

# Articles

## WHAT DO PRODUCTIVITY SHOCKS TELL US ABOUT THE SAVING-INVESTMENT RELATIONSHIP?

Lutfi Erden, Ibrahim Ozkan, Burak Gunalp\*

### Abstract:

This study is a contribution to the empirical literature on the significance of productivity shocks in explaining a high saving-investment correlation, using data from a panel of 21 OECD countries over the period 1970–2003. The study looks at the distributional properties of the productivity shocks in order to test if productivity shocks can relate saving to investment. To this end, we divide the countries into three groups with respect to the distributional characteristics of productivity shocks in each country with an application of the Fuzzy-c-means (FCM) clustering technique. The results provide some support for the productivity shock argument, indicating that the saving retention coefficients are greater for the countries subject to large productivity shocks in magnitude.

**Keywords:** Feldstein-Horioka Puzzle; Productivity Shocks; Fuzzy Clustering; International Capital Mobility; Kolmogorov-Smirnov Statistics

**JEL Classification:** C23; C40; F21; F41

### 1. Introduction

Feldstein and Horioka (1980; F-H hereafter) were the first to investigate the relationship between domestic saving and investment (S-I hereafter), in order to test the degree of international capital mobility. In a closed economy, domestic investment is financed entirely by domestic saving. Therefore, the correlation between the two variables may be expected to be high. However, in an open economy, where domestic investment can be financed not only by domestic saving but also by foreign saving, the high association between S-I may be expected to be quite low. F-H estimated the following equation to test their assertion:

$$inv_i = \alpha + \beta sav_i + u_i \quad (1)$$

where  $inv_i$  and  $sav_i$  are the ratios of investment and savings to GNP, respectively, and  $u_i$  is the random error term. The parameter  $\beta$  is known as the saving retention

\* Hacettepe University, Department of Economics, Beytepe, Ankara, Turkey (lerden@hacettepe.edu.tr; iozkan@hacettepe.edu.tr; gunalp@hacettepe.edu.tr)

coefficient considered to be indicating the degree of international capital mobility. F-H estimated this equation for 16 OECD countries and found a large coefficient. They interpreted this result as evidence of imperfect capital mobility across these relatively industrialized economies, a surprising result that is known as the F-H puzzle.

This has stimulated a growing interest in the literature on the degree of international capital mobility. One strand of studies has put forward productivity shocks as well as country size to explain how a high correlation between S-I can be compatible in a world with highly integrated international capital markets (Obstfeld, 1986; Mendoza, 1991, 1997; Baxter and Crucini, 1993; Glick and Rogoff, 1995; Gregory and Head, 1999; Miniane, 2004).<sup>1</sup> These simulation studies have constructed dynamic stochastic equilibrium models to investigate the association between saving-investment and investment-current account deficits. They have consistently shown that productivity shocks can explain a high saving-investment interaction irrespectively of international capital mobility, a conclusion that is reached under different specifications for the productivity shocks such as temporary or permanent, persistent or not and global or country specific. For instance, if countries experience common (global) shocks that are symmetric, investment demand in each country rises.<sup>2</sup> However, since each country ends up financing the increases in investment expenditure in domestic economy, their current accounts remain stable. Thus, in response to a global shock, investment and savings are closely related. On the other hand, a country specific shock spurs investment spending. This may have positive or negative impact on the current account, which in turn plays a trivial role on the link between saving and investment. In addition, a temporary shock to labour and/or technology that persists some time increases output more than consumption due to consumption smoothing behavior so that the country realizes a rise in saving. Also with the shock, marginal productivity of capital goes up, giving rise to more investment spending. This provides a natural link between S-I. These studies place too much emphasis on the persistence (propagation) of productivity shocks to produce a positive response of saving. Thus, they do not predict that saving responds to a permanent but not persistent shock. However, Miniane (2004) has demonstrated that a permanent shock that is not persistent can also generate a positive saving reaction thanks to the learning process. A learning agent might view part of the permanent shock as temporary due to imperfect information and hence initial response of saving to a permanent shock might be positive. Furthermore, consumption habit provides another explanation for a positive saving response to a permanent shock (Gruber, 2002).

Overall, theoretical studies entertaining different assumptions about the type of productivity shock such as persistent or not persistent and temporary or permanent conclude that productivity shocks provide a natural channel that relates investment to saving. Nonetheless, these studies lack support from the empirical studies that employ econometric methods. There are few studies in the literature emphasizing and empirically testing the effect of productivity shocks as a potential explanation for the

1 Another strand of studies has criticized the F-H interpretation of the high the correlation between S-I as a measure of the degree of financial integration. For example, Coakley *et al.* (1996) argue that the observed high association between S-I may be an indication of current account solvency. Artis and Bayoumi (1990), on the other hand, contend that it could be a result of the policy actions targeting current account balance. See Coakley *et al.* (1998) for an excellent survey of this literature.

