The Distribution Sector and the Development Process: Are There Patterns? Yes.*

James H. Anderson	Roger R. Betancourt
World Bank	D. of Economics
1818 H St., NW	University of Maryland
Washington, DC, 20433	College Pk.., MD 20742
Ph: 202 458 5961	Ph: 301 405 3479
Fax:202 522 2751	Fax 301 405 3542
Email: janderson2@ Worldbank.org	Email: betancou@econ.umd.edu

August 1999

* We are grateful to the World Bank Socioeconomic Data Division for providing the Syrquin and Chenery data. An earlier version of the paper was presented at the meetings of the Western Economic Association in San Diego. We thank R. Driskill, J. Nugent, A. Panagariya and Sanjay Jain for their comments while claiming responsibility for any errors. All correspondence should be addressed to Betancourt. This paper has benefitted from support provided by the U.S. Agency for International Development under Cooperative Agreement No. DHR- 0015-0031-00 to the Center on Institutional Reform and the Informal Sector (IRIS) and administered by the Office of Economic and Institutional Reform, Center for Economic Growth, Bureau for Global Programs, Field Support and Research.

Coauthors: James Anderson, Roger Betancourt

Affiliations: World Bank, U. of Maryland

Title: The Distribution Sector and the Development Process: Are There Patterns? Yes.

Abstract:

In this paper we bring together three different strands of literature to raise and answer the question posed in the title. The literature in the Kuznets tradition applied to services raises the question of whether the pattern observed for services, namely a rising share of GDP, is also observed for the distribution sector. The literature on the retail and wholesale sector, on the other hand, indicates the existence of economies of scale in the provision of the distribution services that constitute a main output of this sector. Hence, one would expect a different pattern for distribution than for services because the standard assumption in explaining the pattern for services is that the services sector experiences constant returns to scale. By linking the literature on economic development and specialization with the literature on the distribution sector, we can show that an inverted-U pattern between the share of distribution and the level of development should be expected to arise. We construct a cross-section time series data set for 74 countries using UN national income accounts data to calculate the share of retail and wholesale trade in GDP and merge this data set with the one used by Syrquin and Chenery (1989)to analyze the service sector in general. In contrast to their results for services, we find an average time series relation between the share of distribution and the level of development that exhibits an inverted-U pattern with respect to the level of development. We also analyze the average cross-section (country) relation between the share of distribution and the level of development. In contrast to the results of Kravis, Heston and Summers(1983) for services, we also find an inverted-U pattern between the share of distribution and the level of development in this setting.

JEL Classification Numbers: O10;L81

Key Words: distribution sector; development process; increasing returns; specialization.

Introduction.

In this essay we bring together three separate strands of literature that have a bearing on the issue of whether or not there are patterns in the role of the distribution sector during the course of economic development. The first strand of literature extends the Kuznets (1966) tradition by searching for empirical regularities in the evolution of the share of services in GDP during the process of development. Both Chenery and Syrquin (1975, p.176) and Syrquin and Chenery (1989,pp.21-26) use time-series data for a cross-section of countries to conclude that the share of services rises during the course of economic development, measured by per capita income. Kravis, Heston and Summers(1983) corroborate this finding for a given year (1975)when per capita GDP expenditures are converted into a common unit using exchange rates but find the share of services to be constant across per capita income levels when per capita GDP expenditures are converted at purchasing power parity rates. Panagariya (1988) develops a general equilibrium three sector model that explains how economies of scale in the industrial sector and constant returns in agriculture and services can generate both of these results.

Since services represent a rather heterogeneous category, including telecommunications and domestic service within its fold, the question immediately arises as to whether any of the services subcategories exhibits the same or a different pattern as the aggregate category. The distribution sector, defined as retail plus wholesale trade, is a well established subcategory of services that accounts for a substantial portion of GDP in most countries (13.5% is the average share over the 74 countries and time period, 1950-1983,in our data set). Should we expect the share of the distribution sector in GDP to exhibit the same rising pattern with the level of development as the share of services? By juxtaposing two strands of recent literature we shall argue, on the contrary, that one should expect an inverted-U pattern instead of a rising pattern for the share of the distribution sector as development proceeds.

One strand of recent literature focuses on the economic function of the distribution sector. In this second strand of literature, Betancourt and Gautschi(1988), relying on earlier work by Bucklin (1973), argue that the economic function of the retail sector is to provide consumers with a set of distribution services or implicit outputs together with the explicit items bought at retail. Betancourt (1993) adapts the analysis to encompass the wholesale sector. A key finding in this literature is the existence of economies of scale with respect to the provision of distribution services, Ofer (1973), Oi (1992) and Betancourt and Malanoski (1999). This finding contradicts directly one of the basic assumptions upon which Panagariya bases his explanation of the pattern for services.

Another strand of recent literature focuses on development as specialization and stresses the increasing returns associated with specialization. Ray (1998, Chs.4,5) provides an overview of various aspects of this literature. Romer (1987) is an early contributor to this third strand of literature in the context of endogenous growth models. Locay (1990) stresses that the association of development with specialization leads to an increase in the variety of products available. This increasing variety implies higher levels of assortment need to be provided by the distribution sector and it provides the link between the two recent strands of literature that allows us to explain how the inverted-U pattern emerges. The production of assortment is subject to increasing returns and by embedding this production in a simple economy wide model we generate a share for the distribution sector with the following properties. In the early stages of development it will increase due to specialization; nonetheless, eventually it must decrease due to both an exhaustion of the increasing returns in the production of assortment and to a decline in substitution possibilities between the production of assortment and the production of the items explicitly consumed.

