COUNTRY SIZE AND INVESTMENT-SAVING CORRELATION: A PANEL THRESHOLD ERROR CORRECTION MODEL

Teung-wu Ho
Shih Hsin University

and

Ru-Lin Chiu
Shih Hsin University

INTRODUCTION

The investment-saving regression parameter is known as the saving-retention coefficient. Whether the saving-retention coefficient explains the degree of international capital mobility has sparked a large body of debate in the literature. The objective of this paper is to examine the country-size approach to the interpretation of the saving-retention coefficient.

Feldstein and Horioka [1980] first documented the idea that perfect capital mobility can be featured by the investment-saving regression coefficient (or saving-retention coefficient) in each country, which is statistically insignificant from zero. Similarly, in a world of perfect capital immobility, this regression coefficient will be insignificantly different from one (or saving-retention coefficient equals 1). Following Feldstein and Horioka [1980], the investment-saving correlation is:

\[ (I/GDP)_t = \alpha + \beta (S/GDP)_t + \epsilon_t \]

where \( I \) is the real investment, \( S \) is the real saving, \( GDP \) is the real gross domestic product, and \( \epsilon_t \) is i.i.d. \( N(0, \sigma^2) \). This regression coefficient is interpreted as a measure of international capital mobility. Feldstein and Horioka [1980] used data for sixteen industrial OECD countries to test the hypothesis that there is low correlation of domestic saving and investment rates. It appeared that a sustained increase in domestic saving has a roughly proportional long-term effect on domestic investment; the regression coefficient is interpreted as the proportion of the incremental saving that is invested domestically. They concluded that the high saving-investment correlation across these countries is inconsistent with the hypothesis of world capital market integration. Their finding was confirmed by Feldstein [1983] for a longer time period.
Subsequent studies by Penati and Dooley [1984], Vos [1989], Dooley et al. [1997], Feldstein and Racerbett [1989], and Tsear [1990] verified the close association between domestic saving and investment rates both for industrial and developing countries. Empirical findings also indicate that domestic saving and domestic investment are highly correlated across OECD countries and over time, thus overwhelmingly rejecting the implication of perfect capital mobility.

These findings have been interpreted as indicating that the world is characterized by capital immobility, yet most economists believe that the world is characterized by an increasingly high degree of international capital mobility. In particular, Sachs [1981] presents empirical evidence that current account deficits are associated with investment booms, implying that increases in domestic investment are at least partly financed by capital inflows. Taken together, this is the Feldstein-Horioka puzzle, which has sparked a substantial disagreement about the interpretation of the saving-retention coefficient. Cooley and Smith [1990] argue that the saving-retention coefficient is irrelevant to the analysis of the hypothesis of international capital mobility, but is instead useful as a proxy for current account solvency. This is because, in the presence of a current account solvency constraint, the current account balance should be stationary and thus investment and saving rates are cointegrated with a unitary coefficient. Sachs and Caetano [2000] argue that the saving-retention coefficient does not even reflect capital mobility, but merely the variability between external and domestic saving.

Murphy [1984] questioned the robustness and generality of Feldstein-Horioka model, relaxed the small country assumption implicitly assumed in this model, and argued that the estimated regression coefficient is sensitive to country size as measured by GNP. Recently, Baxter and Crucini [1993] proposed a general equilibrium model to explain that high correlation can be an evidence of high capital mobility if country size, as measured by GNP, is considered because larger countries have a greater impact on the world interest rate. Baxter and Crucini also emphasize that their theory cannot be applied to small country, nor can it be inferred that small countries have zero correlation.

Baxter and Crucini [1993] estimate the individual saving retention coefficients and compare their magnitude by the country-size ranking. They find that larger countries have larger saving retention coefficients.