2 With a shock, we mean a positive productivity shock throughout the paper.

F-H puzzle (Kim, 2001 for OECD countries; Kim *et al.*, 2007 for East Asian countries; and Fouquau *et al.*, 2007). Kim (2001) and Kim *et al.* (2007), after eliminating productivity shocks from investment and saving series, found that the saving retention coefficient declines only slightly. Hence, they conclude that productivity shocks are not responsible for high S-I correlation. Fouquau *et al.* (2007) employed a panel smooth transition regression S-I model in which one of the transitioning variables is output growth used as a proxy for productivity shocks. They found that the saving retention coefficient does not alter in response to productivity shocks.

In this paper, since the simulation studies suggest that productivity shocks potentially explain the puzzle irrespective of their types, we do not attempt to assess which types of productivity shocks have an impact on the S-I relationship. Therefore we aim to empirically test the significance of productivity shocks themselves in explaining the relationship between saving and investment. In doing so, we adopt a novel approach in which the distributional characteristics of productivity shocks play a key role since all information about a random variable is contained in its distribution. If the productivity shocks are important, then different distributions from which the productivity shocks are drawn lead to different saving retention coefficients. To this end, we apply Fuzzy c-means (FCM) clustering algorithm to panel data from 21 OECD countries over the period between 1970 and 2003. To the best of our knowledge, this paper constitutes the first effort using fuzzy system modelling to investigate whether the productivity shocks are important for the saving retention coefficient.

We conclude with the following insights: (i) productivity shocks are important but (ii) the variability of the productivity shocks is not, (iii) it is the expected value of the productivity shocks that matters. The saving retention coefficients are higher for the countries with greater expected productivity shocks. The rest of the paper is organized as follows. Section 2 introduces data and the methodology. The results are presented in Section 3. Section 4 concludes.

## 2. Data and Methodology

The data used in this study include the ratios of saving and investment to GNPs and labour productivity indices for 21 OECD countries over the period 1970-2003. All series are obtained from the OECD database ([www.sourceoecd.org](http://www.sourceoecd.org)). In order to remove the deterministic trend in productivity index, the first difference of the index is taken, which is used as a proxy for productivity shocks.

The methodology consists of the following steps. The first step is to identify the different groups based on the distribution of productivity shocks through applying Fuzzy c-means (FCM) clustering algorithm to panel data from 21 OECD countries over the period between 1970 and 2003. In the second step, we obtain the dummy variables that differentiate the countries based on their shock distributions via the membership function. The membership values of all countries for each year are chosen as the dummy variables instead of defining a dichotomous variable for the distribution that takes a value of 0 or 1 only. This technique produces the degree of belongingness of countries within the range of 0 and 1 based on the distribution of productivity shocks. After the application of the FCM, new variables (dummies) representing the distribution are produced. In order to measure similarities of distribution of productivity

shocks, we use a modified version of the D-statistic of Kolmogorov-Smirnov. Finally, conventional panel estimation methods are applied to the model that incorporates the productivity shock dummies.

## 2.1 Fuzzy c-means

One of the most important aspects of fuzzy system analysis is that it can capture the underlying behaviour of data without relying on excessive *ad hoc* axioms. One can highlight at least two advantages of FCM: (i) it has a high potential for linguistic explanation (such as big, small, high, low, *etc.*) that brings about a considerable ease of understanding, and (ii) it can cope with noisy and imprecise data (since the boundaries of fuzzy sets are vague) thereby providing flexibility and stability for classification and prediction.<sup>3</sup> For example, classification of the countries based on their productivity shocks distributions may lead to groups such as, “*high*”, “*low*” and “*high and variable*” shocks countries. Even though the difference between these words is not clear-cut, it can be appropriately analyzed by using fuzzy logic. A “*crisp*” differentiation of the groups implies that the gray areas between these groups are simply ignored. A result of classifying the groups into sharp clusters is that a very small change in similarity causes a change in the groups. This implies that the dichotomous variable representing the classification changes its value from 0 to 1 even for a negligibly small change in similarity. Besides, observations may involve some errors as a result of data collection and data aggregation. Consequently, the productivity indices of the countries should be regarded as imprecise data, and hence they should be treated as imprecise.

As defined by Zadeh (1965), a fuzzy set of a set  $X$  is a mapping  $\mu : X \rightarrow [0,1]$ . The set of all fuzzy sets of  $X$  is denoted by  $F(X) := \{X \mid \mu : X \rightarrow [0,1]\}$ . Membership value  $\mu_C \in [0,1]$  then represents the degree of belongingness to  $C$  where a value close to 0 means a low degree of membership while a value close to 1 means a high degree of membership; and in between 0 and 1 there is a smooth transition of membership values.

Hence the fuzzy analysis is defined as the further analysis space,

$A_{Fuzzy}(D, R) := A(D, \{F(K) \mid K \in R\})$  where  $D$  is a data space and  $R$  is a result

space. The result of an analysis is then  $f : X \rightarrow F(K)$  for  $X \subseteq D$  and  $K \in R$ . If

$\forall x \in X : \sum_{k \in K} f(x)(k) = 1$  and  $\forall x \in X : \sum_{x \in X} f(x)(k) > 0$  where  $f(x)(k)$  is interpreted as the degree of membership of the datum  $x \in X$  to the cluster  $k \in K$ , then this type of clustering is called probabilistic clustering (Höppner et al., 1999).