The previous arguments are elaborated in detail in the next section. Subsequently, in Section 2, we present empirical evidence showing that a replication of the Syrquin and Chenery analysis for the distribution sector generates an inverted-U pattern empirically rather than the rising share associated with services. More generally, the average time series relation in our data exhibits an inverted -U pattern. In Section 3 we explore the sensitivity of this result to country coverage, yearly coverage, the use of balanced panels and different functional forms for the dependent and independent variables. In all cases our main finding is confirmed. One by-product of this finding is that country specific factors emerge as important determinants of variations in the share of distribution in GDP.

In Section 4 we confront the inverted-U pattern with cross-section data. This allows us to incorporate variables into the empirical analysis that are suggested by the theoretical discussion and are not available on a time series basis. Some may be viewed as country specific factors. Our main finding in this empirical analysis is that the share of the distribution sector also follows an inverted-U pattern with respect to the level of economic development when the latter is measured in terms of purchasing power parity rates. The only other factor that we found systematically related to the share of the distribution sector in this setting is the level of media penetration in a country.

Succintly put, our empirical results show that there is a stable inverted-U pattern in the role of the distribution sector in the process of economic development, which is a substantially different result than the one found for services. And, our theoretical analysis shows the economic rationale behind this result to stem from three features. First, the increasing specialization associated with development; second, the nature of increasing returns in the distribution sector; and, third, the limits on substitution possibilities between this sector of the economy and other sectors.

1. The Economic Function of the Distribution Sector, the Development Process and Their Interaction.

The economic function of the distribution sector can be characterized as providing a set of distribution services, D, together with the usual explicit items or services that are purchased from a distributor, Q. These distribution services have been classified into five broad categories:[1] Accessibility of location, breadth and depth of assortment, assurance of product delivery at the desired time and in the desired form, information and ambiance. The provision of higher levels of each of these services entails higher costs for distributors, since each one of them can be viewed as an output.[2] Moreover, there is evidence that there are economies of scale or increasing returns in the provision of these distribution services.

Ofer (1973) argued that the value added of a store is a measure of its service output and proceeded to estimate value added production functions for the Israeli retail sector. He found substantial economies of scale in all the sectors he considered. Oi (1992) argues in favor of the existence of economies of scale in this sector along the following lines. There is empirical evidence of a positive association between store size and transaction size. By definition the total quantity bought by a consumer, Q, is equal to hq, or the number of transactions (h) times the average size of a transaction (q). If costs go up more rapidly with the number of transactions than with the average size of the transaction, this generates economies of scale as store size increases. Extending this argument, if the provision of additional

---

[1]See Betancourt and Gautschi (1988) for a detailed discussion of this classification or Oi (1992) for a detailed discussion of a similar one.

[2]In addition, each one of these services can be viewed as a fixed input into the production functions of consumers (in the case of retailers) or retailers (in the case of wholesalers). For an analysis of the implications of this interpretation see Betancourt and Gautschi(1992).

distribution services, for example broader and deeper assortments, is the reason for the larger store size, it becomes the source of scale economies. Recently, Betancourt and Malanoski (1999) found evidence of constant returns to scale with respect to output or turnover and increasing returns to scale with respect to the provision of distribution services for a sample of U.S. supermarkets.

Locay (1990), for example, shows that as development proceeds specialization increases and there is an increase in the variety of products available. Hence, the levels of assortment that have to be provided by the distribution sector increase. This view of development as increasing specialization is becoming widespread and has been incorporated in a variety of models. Thus, Rodriguez-Clare (1996)and Ciccone and Matsuyama (1996)argue in favor of development as increasing roundabout methods of production with specialized suppliers and develop models incorporating this view that can generate development traps. Bardhan (1996)adopts a similar view and shows how it can lead to persistent differences in wages between rich and poor nations.

The literature on development and specialization implies that the distribution sector has to provide increasing levels of assortment as development proceeds. The literature on the economic function of the distribution sector implies that the production of assortment is subject to increasing returns. We will proceed by first specifying a production function for assortment that is subject to increasing returns and then embedding this production function in a simple model of an economy. This will allow us to derive an explicit expression for the share of the distribution sector that can be used to show, in contrast to the case for services, how an inverted-U pattern would be expected to arise as development proceeds.

The production of assortment can be characterized as follows:

$$D = [(X_1)^á + ...+(X_n)^á]^{1/á} \qquad (1),$$

where X_i is the number of items of type i distributed by the sector. Distribution of each of these items can be thought of as produced with a fixed unit of capital and L_i units of labor. Labor will be assumed to be homogeneous so that each unit can produce B units of each type of item. á is a measure of the degree of substitution among the type of items, i.e., the higher the value the greater the substitutability, and 0 < á < 1. Restricting á to be less than unity ensures diminishing returns to the addition of quantities of any particular item. No type of item is essential to the production of assortment.

This characterization borrows the assumptions on technology normally used in the specialization literature and applies it to the distribution sector, where they seem at least as appropriate as in their original application. For any given amount of labor allocated to the distribution sector, L_D = Ó L_i, and any given number of type of items, n, it will always be optimal (in the sense of maximizing the output of assortment) to allocate the same amount of labor, R, to the production of each type of item. This feature of the model allows a rewrite of (1) in reduced form as follows:

D = $n^{(1/á)}$ BR = $n^{è}$ B L_D (2),

where è = (1-á)/á. The first equality brings out clearly the source of increasing returns in the model. If we increase B or R, given n, we have constant returns to scale, but if we increase n, given B and R, we have increasing returns to scale. The second equality is useful below, but it also brings out another aspect of the model.

With a given amount of labor allocated to the distribution sector, L_D, the impact of an increase in the number of types on the output of assortment is greater, the smaller is á or the lower the degree of substitutability among the types of items in the assortment. In discussing assortment a distinction is often made between breadth and depth. Breadth is usually viewed in terms of different product lines whereas depth is usually viewed in terms of different varieties within a product line. It is plausible to argue that the degree of substitutability (á)would be higher when adding varieties within a product line than when adding product lines, and that in the early stages of development increases in n usually represent increases in product lines whereas in the later stages of development increases in n usually represent increases in varieties within a product line. Since this argument strengthens our explanation of how the inverted-U pattern arises, it is noted here to emphasize that the reasoning which immediately follows does not rely on this feature.