However, Baxter and Crucini's approach is defective, since country size is measured by the country-specific GNP, it is implausible for a country's relative output to be the same during the entire sampling period. That is, it is inappropriate to use country as a unit to characterize the country-size effect on the saving-retention coefficient. A plausible formulation is to assume that a given country's size changes over the sampling period. Therefore, the objective of this paper is to empirically examine the threshold effects of country size on the saving-retention coefficients, under which the saving-retention coefficient is subject to change by relative output share. We apply a threshold regression model, developed by Hansen [1999]. Twenty-four OECD countries form our panel, and the time series data are derived from AREMOS/OECD, ANIA (Annual National Income Accounts) database; the sample period extends from 1961 to 1997 for each country. The next section analyzes the time series properties of each country followed by a presentation of the panel threshold estimation results.

**ECONOMETRIC METHODOLOGY**

The **Time-Series Properties**

First of all, we test for the presence of a panel unit root. The standard approach to test for nonstationarity of each observed time series \( y_t \) over \( T \) time periods in a panel of \( N \) countries is to estimate an augmented Dickey-Fuller (ADF) regression, including a time trend:

\[
\Delta y_{it} = \alpha_i + \delta_t + \beta y_{i,t-1} + \sum_{j=1}^{k-1} \rho_j \Delta y_{it-j-1} + \varepsilon_{it}, \quad i = 1, \ldots, N; \quad t = 1, \ldots, T,
\]

where \( \rho_j \) is selected to make the residuals of country \( i \) uncorrelated over time. The null hypothesis for the presence of unit root is \( H_0: \beta = 0 \). The panel-based ADF, extended by Im et al. [1999], allows complete heterogeneity between individuals. In the panel setup, the ADF equation is estimated separately for each individual and thus allows for differing parameter values, residual variance and even different lag lengths. Hence the ADF in panel data is based on the average of the \( N \) country-specific ADF t-statistics as

\[
\hat{t}_N = (1/N) \sum_{i=1}^{N} t_i(p_i)
\]

where \( t_i(p_i) \) is the ADF t-statistic for country \( i \) based on the inclusion of \( p_i \) lags in the country-specific ADF regression. To avoid the small-sample bias, we calculated the critical values of the panel-based test by using Monte-Carlo simulations calibrated for our sample size. These critical values are also reported in Table 1. Evidently, the presence of a panel unit root is confirmed in two series. This result is consistent with current panel unit root test results regarding these two ratios (Oh et al., 1999). To account for non-stationarity, the simplest representation, commonly adopted in the literature, is the first-order error correction model. The **Fixed-Effect Interaction Model**

Equation (1) can be easily modified to an error correction model in the panel data below. To introduce the country-size effect, we use interaction terms
\[
\Delta \text{GDP}_t = \alpha_0 + \beta_1 \Delta \text{GDP}_{t-1} + \beta_2 \Delta \text{GDP}_{t-2} + \ldots + \beta_k \Delta \text{GDP}_{t-k} + \epsilon_t
\]

where \( \epsilon_t \) is the index of country size measured by (GNP/GNP)_c, \( \Delta \text{GDP}_t \) is the individual country index. Appendix A summarizes the selected years of the GNP ratio of each country. The denominator is the sum of GNP of 24 OECD member countries of a given year. This measure does not precisely calculate the GNP share of individual countries. However, this cannot be a serious problem given that this index measures relative country size across the panel and changing the aggregation of the denominator does not change its relative position in respect to the panel.

It is immediately obvious that Equation (4) embodies a long-run equilibrium relationship between saving and investment prescribed by theory, where, in the steady-state equilibrium, the implied relationship in levels is derived easily. Parameter restrictions may be used to test various natural hypotheses concerning the nature of this equilibrium. The size of \( \epsilon_t \) is a measure of the sustainability of current-account imbalances, which has a natural interpretation: it measures the speed of convergence of the system toward equilibrium. If \( \epsilon_t = 1 \) then the long-run equilibrium current account balance is equal to a constant, \( \Delta \text{GDP}_t = (\text{GDP}/\text{GDP})_t \), where the asterisk denotes the steady-state equilibrium. Further, if \( \epsilon_t \neq 0 \) then this constant is zero. Current account measures changes in net national indebtedness; hence, for most open economies, the importance of inter-temporal plans lies in the time path of the current account. Temporary current account deficits may not be "bad", because they imply capital flows to the countries where capital is more productive. However, persistent deficits can have serious effects: first, they can raise domestic interest rates to attract foreign capital. Second, the subsequent accumulation of external debt will imply increasing interest payments, which may impose an excess burden on future generations. Policymakers are thus concerned about the aggravating effect of persistent current account deficits.