Fuzzy c-means (FCM) clustering generates partitions by means of a membership value between 0 and 1 assigned to each observation to each cluster. The membership values indicate the degree of belongingness of each observation to each of the clusters. As the membership value gets higher, the degree of belongingness increases. Bezdek (1973) showed that in the minimization of the objective function;

3 Due to these attracting features, fuzzy system analysis has found widespread application in a variety of fields such as computer science, system analysis, electronic engineering, pharmacology, finance and more recently social sciences (*i.e.* Artis and Zhang (2001); Bezdek (1973); Ozkan and Turksen (2007); Sproule *et al.* (2002) and Tron and Margaliot (2004)).

$$J_m(U, V; X) = \sum_{k=1}^{nd} \sum_{c=1}^{nc} \mu_{c,k}^m \|x_k - v_c\|_A^2 \quad (2)$$

where,  $\mu_{c,k}$  is the membership value of the  $k^{\text{th}}$  vector in the  $c^{\text{th}}$  cluster such that  $\mu_{c,k} \in [0,1]$ ,  $nd$  is the number of vectors,  $nc$  is the number of clusters,  $\|\cdot\|_A$  is norm (or the difference based on the similarity defined) and  $m$  is the level of fuzziness, the membership function is calculated as:

$$\mu_{i,k} = \left[ \sum_{c=1}^{nc} \left( \frac{\|x_k - v_i\|_A}{\|x_k - v_c\|_A} \right)^{\frac{2}{m-1}} \right]^{-1} \quad (3)$$

where  $\sum_{c=1}^{nc} \mu_{k,c} = 1$  for some given  $m > 1$ .

In fuzzy clustering analysis, the number of clusters and the level of fuzziness need to be identified before clustering. In the literature, several cluster validity indices have been introduced to identify the number of clusters (Bezdek, 1974, 1975; Fukuyama and Sugeno, 1989); and very limited studies have been carried out for the level of fuzziness (Ozkan and Turksen, 2004, 2007). The most widely used value for the level of fuzziness is two. This value is usually accepted as the rule of thumb. However it was shown that the proper values for upper and lower bounds of level of fuzziness are 1.4 and 2.6 respectively (Ozkan and Turksen, 2004, 2007). If the clustering algorithm does generate relatively different groups, then the analysis should be performed with both upper and lower values of level of fuzziness since the information gained would be important.

## 2.2 Kolmogorov-Smirnov Test (KS test)

Because we are interested in the distribution groups, the similarity of distributions is important in this analysis. A modified version of the D-statistic of the KS test can be used as the similarity measure between any two distributions. The KS test is a non-parametric statistical test that enables us to determine whether two probability distributions come from the same distribution. More formally, the null hypothesis that two samples are withdrawn from the same distribution is tested against the alternative.

$$H_0 : F_i = F_j$$

where  $F_i$  and  $F_j$  are empirical cumulative distribution functions defined as

$$F(x) = \frac{1}{N} \sum_{i=1}^N I_{X_i \leq x} \quad \text{where } N \text{ is the sample size, } I_{X_i \leq x} = 1 \text{ if } X_i \leq x \text{ and } I_{X_i \leq x} = 0$$

otherwise. The KS test is performed based on the D-statistic which is defined as:

$$D = \sup_x |F_i - F_j|$$

The D-statistic is the absolute value of the maximum difference between two empirical

cumulative functions. The null hypothesis is rejected at level  $\alpha$  if  $\sqrt{\frac{N_1 N_2}{N_1 + N_2}} D > cv_\alpha$ ,

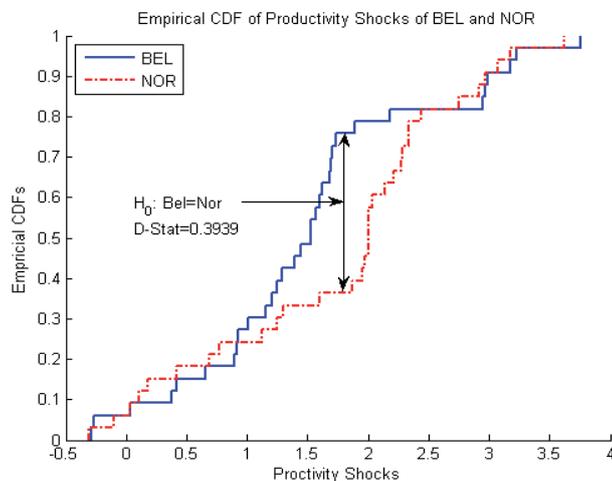
where  $N_1$  and  $N_2$  are the sizes of samples  $i$  and  $j$ , respectively, and  $cv_\alpha$  is the critical value of the Kolmogorov distribution.

As a similarity measure for the productivity shocks distributions, the modified D-statistic is used. Since this work is expected to be sensitive to the mean values of the productivity shocks, the D-statistic is modified so that it can take negative values. Hence the D-statistic for the null hypothesis  $H_0 : F_i = F_j$  is the negative of the D-statistic for the null hypothesis  $H_0 : F_j = F_i$ .