Aggregate output in our economy (Y) will be given by a constant returns to scale production function of the following form:

$$Y = F(D,Q) \qquad (3),$$

where D is the output of assortment and Q is the output of other sectors. Both D and Q are essential inputs in the production of aggregate output, Y. For simplicity we will assume that the output of other sectors is produced with a simple technology that employs just labor, i.e., $Q = L_Q$.

With a fixed amount of labor, L, in the economy, perfect competition in the market for labor will lead aggregate output Y to be maximized at the point where the following condition holds:

$$F_D\, n^{è} B = F_Q = w \qquad (4),$$

where w is the economy wide wage rate. The share of distribution in aggregate output can then be written as follows:

$$wD/Y = (F_D D/F) n^{è} B \qquad (5).$$[3]

[3]Note that this economy is characterized by imperfect competition in the distribution sector due to the increasing returns; thus, labor in this sector will be paid less than the value of product generated by the sector, i.e., $(wD)/(wL_D) = n^{è}B > 1$.

The term in parenthesis is simply the elasticity of aggregate output with respect to distribution services, let us call it â. Hence, the evolution of the share of the distribution sector as development proceeds will depend on what happens to n, â, and B. The latter term (B) will be assumed constant.

If we adopt the view of development in the specialization literature, then higher levels of development, as indicated by higher levels of per capita income, also imply a higher level of specialization, as indicated by n, and equation (5) shows how the share of the distribution sector evolves as a result. First, there will be a direct effect through n that leads to an increase in the share; Second, there will be an indirect effect through â (as n increases D increases and this affects â) that leads to an eventual decrease in the share. The latter effect, which is less obvious, comes about because for any homogeneous of degree one production function the sign of Mâ/MD is determined by the elasticity of substitution between assortment and the output of other sectors, ó. That is, Mâ/MD= (ó- 1)(1-â). But in general the elasticity of substitution decreases as the ratio of assortment to the output of other sectors (D/Q) increases; hence, it must become less than unity as it approaches zero. Intuitively, as additional resources are allocated to the production of higher levels of assortment, the ability of these resources to increase aggregate output by increasing assortment diminishes relative to their ability to increase aggregate output through the expansion of other sectors.[4]

Reality is a far more complex tapestry that can be captured by this simple model, both with respect to the characteristics of the distribution sector and to the changes that take place in the course of development. Notwithstanding, the factors leading to increases in the share of distribution in the early

[4]This argument is strengthened by the point made earlier. Namely, if increases in variety at later stages of development imply increases in assortment depth rather than breadth, the value of á is bigger in the later stages of development and the direct impact of n on the share is thus smaller.

stages of development and to decreases in this share in later stages of development will continue to exist in more complex characterizations. We postpone a discussion of these complexities to Section 4 and turn now to the empirical evidence that arises by using the approach employed by Syrquin and Chenery(1989).

2. Empirical Evidence: 'A Replication'.

A convenient frame of reference for starting our analysis is the set of results in Chenery and Syrquin (CS,1975) as well as in Syrquin and Chenery (SC,1989) with respect to services. SC found an increasing pattern with respect to income and a regular but statistically weaker U pattern with respect to size for an unbalanced panel of countries over the 1950-1983 period. Initially CS found an inverted U pattern for the share of services with respect to the level of development and a regular U pattern with respect to size for an unbalanced panel of countries over the 1950-1970 period. Anomalies in the results for services in earlier studies, however, led them to argue that the results were sensitive to the high share of services in some medium income countries; finally, they concluded (p. 176) on the basis of regressions with country dummy variables for large and small countries separately that "The results (in Syrquin, 1974) indicate a rising share of services at high income levels."

Our analysis focuses exclusively on the wholesale and retail trade. Our dependent variable is the share of GDP of the distribution sector in current prices, D.[5] Our main independent variables were taken from SC (1989), i.e., real income per capita (Y) using exchange rates and population (N), which was measured in millions of persons. Net resource flow, measured as imports minus exports of goods and non-factor services, as a share of GDP (F) was calculated from the same source as our dependent variable.

---

[5]The numerator in this ratio is the value added by the distribution sector. See the Data Appendix for a complete description of variables and sources of data.

The sample period is the same as SC, 1950-1983, but the set of countries is somewhat smaller (74) due to the inability to obtain data on the dependent variable for a few countries. Thus, we have an unbalanced panel which is somewhat smaller than SC.

In the patterns spirit of SC and CS, we first estimate the following semi-log quadratic model.

$$D = \hat{a}_0 + \hat{a}_1(\ln Y) + \hat{a}_2(\ln Y)^2 + \hat{a}_3(\ln N) + \hat{a}_4(\ln N)^2 + \hat{a}_5 F + 3\ \ddot{a}_i T_i + u \qquad (6)$$

where T_i are four time dummy variables defined, following SC, as $T_1 = 1$ for year ≥ 1960; $T_2 = 1$ for year ≥ 1967; $T_3 = 1$ for year ≥ 1973, $T_4 = 1$ for year ≥ 1979.[6] Table 1 presents the results and it reproduces an equation from SC for comparison purposes. Several conclusions emerge from the first five rows of Table 1. First, there is an inverted-U pattern between per capita income and the share of distribution in GDP. This pattern holds whether or not one controls for size or net resource flows and with or without the SC time dummies. Second, there is no discernible pattern with respect to size. Third, net resource flows increase the share of distribution in GDP and this effect is not sensitive to the inclusion or exclusion of size or the SC time dummies. Finally, the SC time dummies are "statistically insignificant" individually, which can be seen from the table, and collectively.[7] Since these SC time dummies are also defined on a somewhat subjective basis, they are dropped from all subsequent analyses of the distribution sector.

