Are saving and investment cointegrated? The case of Malaysia (1965–2003)

James B. Ang

The Australian National University and Monash University
E-mail: james.ang@buseco.monash.edu.au

This article examines whether domestic saving rate leads to higher domestic investment rate in the case of Malaysia. We argue that the results obtained from cross-sectional studies are not able to address this issue satisfactorily and highlight the importance of individual country case studies. Using the recently developed autoregressive distributed lag bounds testing procedure, the results reveal a robust cointegrated relationship between domestic saving and investment rates during the period 1965 to 2003.

I. Introduction

An understanding of the relationship between saving and investment provides an important insight into the process of economic development. This is because economic growth critically depends on capital accumulation and capital accumulation stems from investment which depends on domestic and foreign capital. Hence, increased saving leads to higher economic growth through capital formation.

The relationship between saving and investment has been the subject of intense research over the past 2 decades. In a seminal study, Feldstein and Horioka (1980) examine the extent of correlation between saving and investment across 16 organization for economic cooperation and development (OECD) countries. They argue that there should be no relationship between a country’s domestic saving and its domestic investment in the presence of perfect capital mobility. Extra saving in any country will be channelled to the world capital market to fund other countries with favourable investment climate. Using a cross-sectional analysis, they show that 85–95% of domestic savings are transformed into investment in the domestic economies. Furthermore, the regression coefficient of saving on domestic investment is statistically not different from one, indicating that international capital mobility is rather low. This observed phenomenon is widely known as the Feldstein–Horioka puzzle.

There are two main strands of literature that attempt to resolve the Feldstein–Horioka puzzle. The first strand of literature, which is in line with the Feldstein–Horioka interpretation, argues that higher correlation between saving and investment implies greater international capital immobility. Using a cross-sectional framework, the findings of Feldstein (1983), Penati and Dooley (1984) Dooley et al. (1987), and Vos (1988), among others, suggest a close relationship between domestic saving and investment rates. On the other hand, time series studies in this strand of literature examine the dynamic saving-investment nexus over time and across different exchange rate and capital control regimes (Miller, 1988; Alexakis and Apergis, 1994; Ho, 1999; De Vita and Abbott, 2002; Ozmen and Parmaksiz, 2003; Schmidt, 2003; Narayan, 2005). The results of these empirical studies are mixed.

The second strand of literature, which departs from the Feldstein–Horioka approach, has tried to argue that the saving-investment correlation is due to other macroeconomic factors such as country
size (Baxter and Crucini, 1993), current account solvency (Coakley et al., 1996), financial structure (Kasuga, 2004) and nontraded goods (Murphy, 1986; Wong, 1990). Although various theoretical explanations have been proposed to account for a strong saving-investment correlation in the presence of high capital mobility, the empirical results remain ambiguous.

Solow (2001) argues that an economic model should be dynamic in nature so that it can explain the evolution of the economic behaviour observed over time. Since the saving-investment nexus is largely determined by the nature and operation of the financial institutions and economic policies pursued in each country, it is more appropriate to carry out country-specific studies by examining the evolution of the variables of interest over time and relate the findings to policy designs.

A major constraint impeding research on the dynamic relationship between saving and investment until recently has been the lack of sufficient time series data for developing countries. As a result, cross-sectional studies have dominated the literature. However, results derived from the cross-sectional estimations are only valid in an average sense. Such cross-sectional empirical analyses conducted at the aggregate level are unable to capture and account for the complexity of the financial environments and economic histories of each individual country. Hence, any inference drawn from these studies provides only a general understanding of how these two variables are broadly related and may be subject to aggregation bias. As Athukorala and Sen (2002) argue, cross-country studies are based on a restrictive assumption of ‘homogeneity’ in the observed relationship across countries to which there are always exceptions. Since the nature and quality of data vary significantly across countries, cross-country comparison is unlikely to yield sensible results (Deaton, 1989). These considerations point to the need of more country-specific in-depth studies. Unfortunately, research on the saving-investment link has largely focused on OECD countries. Systematic time series case studies on developing countries in this area are far and few in between.