In this way, it becomes important in the test whether the empirical cumulative distribution is on the left or right side of the other empirical cumulative distribution.<sup>4</sup>

Figure 1

**D-statistic for  $H_0 : F_{BELGIUM} = F_{NORWAY}$**



The D-Statistic calculated for the null hypothesis  $H_0 : F_{BELGIUM} = F_{NORWAY}$  is given in Figure 1 to exemplify the test. We modified the D-statistic in order to have the negative value. Then, the D-statistic for  $H_0 : F_{NORWAY} = F_{BELGIUM}$  is  $-0.3939$ , which is exactly the negative of the D-statistics for  $H_0 : F_{BELGIUM} = F_{NORWAY}$ .

### 2.3 Panel Regression

The effect of productivity shocks is incorporated into equation (1) as the threshold variable. The saving-retention coefficients are then distinguished by differing productivity shock regimes. Hence, equation (1) can be redefined as follows,

$$inv_{it} = \alpha_i + \beta_1 sav_{it} * d_{1,it} + \beta_2 sav_{it} * d_{2,it} + \beta_3 sav_{it} * d_{3,it} + u_{it} \quad (4)$$

4 P-value of an observed value of D is calculated approximately as:

$$P(D > observed) = 2 \sum_{j=1}^{101} (-1)^{j-1} e^{-2\lambda^2 j^2}, \text{ where } \lambda = \max \left\{ \left( \sqrt{N} + 0.12 + \frac{0.11}{\sqrt{N}} \right) D, 0 \right\} \text{ and } N = \frac{N_1 N_2}{N_1 + N_2}$$

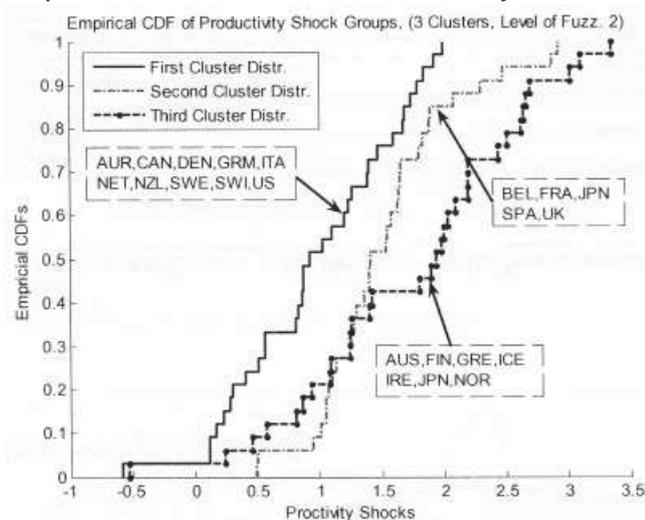
where  $N_1$  and  $N_2$  are the sample sizes. As  $N$  gets large, the approximation becomes asymptotically accurate (Press et al., 2007). However, the calculation gives reasonably accurate result even for  $N \geq 4$  (See Press et al., (2007) and Stephens (1970) for more details).

where  $d_{j,i} \in [0,1]$  is the degree of belongingness of country  $i$  to the cluster  $j$  at year  $t$ , and  $\sum_{j=1}^3 d_{j,it} = 1$  and  $t = 1, \dots, T$ . As mentioned previously, the FCM clustering algorithm is applied to the productivity shocks of each country over all years from 1970 to 2003 in order to obtain these membership values. This specification allows the saving retention coefficients to smoothly change depending on productivity shock clusters.

### 3. Empirical Results

FCM with modified D-statistic as a similarity measure is applied to the distributions of the productivity shocks. In our analysis, the number of clusters is found to be three and the countries are partitioned as, “*high*”, “*low*” and “*high and variable*” shocks countries. The levels of fuzziness are chosen to be 1.4, 2.0 and 2.6 in value. Cluster 1 contains Australia, Canada, Denmark, Germany, Italy, the Netherlands, New Zealand, Sweden, Switzerland and the United States. Cluster 2 contains Belgium, France, Japan, Spain and the United Kingdom. Finally, Cluster 3 contains Austria, Finland, Greece, Iceland, Ireland, Japan and Norway.<sup>5</sup> Figure 2 shows the clusters’ representative empirical cumulative distributions obtained as the shocks weighted by membership values. As it is observed from the figure, the mean of the shock distributions belonging to Cluster 1 is significantly smaller than those belonging to the other clusters. Third cluster has a significantly different variability of the productivity shocks. This type of clustered structure provides us to gain more insights since two clusters, Cluster 2 and 3, have similar means and cluster 1 and 2 have similar variances. Descriptive statistics of the three clusters are given in Table 1.

Figure 2  
Empirical Cumulative Distribution of Productivity Shock Clusters



5 To separate the cluster, the threshold for membership values is set to 1/3 where 3 is the number of clusters.

Table 1  
**Descriptive Statistics of the Cluster**

	<b>Cluster 1</b>	<b>Cluster 2</b>	<b>Cluster 3</b>
Minimum	-0.5750	0.4853	-0.5273
Maximum	1.9975	2.9270	3.3527
Mean	0.9758	1.5247	1.7420
Standard Deviation	0.6416	0.5648	0.9024

This characterization helps us to understand what is important for the S-I relationship. If the mean value of the shocks is important, then the countries belonging to Cluster 1 have different saving retention coefficients than those belonging to Clusters 2 and 3. If the variability is important rather than the mean, then Clusters 1 and 2 have similar coefficient values while Cluster 3 should differ. If both the mean value and the variability of shocks are important, then all clusters should have different saving retention coefficients.