In the next three rows of Table 1 we present the results of estimating three fixed effects equations for the distribution sector. The first one (6) introduces country dummies, the second one (7) introduces

[6] This equation replicates the basic equation in Syrquin and Chenery (1989, p.11) with respect to the choice of explanatory variables and functional form.

[7] An F-test of restrictions between equations 1 and 5 leads to a value of the test statistic of 1.22, which is well below the critical value for rejection of the null hypothesis that the coefficients of the dummies are zero at all standard significance levels.

standard time dummies and the third one (8) introduces both into equation (1). Two results emerge from these equations and their comparison with equation (1). First, the time dummies are largely irrelevant to the analysis. F-tests accept the null hypothesis that the dummies in (7) do not add to the explanation in (1) or that the dummies in (8) do not add to the explanation in (6) at conventional significance levels. Even the magnitudes of the coefficients on income remain the same in (7) as they were in (1)! Second, and most important, the country dummies (6) have a substantial effect on the results both economically and statistically. Explanatory power increases by a factor of 10 as a result of introducing the country dummies.[8] While an inverted-U pattern between distribution and the level of development seems to emerge from these results, it is different from the previous one in that the share of distribution rises steeply and peaks late. The difference can best be seen in Figure 1, where we plot the share of distribution against the log of per capita GNP for both equations (1) and (6). The turning point occurs at a higher level of development when country specific factors are controlled for in the analysis.[9]

The last two rows of Table 1 present the results of two comparable equations: the first one based on our data for the distribution sector and an identical one from the analysis in SC (1989, p.107) for the services sector. Equation (9) is the same as (10) except the dependent variable in (10) is the share of <u>services</u> in GDP in current prices and (10) is based on the complete set of observations available to SC. In (10) there is no inverted-U pattern with respect to services in general and size and the SC time dummies do matter. In (9) there is an inverted U pattern for the distribution sector, the SC dummies do not matter

[8] An F-test of the null hypothesis that the country dummies are zero generates a value of 82.05 for the test statistic.

[9] Incidentally, the analysis in the previous section predicts an inverted-U but it has no specific prediction for the level of per capita income at which the turning point occurs.

by and large, and size does not matter.

What is the interpretation of these comparisons? Simply put, the patterns for the distribution sector are different than the patterns for the services sector. In particular, the share of distribution does not rise monotonically with the level of per capita income converted using exchange rates. It increases at lower levels of development but eventually declines.

3. Sensitivity Analysis.[10]

One issue that arises with respect to the previous results is that they are based on an unbalanced panel. That is, the number of time series observations for every country varies across countries. Thus, we want to ascertain the sensitivity of the results to this characteristic of the data. It is also useful to know the sensitivity of the results to the addition or subtraction of years or countries to the data. Hence, we constructed three different panels that vary with respect to country coverage and yearly coverage. Table 2 presents the results of estimating equations (1), (7) and (6) of Table 1 for these three different panels. The first one is a balanced panel of countries selected to include countries with the longest time series. Indeed, it coincides with the original time period in Section 1, but it contains only 12 countries. The second one is a balanced panel that has the maximum number of countries (33) for which we could find reasonably long time series, which turned out to be 19 years. The last one is a balanced panel constructed so as to maximize the number of observations over both time and countries for which we could have reasonably long time series for every country. It covers 27 countries over the period 58-83.

In all three cases equation (6) outperforms equations (1) and (7) in terms of predictive

[10]We also tested the sensitivity of our results to variations in data collection procedures. These results are reported in the Data Appendix.

performance. The explanatory power of (6) is at least twice the explanatory power of (1) or (7) in all three cases. Thus, the importance of country specific factors in explaining the share of distribution in GDP is robust to the use of balanced panels and to changes in country and yearly coverage.[11] Net resource flow is also robust. The variable comes in positive and statistically significant in all three cases. Finally, the impact of GNP per capita on the share of distribution is also robust in terms of signs and statistical significance and overall pattern. For each of these three alternative panels, we present in Figure 2 the pattern between the share of the distribution sector in GDP and the level of GNP per capita that emerges from equation (6). They are similar to what we found in Section 1 for equation 6.[12]

Next we considered the sensitivity of our results to the specification of the functional form of our independent variables and of our dependent variable. Our choice of a quadratic in the logs followed Syrquin and Chenery but there is no theoretical rationale for this choice. Hence, we reestimated the model in Section 1 with a straight quadratic (straight Y) in income and size and with an inverse quadratic (inverse Y) in income and size.[13] It is also well known that when the value of the dependent variable lies between 0 and 1, the disturbances can not be normally distributed. A common correction for this is to use a logistic transformation of the dependent variable.

---

[11]Incidentally, the relationship between the share of distribution and the level of development as estimated in (1) or (7) is not robust to the use of the balanced panels. Hence, these two relationships will not be discussed further.

[12]The only result from (6) that is affected by the use of these alternative panels is that size shows a regular U pattern in the first balanced panel in Table 2, but not elsewhere.

[13]Anand and Kanbur(1993) have shown that the famous Kuznets inverted-U relationship between inequality and development disappears when a straight quadratic rather than a quadratic in the logs is applied to the same data.

Table 3 presents the results of estimating these alternative possibilities. The first of the six panels reproduces equations (1) and (6) from Table 1 for ease of comparison. First, the specification with country dummies is also the relevant one with the two new functional forms and with both specifications of the dependent variables. Similarly, net resource flow has the same effect economically and statistically as in Section 1. Once country specific factors are allowed for, the logarithmic specification and the quadratic specification generate an inverted U pattern with respect to per capita income for both specifications of the dependent variable.[14] Hence, the results in Section 2 are robust to these alternative specifications of functional form for the independent and dependent variables.