Instead of emphasizing the size of current account deficits at any particular point in time, economists are more concerned with the country's inter-temporal solvency constraint. This constraint emphasizes the long-run path of the current account [Husted, 1992; Ghosh, 1995; and Wu et al., 1996]. A stationery current account implies a finite external debt-to-GNP ratio, and hence the sustainability of external debts. In this case, therefore, there is no reason for the country to default on its international debts. Moreover, the stationarity of the current account is also important to the validity of the modern inter-temporal model of the current account. Theoretically, the modern inter-temporal model of current account determination combines the assumptions of perfect capital mobility and consumption-smoothing behavior to explain that the current account acts as a buffer to smooth consumption in response to shocks.

The saving-retention coefficient \( \beta \) has a natural interpretation: given a long-run equilibrium relationship between saving and investment, these coefficients measure to what extent a temporary annual shock to domestic saving will pass through into domestic investment. Moreover, following the test outlines by Kremer et al. [1992] and Jansen and Schulze [1996], we may interpret a hypothesis test that \( \epsilon_t \) is non-zero as equivalent to a test of cointegration. Converted into a "half-life" measure, this offers an indication of the sustainability of current-account disequilibria for our panel in the sample period under study.

To perform PFEs, seemingly unrelated regression (SUR) weighting is applied. The SUR-weighted least squares, sometimes referred to as the Parks estimator, is the FGLS estimator when the residuals are both cross-section heteroskedastic and contemporaneously correlated. Let \( \Sigma \) denote the symmetric matrix of contemporaneous correlations, assuming typical element \( \sigma_{ij} = \text{R}(\xi_i | \xi_j) \) is constant, the covariance matrix is:

\[
\Sigma = \Sigma \otimes I_P = \begin{bmatrix}
\sigma_{11} & \cdots & \sigma_{1P} \\
\sigma_{21} & \cdots & \sigma_{2P} \\
\vdots & \cdots & \vdots \\
\sigma_{P1} & \cdots & \sigma_{PP}
\end{bmatrix}
\]

where \( \sigma_{ij} \) is estimated from a first-stage pooled OLS regression. Moreover, White's heteroskedasticity consistent covariance estimates are employed also.

The Threshold Modeling

The threshold model assumes the data is given by \( (y, x, D) \), where \( D \) is a threshold variable, respectively, and \( x \) is a \( p \times 1 \) vector of independent variables. The threshold variable \( D \) splits the sample into different groups and may be part of \( x \). According to Hansen [1999], the single threshold model in panel data takes the form below:

\[
y_{it} = x_{it}' \beta_{i} + \epsilon_{it}, D_{it} = \gamma
\]

(6)

\[
y_{it} = x_{it}' \beta_{i} + \epsilon_{it}, D_{it} > \gamma
\]

(7)

The observations in equations (6) and (7) are divided into two regimes depending on whether the threshold variable is smaller or larger than the threshold value \( \gamma \). The regimes are distinguished by varying regression slopes, \( \beta_{i} \) and \( \beta_{i}^* \). For the identification of \( \beta_{i} \) and \( \beta_{i}^* \), \( x_{it} \) and \( D_{it} \) must be time-varying. The error term is assumed to be independent and identically distributed with mean zero and finite variance \( \sigma^2 \).