In this article, we re-examine the long-run relationship between saving and investment in the case of Malaysia, which has undergone a fundamental reform in its financial sector. The choice of this country is motivated by the fact that research on this issue for Malaysia has been very limited. In addition, the database for Malaysia is considered relatively good by developing country standards. The use of annual data covering the period 1965 to 2003 is sufficiently long to allow for a meaningful time series investigation. These therefore address the concerns raised about the lack of time series-based individual country study. The study uses the recently developed bounds testing procedure to analyse the level relationship between saving and investment within an autoregressive distributed lag (ARDL) framework. The results indicate a robust cointegrating relationship between domestic saving and investment rates.

The rest of this article is organized as follows: Section II provides an overview of the Malaysian experience. Model and econometric methodology are set out in Section III. In Section IV, we present the empirical results. Finally, we briefly discuss the policy implications of the results and conclude in the last section.

II. A Brief Overview of the Malaysian Experience

This section examines the behaviour of Malaysia’s saving and investment rates over the period 1960 to 2003. Malaysia can be characterized as a very open- and high-growth economy. The average annual rate of growth in real gross domestic product (GDP) for Malaysia during the period 1961 to 1980 was 7.2% and for 1981 to 1996, the average annual growth rate was 7.4%. The success of Malaysia’s sustaining high economic growth is mainly attributed to a rapid increase in domestic capital accumulation. However, rapid capital formation did not exert excess pressure on balance of payments due to the impressive saving records of the country (Athukorala, 2001). From 1997 to 2003, the average annual growth rate was much lower at 3.5% due to the impact of the Asian financial crisis.

Figure 1 presents the time series plots of gross domestic saving rate, gross domestic investment rate and M2 as a ratio of GDP. The data were obtained from World Development Indicators (2005). As shown in Fig. 1, it is evident that Malaysia has been very successful in mobilizing saving. The share of gross domestic saving to GDP rose from a low of 25.9% in 1960 to a high of 42.9% in 1996, before the onset of the Asian financial crisis. Gross domestic investment was primarily funded by domestic saving (which includes household saving, corporate saving and government saving) and supplemented by foreign saving. It increased from 13.8% of GDP to 41.5% of GDP during the same period. However, gross domestic investment was severely affected in the aftermath of the global
economic recession and the recent Asian financial crisis.

In order to channel saving to productive activities, it is necessary to mobilize saving in which an efficient financial system is indispensable. Accompanying the rise in saving rates has been the rapid pace of financial deepening. The share of M2 to GDP, a commonly used measure of the level of financial development, rose from 22.3% in 1960 to 83.3% in 1996.

Development of the financial system may be shaped by the financial sector policies. Malaysia followed a gradual approach in its financial sector reforms, which started in the 1970s by cautiously liberalizing the interest rates. The major phase of interest rate liberalization occurred in 1978 when the commercial banks were allowed to set the deposit and lending rates freely. The market-determined interest rate mechanism was interrupted from 1985 to 1987 to mitigate the world economic recession impacts on Malaysia. In 1987, the central bank turned to the use of the base-lending rate to control interest rates. However, these interest rate controls were removed in 1991 (Yusof et al., 1994).