For completeness, we performed a J-test on these non-nested alternatives, Greene (1993). The results are presented in Table 4. We will limit the discussion to the model that allows for country specific factors. The inverse Y specification is dominated by the logarithmic Y specification whether the dependent variable is the share of the distribution sector or the logistic of this share. The inverse Y specification is neither dominated nor does it dominate the straight Y specification whether the dependent variable is the share or the logistic of the share. Interestingly, the logarithmic Y specification neither dominates nor is it dominated by the straight Y specification.

Summing up, three robust empirical regularities emerge from our sensitivity analysis. Country specific factors are important in explaining the share of distribution in GDP; across a wide variety of countries and time periods the average time series share of the distribution sector increases initially with the level of development but eventually declines; and, the greater the excess of imports over exports relative

---

[14]While this is not true for the inverse quadratic, we will see below that a nonnested test rejects this functional form specification against the logarithmic one.

to GDP the greater this share. The last result is induced almost mechanically by the nature of the distribution activity and how it shows up in the national income accounts. Thus, it is discussed further here.

Economic agents have several ways of making their products directly available to consumers in foreign countries. For instance, three important ones are: They can develop their own foreign distribution network; they can use importers in the foreign country that are merchant wholesalers; and, they can use exporters in their own country that are merchant wholesalers. Only if they use the last mechanism will their activity be counted as part of the distribution sector in the exporting country. When exports increase the activity of the distribution sector will increase only in proportion to the fraction of the products that are exported through merchant wholesalers that are exporters.[15] By contrast an increase in imports will lead to an increase in the activity of the distribution sector in the importing country through the first two of the three channels mentioned above as well as through others, for example sales of foreign manufacturers directly to retailers in the importing country lead to an increase in the level of the distribution sector in the importing country. This inherent characteristic of the distribution sector, we believe, is the main reason for the positive and 'statistically significant' coefficient of F in these average time series regressions. A belief that is supported by the empirical analysis in the next section.

4. Empirical Evidence: Cross-Section Data.

Our empirical results demonstrate the importance of country specific factors in determining the share of distribution across countries. This raises the question of whether or not there are any empirical regularities that explain these country specific factors. One possibility that must be considered is the level

[15] In the U.S. in 1987, the sales of importers who were merchant wholesalers were close to $225 billions of 1987 $. The sales of exporters who were merchant wholesalers, by contrast, were $ 80 billions of 1987 $. Both numbers come from the 1987 Census of Wholesale Trade (WC87-S-4, Table 7).

of development itself. The only variation in income captured in our previous empirical results for the preferred specification (that is, the models that include the country dummies) was variation within a country over time. Nonetheless since the same factors that determine the evolution of the share in equation (5) of Section 1 will be at work across countries, it is interesting to see if the relationship they generate holds in a cross-country setting despite the many other factors that exhibit substantial variation in this setting. Furthermore, the cross-section data will provide evidence against the possibility that the panel data results are 'spurious'.

One of the factors that can vary substantially across countries and that may affect the distribution sector is the level of information available in the society at large. In order to have a demand for variety, consumers must be aware of the existence of this variety. Nonetheless, other sectors will also be affected and the net effect on the share of distribution must be determined empirically. Similarly, the level of urbanization differs substantially among countries. Higher levels of urbanization would have two opposite effects on the distribution sector: since they are usually accompanied by an increase in the number of households, this increases the demand for the services of the distribution sector; since they generate agglomeration economies in the provision of these services, they may allow the satisfaction of this demand without an increase in resources relative to other sectors of the economy.

An important characteristic of the distribution services mentioned at the beginning of Section 1 is that, except for ambiance, they are in general substitutes for the household's use of time in purchase activities; thus, variations in the opportunity cost of time across countries can generate variations in the demand for the services of the distribution sector. On the other hand, one of the main results of the purchasing power literature is that there are substantial variations in the costs of nontradables across countries. The distribution sector is a nontradable and these substantial variations in costs can generate substantial variations in the supply of services by this sector. Both of these factors will also affect other sectors of the economy, however, and their net effect on the share of distribution may be ambiguous.

Finally, institutions can vary substantially among countries and they seem to matter in terms of economic performance, for example North(1990) and Barro and Sala-i-Martin (1995, Ch.12). Presumably, the greater the security of transactions provided by the institutional environment, the greater the number of transactions between economic agents and the distribution sector that will take place; of course, this greater security will also affect the level of activity in other sectors so that the net effect may be ambiguous.

To explore these issues, we have taken for each country the average share of the distribution sector over the years for which it was available in the data used earlier. We also constructed similar averages for the variables used in our previous empirical analysis. This information is supplemented with data from other sources that can be used to measure the factors discussed above. The result is a set of observations on 47 countries.

Since the dependent variable is the average share of the distribution sector in GDP for each country, it is a proportion and it also varies in reliability depending on the number of time series observations used to construct the average. One method of estimation that takes account of the second problem is weighted regression, using the square root of the number of observations on which the average share for each country is calculated as weights. An appropriate method of estimation that takes account of both of these features is a two step logit estimator, e.g., Greene (1993,pp.653-655).

In the first stage, one obtains consistent but inefficient estimates of the shares by applying OLS to the following logit model

$$\ln [D_i/(1-D_i)] = \hat{a}'X_i + u_i, \qquad (7)$$

where D_i is the observed average share for country i, $\hat{a}'$ is a vector of coefficients, X_i is a vector of corresponding independent variables and the disturbances in (7) are heteroskedastic. These estimates are used to construct weights for each observation as follows:

$$w_i = (n_i d_i(1-d_i))^{1/2}, \qquad (8)$$

where n_i is the number of time series observations used to calculate the average share in country i, and d_i is the estimated share for country i from the first stage regression. These weights are used to estimate the following weighted regression in the second step,

$$\ln[D_i/(1-D_i)]w_i = \hat{a}'X_i w_i + \mathring{a}_i . \qquad (9)$$

The disturbances in equation (9) can now be assumed to be homoskedastic.