The model can be written in a single equation form by defining an indicator function \( I(\cdot) \), which takes the value of 1 if the threshold condition is satisfied and 0 otherwise. So that equations (6) and (7) can be expressed as:

\[
y_{it} = x_{it}' \beta_{i} + \beta_{i}^* I(D_{it} = \gamma) + \epsilon_{it}, D_{it} = \gamma
\]

(8)

For any given threshold value, the slope coefficient can be estimated by ordinary least squares. Chan [1993] and Hansen [1996] recommend estimation of \( \gamma \) by least squares. The easiest way to achieve this is by minimizing the concentrated sum of
TABLE 2
Fixed-Effect Panel Model

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>S/GDP in 1960</td>
<td>α</td>
<td>0.800 (0.33)</td>
<td>0.850 (0.33)</td>
<td>0.820 (0.38)</td>
<td>0.730 (0.670)</td>
<td></td>
</tr>
<tr>
<td>S/GDP in 1960</td>
<td>β</td>
<td>0.000 (0.04)</td>
<td>0.100 (0.03)</td>
<td>0.080 (0.02)</td>
<td>0.320 (0.230)</td>
<td></td>
</tr>
<tr>
<td>S/GDP in 1960</td>
<td>γ</td>
<td>-0.550 (0.04)</td>
<td>-0.500 (0.04)</td>
<td>-0.200 (0.04)</td>
<td>-0.100 (0.025)</td>
<td></td>
</tr>
<tr>
<td>S/GDP in 1960</td>
<td>δ</td>
<td>-0.150 (0.02)</td>
<td>-0.150 (0.015)</td>
<td>-0.050 (0.05)</td>
<td>0.000 (0.015)</td>
<td></td>
</tr>
<tr>
<td>Adjusted R²</td>
<td></td>
<td>0.68</td>
<td>0.67</td>
<td>0.52</td>
<td>0.55</td>
<td></td>
</tr>
<tr>
<td>R²-statistic</td>
<td></td>
<td>184.1</td>
<td>182.1</td>
<td>160.2</td>
<td>339.1</td>
<td></td>
</tr>
</tbody>
</table>

Numbers in the parentheses are standard errors.

SCHERER'S ECONOMIC JOURNAL

characterizes the period of high capital controls. The saving-retention coefficient is 0.80; it implies that the one-percentage increase in domestic saving ratio passes through 80 percent domestic investment rate. The interaction between country size and the saving-retention coefficient insignificantly indicates that for every one-percentage increase in relative output share, the saving-retention coefficient increases 0.06 percent. However, since it is not significant, we conclude that there is no country-size effect in this period. The error correction term (ε) is significantly different from zero, which verifies the presence of cointegration. Moreover, the t-statistic for δ = 0 is rejected significantly, implying that the long-run equilibrium current account is not a constant.

Second, we find that the country-size effect is extremely sensitive to sample size. When we add one more year into our samples, we find that β becomes significant. It significantly indicates that, given the saving-retention coefficient of 0.81, for every one-percentage increase in relative output share, the saving-retention coefficient increases 0.14 percent. The t-statistic for δ = 0 is still rejected at 5 percent significance level, implying that the long-run equilibrium current account is not a constant.

Third, when we sample the post-Bretton-Wood period, 1980-1997, characterizing the period that capital controls are gradually released, we find more interesting results. The saving-retention coefficient dramatically drops to 0.58, and the country-size effect substantially increases to 0.88. This drastic change implies that the country-size effect has time-varying effects on the saving-retention coefficient; in particular, its effect becomes stronger when the capital control is lifted. Moreover, the t-statistic for δ = 0 is accepted at the 5 percent significance level, implying that the long-run equilibrium current account becomes a constant.

Finally, the full-sample estimation indicates that the saving-retention coefficient (β) is 0.73, which is statistically significant; β significantly indicates that for every one-percentage increase in relative output share, the saving-retention coefficient increases 0.61 percent. This finding substantially confirms the effect of country size on the saving-retention coefficient. δ, the ECM term, is significantly different from zero, which verifies the presence of cointegration in our model; moreover, the t-statistic for δ = 0 is accepted significantly, implying that the long-run equilibrium current account is a constant.