Although the financial reform programmes appeared to be narrow in scope with the key objective of eliminating interest controls (Bascom, 1994), the liberalization policies adopted by the Malaysian government seem to have worked well at various stages of development in which financial deepening has clearly been observed. This is particularly evident when the financial sector underwent a radical transformation and deepening in the 1980s following the liberalization of interest rates in 1978. The upshot of this transformation is the emergence of a broader, deeper, more organized and better-structured financial system and an increased mobilization of savings. These financial reform measures and the development that follows, can have significant impacts on the saving-investment relationship. Consequently, the relationship of these two variables must be examined due to the changes in the financial environments.\(^1\)

### III. Model and Methodology

To examine the saving-investment nexus for Malaysia, we adopt the generic long-run model of Feldstein and Horioka (1980) which has the following form:

\[
I_t = \alpha + \beta S_t + \epsilon_t
\]

Following the definition used by Feldstein and Horioka (1980), \(I_t\) refers to the ratio of gross capital formation to GDP at time \(t\) and \(S_t\) is the ratio of gross domestic saving to GDP at time \(t\). \(\alpha\) is a constant, \(\beta\) is the regression coefficient between \(I_t\) and \(S_t\) and \(\epsilon_t\) is the error term. A high estimate of \(\beta\) would suggest that most of the saving remains in the domestic economy and therefore provides evidence against the hypothesis of perfect capital mobility.

Such a simple econometric specification is, however, subject to several limitations. If \(I_t\) and \(S_t\) contain unit roots, then regressing Equation 1 in levels may produce ‘spurious’ regression results (Granger and Newbold, 1974). The specification also ignores short-run adjustment dynamics which are required to obtain the long-run relationship.

To account for these limitations, we begin our analysis by maintaining the assumption that the data generating process for the relationship between \(I_t\) and \(S_t\) is a log-linear vector autoregressive (VAR) model at levels, which can be augmented with deterministics such as intercepts and time trends. The VAR model is given as:

\[
x_t = \mu + \sum_{j=1}^{p} \phi_j x_{t-j} + \epsilon_t
\]

where \(x_t = [I_t, S_t]'\). The series is converted into natural logarithms for the usual statistical reasons. They can also be interpreted in growth terms after taking first difference. \(\mu\) is a vector of constant

---

\(^1\) For a more detailed exposition of the relationship between financial development and economic growth in Malaysia, see Ang and McKibbin (2007).
terms where \( \mu = [\mu, \mu^j] \) and \( \phi_j \) is a matrix of VAR parameters for lag \( j \). The vector of error terms \( \varepsilon_t = [\varepsilon_t, \varepsilon_{S,t}] \sim \text{IN}(0, \Omega) \) where \( \Omega \) is positive definite and is given by:

\[
\Omega = \begin{bmatrix}
\omega_{11} & \omega_{1S} \\
\omega_{S1} & \omega_{SS}
\end{bmatrix}
\]

Given this, \( \varepsilon_{t,t} \) can be expressed in terms of \( \varepsilon_{S,t} \) as:

\[
\varepsilon_{t,t} = \omega_{SS}^{-1} \varepsilon_{S,t} + \varepsilon_{t,t}
\]

where \( \varepsilon_{t,t} \sim \text{IN}(0, \omega_{11}) \).

The VAR model in Equation 2 can be transformed into a vector error correction model (VECM) after some mathematical manipulation. Hence, we have

\[
\Delta x_t = \mu + x_{t-1} + \lambda \sum_{j=1}^{p-1} \gamma_j \Delta x_{t-j} + \varepsilon_t
\]

where \( \Delta = 1-L \), \( \lambda \) is the long-run multiplier matrix and

\[
\gamma_j = \begin{bmatrix}
\gamma_{1j} & \gamma_{S1j} \\
\gamma_{S1j} & \gamma_{SS,j}
\end{bmatrix} = -\sum_{k=j+1}^{p} \Phi_k
\]

We include a dummy variable to account for structural break in the investment series that occurs due to the impacts of the 1997–1998 Asian financial crisis. The dummy variable is defined by:

\[
D = \begin{cases}
1 & \text{if } t = 1998-2003 \\
0 & \text{otherwise}
\end{cases}
\]