In order to check whether the clusters' sample distributions are coming from the same underlying distribution, we perform the hypothesis tests shown in Table 2. It can be seen from the table that all clusters' productivity shocks are from different distributions. The null hypotheses that they are drawn from the same distributions are rejected in all three tests with at least 99% confidence. Hence, we can assume that all clusters have different underlying distributions.

Table 2  
**Tests for the Equality of Clusters' Distributions**

$H_0 : F_2 = F_3$ $H_a : F_2 \neq F_3$	$H_0 : F_1 = F_2$ $H_a : F_1 \neq F_2$	$H_0 : F_1 = F_3$ $H_a : F_1 \neq F_3$	$H_0 : F_2 = F_3$ $H_a : F_2 \neq F_3$
Test Result	Rejected	Rejected	Rejected
p-value	0.0013	0.0005	0.0082

The continuous membership functions of three clusters based on the D-statistic obtained with the countries' productivity shocks and the average of the all countries' productivity shocks are given in Figure 3 without removing the harmonics. The first center is near -0.35, the second is about at 0.2, and the third one is about 0.37.

As an example, the membership matrix obtained from the FCM clustering technique with the level of fuzziness set to two is given in Table 3. The degrees of membership of Austria and Finland to Clusters 2 and 3 are almost the same. Japan as a very big economy in the sample has a significant degree of membership to Cluster 3 which is composed of relatively small countries.

Figure 3  
Membership Function for D-statistic

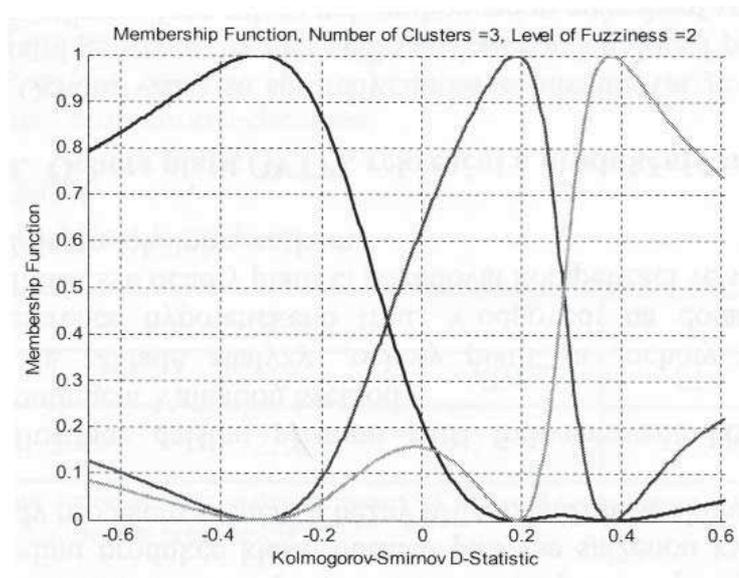


Table 3  
Membership Matrix

Country	Cluster 1	Cluster 2	Cluster 3	Country	Cluster 1	Cluster 2	Cluster 3
Austria	0.00996	0.30487	0.68518	Italy	0.98118	0.01252	0.00630
Australia	0.94864	0.03462	0.01674	Japan	0.01185	0.61157	0.37659
Belgium	0.00023	0.9982	0.00154	Netherlands	0.97852	0.01368	0.00780
Canada	0.96019	0.02513	0.01468	Norway	0.00354	0.03612	0.96034
Denmark	0.99960	0.00026	0.00012	N. Zealand	0.99699	0.00197	0.00103
Finland	0.00996	0.30487	0.68518	Spain	0.00173	0.98027	0.01800
France	0.02145	0.91418	0.06437	Sweden	0.94864	0.03462	0.01674
Greece	0.00354	0.03612	0.96034	Switzerland	0.91701	0.05155	0.03144
Germany	0.99442	0.00368	0.00190	UK	0.00023	0.99823	0.00154
Iceland	0.00431	0.09025	0.90544	US	0.98118	0.01252	0.00630
Ireland	0.00354	0.03612	0.96034				

Note: Values are rounded to the nearest digit. The level of fuzziness is 2.

Before we proceed with estimating equation (4), we apply panel unit root tests to our series to ascertain that the estimation results do not suffer from a spurious regression problem. We test the null hypothesis of a unit root by using three standard panel unit root tests: LLC (Levin *et al.*, 2002), IPS (Im *et al.*, 2003) and Fisher ADF (Maddala and Wu, 1999). LLC assume a common unit root process for the series; in

contrast, IPS and ADF allow the presence of individual unit root processes across cross section units. The following equation provides a basis for all these tests:

$$\Delta x_{it} = \delta_i + \phi x_{it-1} + \sum_{j=1}^{pi} \gamma_{ij} \Delta x_{it-j} + x_{it}' \lambda + \varepsilon_{it} \quad (5)$$

The results from the panel unit root tests are presented in Table 4. The null hypothesis of unit root can be rejected at least at the 5% level of significance for both *inv* and *sav* series. Therefore, equation (4) can be estimated in the levels of the variables by employing standard panel estimation techniques.