Therefore, according to Table 2, we have three findings that the country size has a complex and time-varying effect on the magnitude of saving-retention coefficient, indicating the existence of possible non-linearity. Moreover, the current account equilibrium is also sensitive to the period, but it exhibits a consistent adjustment to long-run equilibrium. In a nutshell, the relationship between the saving-retention coefficient and country size is very complex, which motivates the subsequent threshold analysis.

Table 3 reports the test statistics for the number of thresholds and their values. It clearly indicates the presence of a double threshold, and accepts the null of no triple threshold. The threshold estimates are 0.00049 and 0.0045. Hence, there are three saving-retention coefficients with respect to three country-size regimes: less than 0.0049 percent, between 0.049 percent and 0.45 percent, and greater than 0.45 percent.

EMPIRICAL RESULTS

Table 2 presents the estimation results of a conventional fixed-effect model of equation (4). Because the 1960s was a period where capital controls were in place and the United States dominated in the international capital markets, we divide our samples into several periods and estimate the parameters separately. We have several important results: First, we estimate the period ranging from 1960 to 1971, which
COUNTRY SIZE AND THE INVESTMENT-SAVING CORRELATION

Table 3 presents the parameter estimates. Three points are worth mentioning. Firstly, $\beta_1$ is significantly different from zero, which indicates the presence of cointegration in our model. Second, the t-statistic for $\beta_2 = 0$ is accepted significantly, implying that the long-run equilibrium current account is a constant. Finally, the short-run saving-retention coefficients of small countries ($\beta_1$), the medium countries ($\beta_2$), and large countries ($\beta_3$) are, respectively, -0.043, 0.0052, and 0.014. These parameter estimates increase with country size. Moreover, small countries, whose output shares are less than 0.049 percent, have insignificant parameter estimates. The medium-sized countries have relatively higher saving-retention coefficient. The big countries have the highest coefficients. The magnitudes of these estimates are consistent with relevant research. For example, the parameter estimates of Coïtoux and Olivier (2005) and Jansen (2006), similar to those reported in Table 5 of their article.

In general, the estimation results support the Baxter-Crouzet country-size hypothesis. However, the most important finding concerning the saving-retention coefficient is that the largest saving-retention coefficient is around 0.6, which is 10 percent lower than the estimate of Table 2. The discrepancy can be interpreted by the non-linear effect of country size on the saving-retention coefficient.

SUMMARY AND CONCLUSION

Economic theory suggests that, in a world of perfect capital mobility, capital flows among countries equalize the yield to investors. In the presence of such arbitrage, it is possible for domestic saving and domestic investment to be uncorrelated to each other. Therefore, saving in each country reacts to cross-country differences in rates of return on capital while the level of investment is financed from the world capital market through a current account deficit. Such a theory has typically been interpreted as providing evidence that capital is highly internationally mobile, while high time-series correlation of investment and saving have typically been interpreted as evidence that capital is not highly mobile internationally. We extend the conventional Feldstein-Horioka model to allow data to be drawn from three country-size regimes (or two thresholds), known as country size intervals. The hypothesis that the saving-retention coefficients vary with country size over time could not be rejected, which justifies the use of a double threshold model to examine the country-size argument of the saving-investment relationship.

This paper provides substantial evidence to reconcile this empirical puzzle and confirms the country-size hypothesis. In other words, the domestic capital markets of twenty-four OECD countries tend to evolve internationally. The reason for these findings cannot be interpreted as a higher interpretative power of the panel threshold model than that for the conventional model.