Since the fraction of the nonzero dummy variables is only 15.4% over the entire sample period, 1965 to 2003, the inclusion of such dummy variable does not affect the asymptotic properties of the ARDL test. Similarly, an interactive dummy variable with the lagged changes of \( S_t \) is also included in the specification. The conditional error correction model (ECM) of interest can be written as:

\[
\Delta I_t = \alpha_0 + \alpha_1 D + \alpha_2 I_{t-1} + \alpha_3 I_{S,t-1} + \sum_{j=1}^{p-1} \beta_j \Delta I_{t-j} + \sum_{i=0}^{p-1} \delta_i (D \Delta S_{t-i}) + \varepsilon_t
\]

The earlier-mentioned econometric specification hinges upon the assumption that the disturbance term \( \varepsilon_t \) is serially uncorrelated. Hence, it is critical to ensure that optimal lag order \( p \) of the underlying VAR is chosen appropriately in order to strike a satisfactory balance between the two competing concerns, in which the lag order is high enough to ameliorate the residual serial correlation problems while low enough so that the conditional ECM is not subject to over-parameterization problems. This is particularly important, given that a small sample is used in the estimation of this study. Augmenting the specification with an adequate number of lagged changes in the regressors also corrects the problems of omitted variables bias (Inder, 1993; Hendry, 1995). The determination of an appropriate and correctly specified ARDL model is based on test criteria such as the Akaike's information criterion (AIC), Schwarz's Bayesian criterion (SBC) and other diagnostic tests for econometric problems.

Next, we have to decide whether a time trend should be incorporated into the conditional ECM. In Fig. 1, it is clear that there is an upward trend associated with \( I_t \) and \( S_t \). Hence, there is no compelling reason to ignore a time trend in our empirical analysis during the sample period under consideration. Annual time series data covering the period 1960 to 2003 are available for estimation. However, to ensure comparability in model selection, we use the same sample period, 1965 to 2003 in all estimations, with the first five observations reserved for the construction of lagged and differenced variables. Hence, we are left with 39 observations in all estimations.

The ARDL bounds approach developed by Pesaran et al. (2001) can be used to establish the short-run and long-run relationships between saving and investment rates. The advantage of this approach is that it can be applied to the model irrespective of whether they are purely \( I(0) \), purely \( I(1) \), or mutually cointegrated. But to apply the bounds procedure, it is important to ensure that the variables under consideration are not integrated at an order higher than one. In the presence of \( I(2) \) variables, the critica value provided by Pesaran et al. (2001) are no longer valid.

To assess the degree of integration of the variables under investigation, we perform three unit root tests – augmented Dickey–Fuller (ADF) test, Phillips–Perron (PP) test and Kwiatkowski–Phillips–Schmidt–Shin (KPSS) test. The choice of the KPSS test to complement the widely employed ADF and PP tests is motivated by the argument that tests designed on the basis of the null that a series is \( I(1) \) have low power of rejecting the null. Since most economic time series are not very informative about whether or not there is a unit root, it would be useful to perform tests of the null hypothesis of stationarity as well as tests of the null hypothesis of a unit root (Kwiatkowski et al., 1992).

Two separate statistics are employed to ‘bounds test’ for the existence of a long-run relationship: 1) an \( F \)-test for the joint significance of the coefficients on the lagged-level terms of the unrestricted error correction model \( (H_0: \alpha_3 = \alpha_4 = 0) \) and 2) a \( t \)-test for the significance of the coefficient associated with \( I_{t-1} \).
(\(H_0: \alpha_3 = 0\)). Two asymptotic critical value bounds provide a test for cointegration when the independent variables are \(I(d)\) (where \(0 \leq d \leq 1\)). The lower bound assumes that all the independent variables are \(I(0)\) and the upper bound assumes that they are \(I(1)\). If the test statistics exceed their respective upper critical values, the null is rejected and we can conclude that a long-run relationship exists. If the test statistics fall below the lower critical values, we cannot reject the null hypothesis of no cointegration. If the statistics fall within the band, the statistical inference would be inconclusive. The critical values are provided by Pesaran et al. (2001).

