\Sigma}_{12} = \rho'$, and $\bar{\Sigma}_{21} = \rho$, where $\bar{\Sigma}$ is partitioned conformably with Σ. Finally, let

$$\nu_1 = (\lambda + z_V)'(\lambda + z_V) \tag{A.3}$$

and

$$\nu_1 = (\lambda + z_V)' z_u \tag{A.4}$$

The 2SLS estimate of $(\beta', \gamma')'$ is

$$\left(\hat{\beta}(k)'\hat{\gamma}(k)'\right)' = \left(\bar{X}'\left(I - M_{\bar{Z}}\right)\bar{X}\right)^{-1}\left(\bar{X}'\left(I - M_{\bar{Z}}\right)y\right).$$

By standard projection arguments, the 2SLS estimate of β is

$$\hat{\beta}(k) = \left(Y^{\perp\prime}\left(I - M_{Z^{\perp}}\right)Y^{\perp}\right)^{-1}\left(Y^{\perp\prime}\left(I - M_{z^{\perp}}\right)y^{\perp}\right).$$

The Wald statistic testing $\Pi = 0$ is $W = \operatorname{tr}(G_T)/K_2$, where $G_T = \hat{\Sigma}_{VV}^{-1/2\prime} Y^{\perp\prime} P_{Z^{\perp}} Y^{\perp} \hat{\Sigma}_{VV}^{-1/2}$, with $\hat{\Sigma}_{VV} = Y' M_{\bar{Z}} Y / (T - K_1 - K_2)$. Staiger and Stock (1997) show that the limit distribution of G_T is ν_1 defined in (A.3). As we just have one endogenous variable, y_{t-k}, the F statistic $F_{SS} = G_T/K_2$ converges to a non central $\chi^2_{K_2}/k_2$ with noncentrality parameter $\lambda'\lambda$. In the general case, with more than one endogenous variable, $\lambda'\lambda$ is the matrix of noncentrality parameters of the limiting noncentral Wishart random variable ν_1.

A.4.2 Wu-Hausman Test of Exogeneity

The Wu-Hausman (WH) test (see Hausman, 1978 and Wu, 1973) examines the null hypothesis that Y is exogenous (i.e., $p = 0$) by checking for a statistically significant difference between the OLS and 2SLS estimates of β. The test statistic is

$$F_{\mathrm{WH}} = \left(\hat{\beta}_{\mathrm{2SLS}} - \hat{\beta}_{\mathrm{OLS}}\right)' V^{-1}\left(\hat{\beta}_{\mathrm{2SLS}} - \hat{\beta}_{\mathrm{OLS}}\right)$$

with

$$V = \left[\left(Y^{\perp\prime}P_{Z^{\perp}}Y^{\perp}\right)^{-1} - \left(Y^{\perp\prime}Y^{\perp}\right)^{-1}\right]\hat{\sigma}_{uu,2SLS}$$

Its limit distribution is

$$F_{\mathrm{WH}} \Rightarrow \frac{\left[\Delta_0^*(0)-\rho\right]' \nu_1 \left[\Delta_0^*(0)-\rho\right]}{S_1\left(\Delta_0^*(0)\right)}$$

where $\Delta_0^*(0) = \nu_1^{-1}\nu_2$ and $S_1(b) = 1 - 2\rho' b + b'b$. Under the null hypothesis $\rho = 0$, F_{WH} simplifies to

$$F_{\mathrm{WH}} \Rightarrow \frac{\zeta'\zeta}{\left(1+\zeta'\nu_1^{-1}\zeta\right)} \tag{A.5}$$

where $\zeta = \nu_1^{-1/2\prime}\left(\lambda + z_V\right)'\eta \sim N\left(0, I_n\right)$ with $\eta = (z_u - z_V\rho)/\sqrt{1-\rho'\rho}$ and ζ and ν_1 are independent. Note that since $\zeta'\zeta/\left(1+\zeta'\nu_1^{-1}\zeta\right) \le \zeta'\zeta \sim \chi_n^2$, applying χ_n^2 critical values to F_{WH} results in asymptotically conservative tests. However, as noted by Staiger and Stock (1997), a size adjustment of F_{WH} is infeasible because the distribution depends on $\lambda'\lambda/K_2$.