Our most important explanatory variable is, of course, the level of development (ICP). It was measured by an index of the level of GDP per capita in 1980 ICP $.[16] In light of our earlier results using GNP per capita converted at exchange rates and of the Kravis, Heston and Summers (1983) result on services cited in the introduction, we also estimated every specification with GNP per capita (AVGY) as an alternative measure of the level of development. Our analyses in Sections 2 and 3 suggested the introduction of average size (AVGN) and average net resource flow (AVGF) as explanatory variables.

The discussion at the beginning of this section suggested the inclusion of the following variables: The availability of information measured by an index of media penetration (MEDIA), which was constructed as a simple average of indexes of the number of radio's, television sets and newspapers per 100 persons

[16] Sources and detailed definitions of all the variables are given in the Data Appendix.

in a country; Urbanization (URBAN) measured by the percentage of households in a country that resided in urban areas in 1965; The opportunity cost of time (WOMEN) measured by the rate of labor force participation of women in 1968;[17] The relative price of nontradables (ERDI) measured in inverse form as the ratio of GDP in 1980 ICP $ to GDP in 1980$ converted using exchange rates; Finally, the security of transactions (RULEOFLAW) measured as an index of the extent to which the rule of law prevails in a country.

Table 5 presents the results of three different specifications estimated by weighted least squares (WLS) and by the two step logit estimator (TSLOGIT). The first equation in each panel includes only the logarithm of income and its square as explanatory variables. The second equation in each panel includes, in addition, all the variables listed above without the slope dummy for labor force participation of women in Africa; finally, the third equation in each panel adds this last variable to the second equation. For completeness, the third panel of the table presents the results of estimating the same equations by WLS using as dependent variables the estimated coefficients of the dummies for each country from equation 6 in Table 1. The weights in this case are the standard errors of the estimated coefficients.

An inverted -U pattern with respect to the level of development appears in every specification for all three panels. By itself the level of development explains between 20 % and 38% of the variation in the (transformed)dependent variable and both coefficients are statistically significant at the 1% level. Our search for empirical regularities yields two other results. First, media has a positive and statistically significant impact on the share of the distribution sector. This is true at the 5% level or lower in 5 of the six regressions

[17] To account for the high rate of labor force participation of women in several African countries (higher than in most developed countries) we introduced this variable by itself and with a slope dummy for African countries (WOMENAFRICA).

and at the 10 % level in the remaining one. Second, no other variable is statistically significant at the 10% level. In particular, average net resource flow plays no role in determining the share of distribution in GDP in this setting, which provides empirical support for the mechanical interpretation of the time series result given at the end of Section 3.

Just as in Section 3, we considered two alternative specifications of the functional form of the level of development: namely, a quadratic in ICP and an inverse quadratic in ICP. In this case we also considered the use of the log of GNP per capita (lnAVGY) instead of our index of GDP per capita at purchasing power parity rates for the logarithmic specification. Table 6 presents the results of a J-test of the semilog functional form in Table 5 against each of these three alternative specifications equation by equation. The lnAVGY specification is dominated at the 5% level or above in each of the nine possible cases. The ln (ICP) specification dominates the straight ICP specification at the 5% level or above in 5 of 9 possible cases and the results are inconclusive in the other 4 cases. Finally, the ln ICP specification dominates the inverse ICP specification in every case at the 10 % level or above and in 7 of 9 cases at the 5% level or above shown in the table.

In Table A2 of the Data Appendix we present results comparable to Table 5 for the straight ICP quadratic specification. By themselves, this variable and its square explain less than 5 % of the variation in the share of the distribution sector and neither one is statistically significant at the 5% level. These results improve somewhat statistically when other variables are added, but the 'statistically significant' results for media and women are not stable between the second and third specification in each panel. The results in Table A2 indicate a lack of statistical robustness and our J-tests in Table 6 show that this functional form is statistically inferior to the logarithmic form on which Table 5 is based. Thus, these results support the use of the semilog functional form as well as Anand and Kanbur's (1993) advocacy of the need to consider alternative functional forms and to perform statistical tests to choose among these forms in situations where the theory provides little or no guidance as to functional form.

5. Concluding Remarks.

Economists and policy makers have been equally adept at neglecting or ignoring the distribution sector. Indeed, in the centrally planned economies this sector was systematically suppressed and one characteristic of these economies transition to a market system has been a rapid expansion of the distribution sector. Our analysis shows that this sector warrants greater attention on three grounds. First, we have the sheer economic importance of the sector due to size, since the average share of this sector over our 74 countries during the 34 year period of our data is 13.5%. Second, we have the systematic and robust pattern exhibited by the share of this sector over time and across countries: namely, it first rises and subsequently declines as per capita income increases. Third, this pattern is not merely an statistical artifact but seems to be grounded on fundamental economic forces characterizing both the development process and the functioning of the distribution sector.

When economies grow more complex and produce and consume a greater variety and quantity of items, the distribution sector has to perform more work and provide higher levels of distribution services in general and of assortment in particular. Economies of scale in the provision of these services lead to an increase in the share of this sector in GDP in the early stages of development, but this process contains the seeds of its own demise because it becomes more difficult to substitute between the services provided by this sector and the output of the rest of the economy.