IV. Results

The ADF and PP test the null of a unit root against the alternative of stationarity, while the KPSS tests the null of stationarity against the alternative of a unit root. We allow both intercept and intercept with trends in the testing. As shown in Table 1, the testing results are mixed. Hence, the widely used Johansen cointegration may not be appropriate in this case. Since none of the variables are integrated at an order higher than one, this allows the use of the ARDL bounds procedure.

To determine the optimal lag length \(p\) and whether a deterministic linear trend is required, we estimated the conditional ECM in Equation 4 by ordinary least square (OLS), with and without a linear time trend. Given the sample size, we have considered a maximum lag length of four and then tested down. Table 2 gives Akaike’s and Schwarz’s Bayesian information criteria, denoted by AIC and SBC, respectively.

In Table 2, the lag order selected by AIC (\(p_{\text{AIC}} = 3\)) is larger than that selected by SBC (\(p_{\text{SBC}} = 1\)) when a deterministic time trend is included in the model. The results are not surprising, given that SBC tends to select a smaller lag length. In the model without such a deterministic trend term, both AIC and SBC unanimously point to one lag. In the light of the importance of the assumption of no serial correction in the residuals, it seems prudent to choose the lag length to be three or higher. Since the results may be sensitive to the lag length chosen, we report test results for both \(p = 3\) and \(p = 4\).

In Table 3, the null hypothesis that there exists no level investment equation is rejected at the 5% significance level, irrespective of whether the regressors are purely \(I(0)\), purely \(I(1)\), or mutually cointegrated. The results are robust to the number lags included and the inclusion of a deterministic trend in the specification. Overall, the results strongly support the existence of a long-run relationship between investment and saving. No cointegrating relationship is found if saving rate was used as the dependent variable.

Table 1. Test results for unit roots

<table>
<thead>
<tr>
<th></th>
<th>(S_t)</th>
<th>(\Delta S_t)</th>
<th>(I_t)</th>
<th>(\Delta I_t)</th>
</tr>
</thead>
<tbody>
<tr>
<td>ADF(^a) Intercept</td>
<td>-1.053 (2)</td>
<td>-8.275 (1)**</td>
<td>-1.799 (0)</td>
<td>-5.519 (0)**</td>
</tr>
<tr>
<td>ADF(^a) Intercept and trend</td>
<td>-2.909 (3)</td>
<td>-8.202 (1)**</td>
<td>-1.359 (0)</td>
<td>-6.562 (0)**</td>
</tr>
<tr>
<td>PP(^b) Intercept</td>
<td>-1.608 (1)</td>
<td>-11.217 (10)**</td>
<td>-1.907 (2)**</td>
<td>-5.519 (0)**</td>
</tr>
<tr>
<td>PP(^b) Intercept and trend</td>
<td>-4.300 (0)**</td>
<td>-11.114 (10)**</td>
<td>-1.599 (2)**</td>
<td>-5.619 (1)**</td>
</tr>
<tr>
<td>KPSS(^c) Intercept</td>
<td>0.715 (5)**</td>
<td>0.151 (8)</td>
<td>0.468 (4)**</td>
<td>0.174 (1)</td>
</tr>
<tr>
<td>KPSS(^c) Intercept and trend</td>
<td>0.062 (1)</td>
<td>0.109 (8)</td>
<td>0.117 (4)</td>
<td>0.067 (0)</td>
</tr>
</tbody>
</table>

Notes: \(^a\)\(H_0\) = the series has a unit root. AIC is used to select the lag length. The maximum number of lags is set to be four. \(^b\)\(H_0\) = the series has a unit root. Barlett–Kernel is used as the spectral estimation method. The bandwidth is selected using Newey–West method. \(^c\)\(H_0\) = the series is stationary. Barlett–Kernel is used as the spectral estimation method. The bandwidth is selected using Newey–West method.