**APPENDIX**

**TABLE 3**

<table>
<thead>
<tr>
<th>Tests for the Number of Thresholds and Threshold Estimates</th>
</tr>
</thead>
<tbody>
<tr>
<td>(1) $H_0$: There is no threshold. (Tests for single threshold)</td>
</tr>
<tr>
<td>(2) $H_0$: There is only one threshold. (Tests for double threshold)</td>
</tr>
<tr>
<td>(3) $H_0$: There are only two thresholds. (Tests for triple threshold)</td>
</tr>
</tbody>
</table>

Threshold estimates $\gamma_1=0.0049$ $\gamma_2=0.0045$

Numbers in the parenthesis are critical values at 10%, 5%, and 1% present significance levels.

---

**TABLE 4**

<table>
<thead>
<tr>
<th>Regression Estimates of Double Thresholds</th>
</tr>
</thead>
<tbody>
<tr>
<td>Regressors</td>
</tr>
<tr>
<td>-------------</td>
</tr>
<tr>
<td>$\Delta(GDP_t^H)/D_t=0.00049$</td>
</tr>
<tr>
<td>$\Delta(GDP_t^P)/D_t=0.00049$</td>
</tr>
<tr>
<td>$\Delta(GDP_t^H)/D_t=0.00049$</td>
</tr>
<tr>
<td>$\Delta(GDP_t^P)/D_{t-1}$</td>
</tr>
<tr>
<td>$\Delta(GDP_t^H)/D_{t-1}$</td>
</tr>
</tbody>
</table>

Standard error is the White heteroscedastic standard error.

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Australia</td>
<td>0.01716</td>
<td>0.01838</td>
<td>0.01724</td>
<td>0.01772</td>
<td>0.01591</td>
</tr>
<tr>
<td>Austria</td>
<td>0.00682</td>
<td>0.00939</td>
<td>0.00735</td>
<td>0.00956</td>
<td>0.01042</td>
</tr>
<tr>
<td>Belgium</td>
<td>0.01163</td>
<td>0.01531</td>
<td>0.00800</td>
<td>0.01176</td>
<td>0.01231</td>
</tr>
<tr>
<td>Canada</td>
<td>0.09726</td>
<td>0.03350</td>
<td>0.03992</td>
<td>0.03434</td>
<td>0.02980</td>
</tr>
<tr>
<td>Denmark</td>
<td>0.0744</td>
<td>0.06854</td>
<td>0.09559</td>
<td>0.06800</td>
<td>0.00814</td>
</tr>
<tr>
<td>Finland</td>
<td>0.05021</td>
<td>0.05046</td>
<td>0.05025</td>
<td>0.00928</td>
<td>0.00567</td>
</tr>
<tr>
<td>France</td>
<td>0.06578</td>
<td>0.08369</td>
<td>0.05813</td>
<td>0.07168</td>
<td>0.06907</td>
</tr>
<tr>
<td>Germany</td>
<td>0.06528</td>
<td>0.11435</td>
<td>0.07741</td>
<td>0.09834</td>
<td>0.03069</td>
</tr>
<tr>
<td>Greece</td>
<td>0.0696</td>
<td>0.09012</td>
<td>0.00460</td>
<td>0.00497</td>
<td>0.00922</td>
</tr>
<tr>
<td>Ireland</td>
<td>0.00024</td>
<td>0.00043</td>
<td>0.00032</td>
<td>0.00038</td>
<td>0.00031</td>
</tr>
<tr>
<td>Iceland</td>
<td>0.00197</td>
<td>0.00255</td>
<td>0.00208</td>
<td>0.00275</td>
<td>0.00296</td>
</tr>
<tr>
<td>Italy</td>
<td>0.09404</td>
<td>0.06969</td>
<td>0.04719</td>
<td>0.06859</td>
<td>0.05890</td>
</tr>
<tr>
<td>Japan</td>
<td>0.09890</td>
<td>0.15339</td>
<td>0.14985</td>
<td>0.17858</td>
<td>0.21015</td>
</tr>
<tr>
<td>Luxembourg</td>
<td>0.00956</td>
<td>0.05685</td>
<td>0.00442</td>
<td>0.00682</td>
<td>0.00578</td>
</tr>
<tr>
<td>Norway</td>
<td>0.01563</td>
<td>0.02364</td>
<td>0.00424</td>
<td>0.01791</td>
<td>0.01792</td>
</tr>
<tr>
<td>Netherlands</td>
<td>0.00285</td>
<td>0.00292</td>
<td>0.00249</td>
<td>0.00259</td>
<td>0.00270</td>
</tr>
<tr>
<td>New Zealand</td>
<td>0.00570</td>
<td>0.06709</td>
<td>0.00726</td>
<td>0.00649</td>
<td>0.00660</td>
</tr>
<tr>
<td>Portugal</td>
<td>0.00336</td>
<td>0.02583</td>
<td>0.00963</td>
<td>0.00415</td>
<td>0.00471</td>
</tr>
<tr>
<td>Spain</td>
<td>0.01790</td>
<td>0.02966</td>
<td>0.01843</td>
<td>0.02950</td>
<td>0.02158</td>
</tr>
<tr>
<td>Sweden</td>
<td>0.01550</td>
<td>0.01581</td>
<td>0.01179</td>
<td>0.01373</td>
<td>0.01041</td>
</tr>
<tr>
<td>Switzerland</td>
<td>0.01100</td>
<td>0.01533</td>
<td>0.01793</td>
<td>0.03700</td>
<td>0.02383</td>
</tr>
<tr>
<td>Turkey</td>
<td>0.00825</td>
<td>0.00906</td>
<td>0.00947</td>
<td>0.00902</td>
<td>0.00703</td>
</tr>
<tr>
<td>UK</td>
<td>0.05703</td>
<td>0.00769</td>
<td>0.00580</td>
<td>0.05849</td>
<td>0.04861</td>
</tr>
<tr>
<td>USA</td>
<td>0.04449</td>
<td>0.34113</td>
<td>0.44998</td>
<td>0.35390</td>
<td>0.31646</td>
</tr>
</tbody>
</table>
NOTES