Data Appendix

We begin by providing definitions and sources for each variable used in the analysis. Subsequently we describe some checks devised to account for variations in data collection procedures. The data for Sections 2 and 3 came from two sources: The Syrquin and Chenery (SC) data set, which was kindly provided by the World Bank, and the United Nations series (UN)of yearbooks "National Accounts Statistics". Since we could not find data for the distribution sector in the UN sources for every country in the SC data set, our sample of countries was reduced to 74 countries. In what follows we describe each variable, its construction when applicable, and its source.

D is the share of distribution in GDP. Both numerator and denominator were compiled from the same UN sources. The numerator is value added by wholesale and retail trade in current local currency. The denominator is GDP also measured in current local currency. The main difficulty in constructing this variable was that some countries include hotels and restaurants as part of the distribution sector and others don't. Fortunately, the UN sources indicate whether or not hotels and restaurants are included and in most cases also provide separate information on this item. Hence, we subtracted this information from the aggregate whenever it was included and eliminated the country-year from the analysis if the aggregate contains hotels and restaurants and we could not disaggregate. Thus, our measure does not include hotels and restaurants as part of the distribution sector. Table A0 provides descriptive statistics on this variable and its use for every country. F is net resource flow measured as imports minus exports divided by GDP, all three in local currency. Both numerator and denominator were taken from the same UN sources.

Y was measured as the level of GNP per capita in 1980 dollars and it was taken directly from the SC data set. Similarly, N was measured as the number of millions of persons in the country and it was also taken directly from the SC data set. Both AVGY and AVGN in Section 4 are simply the average over the available time series for each country of Y and N, respectively. The same is true, of course, of the average share of the distribution sector and of AVGF in Section 4.

The remaining variables were used only in Section 4. Three of them come from World Development Reports (WDR). The level of development in purchasing power parity terms (ICP) was measured as the ICP 1980 index of per capita income, WDR (1987). It ranges from 2.4 to 101.5 and the USA = 100. Urban was measured as the percentage of households living in urban areas in 1965, WDR (1987). It ranges from 5 to 93. Finally, ERDI was measured as the ratio of GDP in 1980 ICP $ (WDR, 1987) to GDP in 1980 $ converted at official exchange rates using the World Bank Atlas method (WDR, 1982).

Our last three variables are women, rule of law and media. Women was measured as the percentage of the female labor force in the total labor force in 1968. It ranges from 8.9 to 50.7 and it was taken from the 1994 World Tables. Rule of law was measured as an index of the extent to which the rule of law prevails in a country. It is measured on a scale of 0 to 6, which ranges from 1 to 6 in our sample. This variable is part of the IRIS data base and it was developed from the International Country Risk Guide, International Reports, March 1992, IBC USA Publications, INC. Incidentally, this is the variable used by Barro and Sala-i-Martin (1995, Ch.12) to capture institutional factors in their cross-country regressions. Finally, media was constructed as an average of indexes of the number of radios , tv sets and newspapers per 100 persons in the country. Each individual index was constructed so that the maximum value of the index in the data was unity. Media ranges from 0.008 to 0.802. All of the data to construct the indexes was taken from the Human Development Report, 1994.

Since the data used to construct the dependent variable correspond to a 34 year time series in which there were changes in national income accounting procedures, we performed some robustness tests of the potential impact of these changes in data collection procedures. One issue is that the UN sources calculated national income aggregates using factor costs in some years and producer's prices in others. A dummy variable was constructed to indicate which method was employed. Another issue is that the UN system of national accounts was changed in the middle of this period. Another dummy variable was constructed to capture whether the data came from the former or the present system of national accounts. Finally, the UN data had to be compiled from 18 different yearbooks. Hence, we defined 17 dummy variables (one for all but one of the yearbooks) to capture any unreported changes from one yearbook to another.

We reestimated our models introducing these dummy variables to see if they would affect our substantive conclusions. The results for equations 6 and 8 of Table 1 with and without these 19 dummy variables are presented in Table A1. Our substantive conclusions are not affected by the introduction of these 'robustness' dummies.

References

Anand, S. and R. Kanbur, 1993, "Inequality and Development : A Critique," Journal of Development Economics 41, 19-43.

Bardhan, P., 1996, "Disparity in Wages but not in Returns to Capital between Rich and Poor Countries," Journal of Development Economics, 49, 257- 270.

Barro, R., and X. Sala-i-Martin, 1995, Economic Growth, (Mc-Graw-Hill, Inc., New York).

Betancourt, R., 1993, An Analysis of the U.S. Distribution System, Discussion paper No. 135 (Organization for Economic Cooperation and Development, Paris).

Betancourt, R. and D. Gautschi, 1988, "The Economics of Retail Firms," Managerial and Decision Economics 9, 133-144.

Betancourt, R. and D. Gautschi, 1992, "The Demand for Retail products and the Household Production Model: New Views on Complementarity and Substitutability,"Journal of Economic Behavior and Organization 17, 257-275.

Betancourt, R. and M. Malanoski, 1999, " An Estimable Model of Supermarket Behavior: Prices,Distribution Services and Some Effects of Competition,"forthcoming Empirica.

Bucklin, L., 1973, Productivity in Marketing (American Marketing Association, Chicago).

Chenery, H. and M. Syrquin, 1975, Patterns of Development: 1950-1970 (Oxford University Press, London).

Ciccone, A., and K. Matsuyama, 1996, "Start-up Costs and Pecuniary Externalities as Barriers to Economic Development," Journal of Development Economics, 49, 33-59.

Greene, W., 1993, Econometric Analysis, Second Edition (Macmillan Publishing Co., New York).

Kravis, I., A. Heston and R. Summers, 1983, "The Share of Services in Economic Growth," in F.G. Adams and B. Hickman, eds., Global Econometrics: Essays in Honor of Lawrence R. Klein (MIT Press, Cambridge).

Kuznets, S., 1966, Modern Economic Growth: Rate, Structure and Spread (Yale University Press, New Haven).