*, ** and *** indicate 10, 5 and 1% levels of significance, respectively. The optimal lag length or bandwidth is indicated in the parentheses.
In the hypothesis testing for the presence of a long-run level equation for investment, it is important to ensure that no restrictions on the coefficients of the lagged terms are imposed. Otherwise, the test results are subject to a pre-testing problem. However, to estimate the level effects and short-run dynamics of investment adjustments, the use of a parsimonious specification is more desirable. To this end, we adopt Hendry’s (1995) general-to-specific modelling approach to derive a satisfactory model for the equation $\Delta I_t$. Given the sample size, we begin by formulating an ARDL with four lags. We use a systematic algorithm to search for the ‘best’ model. First, the lag orders of the ARDL with two variables ($I_t, S_t$) were selected by searching across the $4^3 = 64$ ARDL, spanned by $p_i = 1$ to 4, $i = 0, 1, 2$, using AIC criterion. Such a general model is then tested down by dropping statistically insignificant terms. Hence, the exact formulation of the model depends on the time series properties of the data. The final parsimonious model must satisfy various diagnostic tests. This resulted in the choice of an ARDL specification reported in Table 4.

The regression results for the conditional ECM of $\Delta I_t$ show several desirable statistical features. All the coefficients are statistically significant at the 10% level. The regression specification fits remarkably well and passes the diagnostic tests against nonnormal residuals, serial correlation, heteroskedasticity, autoregressive conditional heteroskedasticity and functional misspecification.

The structural stability of the conditional ECM is examined using the cumulative sum and cumulative sum of squares tests on the recursive residuals. The cumulative sum test is able to detect systematic changes in the regression coefficients whereas the cumulative sum of squares test is able to detect sudden changes from the constancy of the regression coefficients (Brown et al., 1975). Figure 2 displays the cumulative sum and cumulative sum of squares results. In the upper panel, although the series appears to be trending upwards after the crisis period, the cumulative sum statistics lie within the 5% confidence interval bands. Similarly, the cumulative sum of squares statistics displayed in the lower panel of Fig. 2 are within the 5% critical bands, suggesting no structural instability in the residuals of the equation for $\Delta I_t$.

Using the results reported in Table 4, the estimated long-run equilibrium relationship in levels is given by:

$$-0.730 + 0.010 \text{ trend} - 0.650 I_t + 0.301 S_t = 0 \quad (5)$$

The coefficient associated with $I_t$, which measures the speed of adjustment back to the long-run equilibrium value, is statistically significant at the 1% level and correctly signed (negative). This implies that the long-run equilibrium deviation has a significant impact on the growth of investment rate. It takes only about 1.5 years to achieve long-run equilibrium whenever there is a deviation from equilibrium. By normalizing the coefficient of $I_t$ to one, the long-run equilibrium relationship between $I_t$ and $S_t$ can be described as:

$$I_t = -1.123 + 0.015 \text{ trend} + 0.463 S_t \quad (6)$$
Hence, the elasticity of $I_t/C_0$ with respect to $S_t/C_0$ is found to be 0.463 in the long-run, suggesting a positive long-run relationship between investment and saving rates. In contrast to the findings of Feldstein and Horioka (1980), the null hypothesis of $\alpha_4 = 1$ (the regression coefficient between investment and saving rates) is firmly rejected in this case.

### V. Conclusions

In this article, we examine the cointegrating relationship between domestic saving and investment in an ARDL framework. Using the Malaysian data for the period 1965 to 2003, we find a fairly robust long-run relationship between domestic saving and investment, after controlling the effects of the Asian financial crisis on domestic investment rate.

The findings of a robust long-run cointegrated relationship between domestic saving and investment suggest that any change in domestic saving will be closely associated with a change in investment. Hence, financial sector policies targeting at mobilizing domestic saving are critical for capital accumulation. However, it should also be highlighted that over reliance on domestic saving may limit the growth opportunity of an economy. As such, policy makers should also focus on attracting foreign capital as part of the development policy while mobilizing resources in the domestic economy.
References