A.4.3 Testing for Serial Correlation at lags $q > k-1$

We describe in some details the test of Cumby and Huizinga (1992) for autocorrelation structure of the OLS residuals (CH-Test). This test is perfectly suited for our purpose as it allows to have under the null hypothesis a regression error process with a moving average of known order $q = k-1 > 0$ against the general alternative that the autocorrelations of the regression error are nonzero at lags greater than q. The CH-Test is a Wald test of the null hypothesis that the regression error is uncorrelated with itself at lags $q+1$ through $q+s$. Consider the general formulation of the CH-Test for an OLS regression. Here we present the Cumby and Huizinga (1992) test for autocorrelation structure of the OLS residuals of equation (1) with no instrumental variables. A general formulation for two-stage least squares and two-step two-stage least squares is presented in Cumby and Huizinga (1992). The model is

$$y_t = X_t\beta + u_t \tag{A.6}$$

where X_t is a vector of the n scalar predetermined regressors. The regression errors, u_t are assumed to be serially correlated up to a known lag $q \ge 0$ and their autocorrelations at all lags greater than q are required to be zero under the null hypothesis. We require the errors u_t to be unconditionally homoskedastic. The CH-Test statistic is

$$l_{q,s} = T\hat{r}'\left[\hat{V}_r + \hat{B}\hat{V}_d\hat{B}' + \hat{C}\hat{D}'\hat{B}' + \hat{B}\hat{D}\hat{C}'\right]^{-1}\hat{r} \sim \chi^2(s). \tag{A.7}$$

where $\hat{V}_r, \hat{B}, \hat{V}_d, \hat{C}, \hat{D}$ are consistent estimates of V_r, B, V_d, C, D. Here, r is a $s \times 1$ vector $r = \left[r_{q+1}, r_{q+2}, \ldots, r_{q+s}\right]'$,

$$r_j = \frac{\sum_{t=j+1}^{T} u_t u_{t-j}}{\sum_{t=1}^{T} u_t^2}.$$

V_d is the asymptotic covariance matrix of the estimator $\hat{\beta}$,

$$V_d = (X'X)^{-1}X'\Omega X(X'X)^{-1}$$

with $\Omega = \lim_{T\to\infty} T^{-1}E[uu']$. Let D be the $k \times k$ matrix $D = p\lim_{T\to\infty} T(X'X)^{-1}$. B is the $s \times k$ matrix with i,jth element

$$B(i,j) = -\left[E\left(u_{t-q-i}X_{j,t}\right) + E\left(u_t X_{j,t-q-i}\right)\right]/E\left(u_t^2\right).$$

Let $\xi_{i,t} = u_t u_{t-q-i}$ for $i = 1, \ldots, s, \omega_{j,t} = u_t X_{j,t}$ for $j = 1, \ldots, k$, the ijth element of the $s \times s$ matrix V_r be given by $V_r(i,j) = \sigma_u^{-4} \sum_{n=-q}^{q} E\left(\xi_{i,t}\xi_{j,t-n}\right)$, and the ij th element of the $s \times h$ matrix C be given by $C(i,j) = \sigma_u^{-2} \sum_{n=-q}^{q} E\left(\xi_{i,t}\omega_{j,t-n}\right)$. In order to consistently estimate the test statistic $l_{q,s}$, we need consistent estimate of the errors u_t. For the Fama regression (13) we use FGLS residuals,

$$\hat{u}_{t+k,FLGS} = y_{t+k} - \hat{\alpha}_{FGLS} - \hat{\beta}_{FGLS} x_t$$

where $y_{t+k} = s_{t+k} - s_t$ and $x_t = f_{t,k} - s_t$. Whereas for the Hansen and Hodrick (1980) regression (18) we use GLS-IV residuals,

$$\hat{u}_{t+k,IV-FLGS} = y_{t+k} - \hat{\alpha}_{0GLS-IV} - \hat{\beta}_{GLS-IV} y_t - \hat{\alpha}_{1GLS-IV} w_t^1 - \hat{\alpha}_{2GLS-IV} w_t^2$$

where $y_{t+k} = s_{t+k}^i - f_t^i$ and $w_t^j = s_t^j - f_{t-k}^j$ for $j \neq i$. We estimate Ω as discussed in Section 4.1.1.