Table 4  
Panel Unit Root Tests

Inv			sav		
LLC	IPS	ADF	LLC	IPS	ADF
-2.49*	-2.12**	57.8**	-4.19*	-3.41*	71.9*

Note: Schwartz information criteria (SIC) is used for optimum lag selection. \* and \*\* show 1% and 5% levels of significance respectively.

Now we can turn to estimation of equation (4) under the fixed effects specification. Table 5 presents the results. Each column of the table reports the estimated coefficients using various levels of fuzziness. The results from equation (4) show that the saving retention coefficient is responsive to productivity shocks. More specifically, the coefficient varies depending on the productivity shocks distributions that the countries are subject to. They significantly differ from each other, ranging from 0.34 to 0.70. This can be taken as evidence that the productivity shocks may be one of the important factors that links domestic saving to investment independently of the degree of financial openness in these countries. Hence, although the productivity shocks do not make the coefficients very close to 1, they provide some explanation for the F-H puzzle as suggested in the theoretical literature.

Table 5  
Estimated Saving Retention Coefficients (equation 4)

	Levels of Fuzziness		
	m = 1.4	m = 2.0	m = 2.6
Cluster 1 ( $\hat{\beta}_1$ )	0.442 (0.036)*	0.403 (0.041)*	0.344 (0.049)*
Cluster 2 ( $\hat{\beta}_2$ )	0.616 (0.069)*	0.662 (0.069)*	0.702 (0.073)*
Cluster 3 ( $\hat{\beta}_3$ )	0.613 (0.053)*	0.616 (0.055)*	0.629 (0.060)*
Adjusted R <sup>2</sup>	0.617	0.619	0.622
Total observation	714	714	714

Notes: 1. \* shows 1% level of significance.

2. The figures in parentheses are White-heteroskedasticity consistent standard errors.

3. Test of the null that all coefficients are equal yields an F-statistics of 6.58 with p value<0.01.

4. Test of the null that the coefficients of group 1 and group 2 are equal yields an F-statistics of 0.26 with p value>0.10.

Furthermore, while the null of all coefficients are equal is rejected, we fail to reject the hypothesis that the coefficients of Cluster 2 and 3 are the same. As mentioned previously, the Clusters 2 and 3 have similar variations in the productivity shocks. This suggests that either saving or investment may respond differently to the productivity shocks with high variability so that the coefficients of Clusters 2 and 3 do not differ. Interestingly though, while the coefficient of Cluster 1 with low expected value of the shocks is found to be small, those of Clusters 2 and 3 with high expected values are large. This finding indicates that the coefficient gets larger for the countries subject to large productivity shocks, a result that is in line with the simulation studies by Mendoza (1997) and Miniane (2004). This implies that it is the magnitude rather than the variability of shocks that matters. Overall, productivity shocks seem to provide an important channel that leads saving and investment to closely hanging together, shedding some light on the F-H puzzle.

Together with the productivity shocks, the country size is considered to be another important reason for the co-movement of S-I (Baxter and Crucini, 1993; Bahmani-Oskooee and Chakrabarti, 2005; Ho, 2002 and Ho and Huang, 2006). Since large countries can influence the world interest rate, any gap between S-I in large countries can be closed up by the movements in interest rates. In this case, it is natural to expect a high saving retention coefficient for a large country. Thus, one needs to check whether the productivity shocks can still provide a mechanism that moves S-I together even after controlling for the country size. To this end, the countries are clustered with respect to their relative gross domestic products taken to represent country size, employing the Fuzzy-c means clustering technique. Although the memberships are time-variant dummies, Table 6 presents the results with the fuzziness level of 2.0 to exemplify how the classification looks like in 2003 only.

Table 6  
Classification of the Countries Based on Fuzzy Size

Groups	Number of Countries	Countries
Small Size	15	Austria, Australia, Belgium, <b>Canada</b> , Denmark, Finland, Greece, Iceland, Ireland, Norway, Netherlands, New Zealand, <b>Spain</b> , Sweden, Switzerland
Medium Size	6	<b>Canada</b> , France, <i>Germany</i> , Italy, <b>Spain</b> , UK
Large Size	3	<i>Germany</i> , Japan, US

Note: Countries in bold belong to both small and medium groups since their membership values are greater than  $\frac{1}{4}$  (1/# of clusters). Countries in italic belong to both medium and large groups since its membership values are greater than  $\frac{1}{4}$ . Since the US is the only member of the fourth cluster, presenting itself as an outlier, we merge it into the large size cluster.

Since there are too many overlaps in the size and productivity clusters, equation (1) cannot be modified to incorporate both size and productivity dummies. Instead, we perform this test by dividing the sample into clusters with respect to size that do not overlap with the productivity shock clusters. It turns out that 15 small sized countries meet this requirement. Also, because it is already evident that the productivity shock Clusters 2 and 3 have similar coefficients, these clusters are merged and named Cluster

2. Table 7 reports the results. As seen, within the cluster of the same sized countries, coefficients get larger for those that incur large productivity shocks. Therefore, the productivity shocks can affect the S-I relationship independently of the country size, a result that is in agreement with Baxter and Crucini (1993).