The author gratefully acknowledges two anonymous referees whose comments helped to improve the content and style of this work.

REFERENCES


THE THEIL INDEX IN SEQUENCES OF NESTED AND HIERARCHIC GROUPING STRUCTURES: IMPLICATIONS FOR THE MEASUREMENT OF INEQUALITY THROUGH TIME, WITH DATA AGGREGATED AT DIFFERENT LEVELS OF INDUSTRIAL CLASSIFICATION

Pedro Conceição
Institut Superior Teneio, Lisbon

James K. Galbraith
The University of Texas at Austin

Peter Bradford
IBM Corporation

The Theil index, when applied to measure the between-industry dispersion of wages, can be used to construct long and dense time-series of inequality. To use the Theil index in such a way measures inequality between sectors, but fails to capture the level of inequality within each of the sectors. However, Conceição and Galbraith [2000] provide formal criteria under which the between sector component of the Theil index tracks the overall movement of inequality, concluding that, under some very general conditions, the dynamics of overall inequality can be captured using only the between sector component of the Theil index.

This paper deepens and extends that argument, explores the fractal properties of the Theil index, and presents a more compelling empirical illustration. The fractal property of the Theil index results directly from its characteristic of perfect decomposability, which allows for the separation of inequality into between- and within-groups components, provided that the groups are mutually exclusive and completely exhaustive (MECE). First, we explore the fractal property of the Theil index in the context of nested and hierarchical groups—of relevance when dealing with groups of industries aggregated at different levels of SIC codes—and also of non-hierarchical grouping structures. Finally, we provide an empirical illustration, based on a monthly time-series of industrial earnings inequality for the United States, from January 1947.

James K. Galbraith: Ed School of Public Affairs, University of Texas at Austin, Austin, TX 78712.
E-mail: galbraith@mail.utexas.edu

Eastern Economic Journal, Vol. 27, No. 4, Fall 2001