Table 7  
The Response of the Coefficient to Productivity Shocks Given Small Sized Countries

Productivity Groups	$\hat{\beta}$
Cluster 1 ( $\hat{\beta}_1$ )	0.433 (0.041)*
Cluster 2 ( $\hat{\beta}_2$ )	0.558 (0.057)*
Adjusted R <sup>2</sup>	0.446
Total observation	510

Note: Test of the null that the coefficients of group 1 and group 2 are equal yields an F-statistic of 3.125 with p value < 0.10.

Finally, according to Baxter and Crucini (1993), productivity shocks can play a more significant role on the S-I relationship in larger countries: *“In a large country there is a secondary effect which stems from the fact that technology shocks in a large country have a nontrivial short-term effect on the world interest rate... Because the large country faces an imperfectly inelastic point-in-time supply curve for capital, and because saving rises with interest rate, ... national saving and national investment are more highly correlated for larger countries”* (p.428).” To test whether the country size can relate S-I above and beyond productivity shocks, we divide the sample with respect to productivity shocks that does not overlap with the size clusters. The results are presented in Table 8. The findings indicate that the coefficient gets larger for larger countries that are subject to similar productivity shocks.

Table 8  
The Response of the Coefficient to Country Size Given Similar Productivity Shock Countries

Size Groups	$\hat{\beta}$
Small ( $\hat{\beta}_1$ )	0.543 (0.039)*
Medium ( $\hat{\beta}_2$ )	0.676 (0.038)*
Large ( $\hat{\beta}_3$ )	0.766 (0.031)*
Adjusted R <sup>2</sup>	0.776
Total observation	510

#### 4. Conclusion

It is not uncommon that one encounters a puzzle in international finance literature. The F-H puzzle is one of the six major puzzles identified by Obstfeld and Rogoff (2000). Although there are several potential explanations put forward for this puzzle,

productivity shock argument appears to be more prominent particularly in real business cycle literature. Simulation studies emphasize the importance of productivity shocks as an explanation for the puzzle without empirical evidence. In this paper, we empirically test the question of whether productivity shocks play an important role in solving the F-H puzzle. To do so we take a distinct approach that consists of three steps.

In the first step, we divide the countries into clusters based on their productivity shock distributions with an application of an unsupervised learning method: Fuzzy C-means. In the second step, the membership values of the countries to the clusters are obtained. In the last step, we estimated the panel model that incorporates the fuzzy membership values as dummy variables representing productivity shocks. The data reveal three productivity shock clusters. Clusters 2 and 3 have similar means while Clusters 1 and 2 have similar variances. The findings indicate that the saving retention coefficient is sensitive to the expected value of productivity shocks but not to the variability of shocks. More specifically, the coefficient gets larger with the magnitude of shocks. In addition, the results show that country size influences the coefficient in a given productivity shock cluster. Further, productivity shocks affect the coefficient in a given size cluster.

In the light of these analyses, this paper provides empirical support to the simulation studies reporting the significance of productivity shocks in explaining the F-H puzzle. Country size also seems to have an effect on the saving and investment correlation independent of productivity shocks. Thus, it appears irrelevant to look at domestic saving-investment interaction in order to test the degree of international capital mobility since these variables are only a part of a complex system in which country size and/or productivity shocks can move them together irrespectively of international financial openness. Since both productivity shocks and size contribute to explaining the puzzle, for future research, this analysis can be extended to consider clustering countries based on these variables together.

## References

- Artis, M. J., Zhang, W.** (2001), "Core and Periphery in EMU: a Cluster Analysis." *Economic Issues* 6 (2), pp. 39-60.
- Artis, M. J., Bayoumi, T. A.** (1990), "Saving, Investment, Financial Integration and the Balance of Payments." IMF Staff Studies for World Economic Outlook, pp. 19-34.
- Bahmani-Oskooee, M., Chakrabarti, A.** (2005), "Openness, Size and the Saving-Investment Relationship." *Economic Systems* 29, pp. 283-293.
- Baxter, M., Crucini, M.** (1993), "Explaining Saving-Investment Correlations." *American Economic Review* 83, pp. 416-436.
- Bezdek, J. C.** (1973), "Fuzzy Mathematics in Pattern Recognition." Ph.D. Thesis, Applied Mathematics Center, Cornell University.
- Bezdek, J. C.** (1974), "Cluster Validity with Fuzzy Sets." *Journal of Cybernetics*, 3 (3), pp. 58-72.
- Bezdek, J. C.** (1975), "Mathematical Models for Systematics and Taxonomy." In: Estabrook, G. (Ed.), *Proceedings on 8th International Conference on Numerical Taxonomy*. San Francisco, CA: Freeman, 1975.
- Coakley, J., Kulasi, F., Smith, R.** (1996), "Current Account Solvency and the Feldstein-Horioka Puzzle." *Economic Journal*, 106, pp. 620-627.
- Coakley, J., Kulasi, F., Smith, R.** (1998), "The Feldstein-Horioka Puzzle and Capital Mobility: A Review." *International Journal of Finance and Economics*, 1998/3, pp. 169-188.

- Feldstein, M., Horioka, C.** (1980), "Domestic Saving and International Capital Flows." *Economic Journal* 90, pp. 314-329.
- Fukuyama, Y., Sugeno, M.** (1989), "A New Method of Choosing the Number of Clusters for the Fuzzy c-means Method." Proceedings of the Fifth Fuzzy Systems Symposium, pp. 247-50.
- Fouquau, J., Hurlin, C., Rabaud, I.** (2007), "The Feldstein-Horioka Puzzle: A Panel Smooth Transition Regression Approach." *Economic Modeling*, doi:10.1016/j.econmod.2007.06.008.
- Glick, R., Rogoff, K.** (1995), "Global versus Country-Specific Productivity Shocks and the Current Account." *Journal of Monetary Economics* 35, pp. 159-192.
- Gregory, A. W., Head, A. C.** (1999), "Common and Country-Specific Fluctuations in Productivity, Investment, and the Current Account." *Journal of Monetary Economics* 44, pp. 423-451.
- Gruber J. W.** (2002), "Productivity Shocks, Habits, and the Current Account." Board of Governors of the Federal Reserve System, International Finance Discussion Papers, Vol. 733.
- Ho, T.** (2002), "The Feldstein-Horioka Puzzle Revisited." *Journal of International Money and Finance* 21, pp. 555-564.
- Ho, T., Huang, H.** (2006), "The smooth-saving-retention coefficient with country-size." *Applied Economics Letters* 13, pp. 247-50.
- Höppner F., Klawonn F., Kruss R., Rungler T.** (1999), *Fuzzy Clustering Analysis*. Chichester, UK: John Wiley & Sons Ltd.
- Im, K. S., Pesaran, M. H., Shin, Y.** (2003), "Testing for Unit Roots in Heterogeneous Panels." *Journal of Econometrics* 115, pp. 53-74.
- Kim, S. H.** (2001), "The Saving-Investment Correlation Puzzle Is Still a Puzzle." *Journal of International Money and Finance* 20, pp. 1017-34.
- Kim, S., Kim, S. H., Wang, Y.** (2007), "Saving, Investment and International Capital Mobility in East Asia." *Japan and the World Economy* 19, pp. 279-91.
- Levin, A., Lin, C. F., Chu, C. S. J.** (2002), "Unit-Root Test in Panel Data: Asymptotic and Finite Sample Properties." *Journal of Econometrics* 108, pp. 1-24.
- Maddala, G. S., Wu, S.** (1999), "A Comparative Study of Unit Root Tests with Panel Data and a New Simple Test." *Oxford Bulletin of Economics and Statistics* 61 (Special Issue, November), pp. 631-652.
- Mendoza E. G.** (1991), "Real Business Cycles in the Small Open Economy." *American Economic Review* 81, pp. 797-818.
- Mendoza, E. G.** (1997), "Robustness of Macroeconomic Indicators of Capital Mobility." In: Leiderman, L., Razin, A. (Eds.), *Capital Mobility: The Impact on Consumption, Investment and Growth*. Centre for Economic Policy Research, London, pp. 83-111.
- Miniane, J.** (2004), "Productivity Shocks, Learning, and Open Economy Dynamics." IMF Working Paper, vol. 88.
- Obstfeld, M.** (1986), "Capital Mobility in the World Economy: Theory and Measurement." Carnegie-Rochester Conference Series on Public Policy 24, pp. 55-104.
- Obstfeld, M., Rogoff, K.** (2000), "The Six Major Puzzles in International Macroeconomics: Is There a Common Cause?" NBER Working Paper Series, Vol. 7777.
- Ozkan, I., Turksen, I. B.** (2004), "Entropy Assessment of Type-2 Fuzziness." Fuzz-IEEE 2004, Special Session on Type-2 Applications, Vol. 2, pp. 1111- 1115.
- Ozkan, I., Turksen, I. B.** (2007), "Upper and Lower Valued for the Level of Fuzziness in FCM." In: Wang, P. P., Ruan, D., Kerre, E. E. (Eds.), *Fuzzy Logic: A Spectrum of Theoretical and Practical Issues*. Springer: Studies in Fuzziness and Soft Computing, Vol. 215, No. 16, pp. 85-105.
- Press, H. W., Teukolsky, S. A., Vetterling, W. T., Flannery, B. P.** (2007), "Numerical Recipes, The Art of Scientific Computing." New York: Cambridge University Press.
- Sroule, B. A., Naranjo, C. A., Turksen, I. B.** (2002), "Fuzzy Pharmacology: Theory and Applications." *Trends in Pharmacological Sciences* 23, pp. 412-417.
- Stephens, M. A.** (1970), "Use of Kolmogorov-Smirnov, Cramer-von Mises and Related Statistics without Extensive Tables." *Journal of the Royal Statistical Society B*, 32, pp. 115-122.
- Tron, E., Margaliot, M.** (2004), "Mathematical Modelling of Observed Natural Behaviour: a Fuzzy Logic Approach." *Fuzzy Sets and Systems* 146 (3), pp. 437-450.
- Zadeh, L. A.** (1965), "Fuzzy Sets." *Information and Control*, Vol. 8, pp. 338-353.