Locay, L., 1990, "Economic Development and the Division of Production between Households and

Markets ," Journal of Political Economy, 98, 955-982.

North, D., 1990, Institutions Institutional Change and Economic Performance (Cambridge University Press, London).

Ofer, G., 1973, "Returns to Scale in Retail Trade," Review of Income and Wealth 19, 363-384.

Oi, W.,1992, "Productivity in the Distributive Trades: The Shopper and the Economies of Massed Reserves," in Z. Griliches, ed., Output Measurement in the Service Sector (University of Chicago Press, Chicago).

Panagariya, A., 1988, "A Theoretical Explanation of Some stylized Facts of Economic Growth," Quarterly Journal of Economics 102, 509-526.

Ray, Debraj, 1998, Development Economics (Princeton University Press: Priceton, NJ).

Rodriguez-Clare, A., 1996, "The Division of Labor and Economic Development," Journal of Development Economics, 49, 3-32.

Romer, P., 1987, "Growth Based on Increasing Returns due to Specialization," American Economic Review, 77, 56-72.

Syrquin, M. and H. Chenery, 1989, Patterns of Development:1950-1985, Discussion paper No. 41(World Bank, Washington, DC).

United Nations, 1994, Human Development Report (Oxford University Press, New York).

United Nations, National Accounts Statistics, various issues, (United Nations Publishing Division, New York).

U.S. Bureau of the Census, 1987, Census of Wholesale Trade, WC87-S-4(Washington, DC)

World Bank,1982 and1987, World Development Report (Oxford University Press, London).

World Bank, 1994, World Tables (The Johns Hopkins University Press, Baltimore).

Table 1. The Distribution Sector and the Development Process: Initial Results[a]

Equation	ln Y	(ln Y)^2	ln N	(ln N)^2	F	T_1	T_2	T_3	T_4	R^2	N
D (1)	0.023	-0.20E-02	-0.43E-03	0.33E-03	0.087					.080	1609
	(2.40)	(3.08)	(0.14)	(0.62)	(6.41)						
D (2)	0.021	-0.20E-02								.056	1609
	(2.32)	(3.17)									
D (3)	0.024	-0.22E-02	-0.11E-02	0.38E-03						.057	1609
	(2.51)	(3.32)	(0.36)	(0.71)							
D (4)	0.019	-0.17E-02			0.084					.079	1609
	(2.08)	(2.79)			(6.31)						
D (5)	0.024	-0.21E-02	-0.48E-03	0.31E-03	0.085	0.57E-02	-0.18E-02	0.51E-02	-0.22E-02	.085	1609
	(2.52)	(3.21)	(0.15)	(0.58)	(6.19)	(1.63)	(0.58)	(1.61)	(0.61)		
D (6)[b]	0.094	-0.55E-02	0.71E-02	-0.13E-02	0.074					.815	1609
	(5.81)	(5.62)	(1.01)	(0.94)	(7.07)						
D (7)[c]	0.023	-0.20E-02	-0.48E-03	0.30E-03	0.084					.088	1609
	(2.46)	(3.15)	(0.15)	(0.56)	(6.08)						
D (8)[d]	0.119	-0.78E-02	-0.019	-0.16E-02	0.070					.819	1609
	(6.82)	(6.77)	(1.86)	(1.21)	(6.54)						
D (9)	0.025	-0.23E-02	-0.13E-02	0.37E-03		0.48E-02	-0.16E-02	0.62E-02	-0.36E-03	.063	1609
	(2.68)	(3.50)	(0.41)	(0.68)		(1.36)	(0.50)	(1.95)	(0.10)		
S (10)[e]	0.02	0.002	-0.019	0.001		-0.016	0.009	-0.006	0.013	.407	2311
	(1.49)	(2.03)	(4.77)	(0.98)		(2.20)	(2.03)	(1.41)	(2.67)		

a. D stands for the share of distribution in GDP as the dependent variable; S stands for the share of services in GDP as the dependent variable; constants were included in all regression but are not displayed here; t-ratios (absolute value) are in parentheses.
b. This equation is (1) with country dummies (not shown here) added to the right hand side.
c. This equation is (1) with time dummies (not shown here) added to the right hand side.
d. This equation is (1) with country and time dummies (not shown here) added to the right hand side.
e. This equation is taken from Syrquin and Chenery (1989, p. 102).

Table 2. Balanced Panel Regressions

Years	Countries	Country Dummies	Time Dummies	ln Y	(ln Y)^2	ln N	(ln N)^2	F	R^2	N
50-83	12	no	no	-0.019 (1.48)	0.12E-02 (1.45)	-0.042 (7.79)	0.74E-02 (6.84)	0.130 (6.02)	.217	408
50-83	12	no	yes	-0.26E-02 (0.20)	0.29E-04 (0.03)	-0.044 (8.35)	0.77E-02 (7.28)	0.111 (5.00)	.310	408
50-83	12	yes	no	0.210 (7.40)	-0.012 (7.51)	-0.032 (2.93)	0.011 (3.99)	0.057 (3.18)	.631	408
60-78	33	no	no	-0.80E-02 (0.50)	0.22E-05 (0.00)	-0.015 (3.02)	0.23E-02 (2.42)	0.069 (2.23)	.133	627
60-78	33	no	yes	-0.68E-02 (0.41)	-0.12E-03 (0.11)	-0.015 (2.98)	0.22E-02 (2.34)	0.061 (1.92)	.146	627
60-78	33	yes	no	0.265 (8.94)	-0.016 (8.94)	-0.026 (1.87)	-0.35E-02 (1.34)	0.103 (4.67)	.809	627
58-83	27	no	no	0.013 (0.89)	-0.145 (1.45)	-0.018 (3.71)	0.30E-02 (3.23)	0.031 (1.02)	.158	702
58-83	27	no	yes	0.025 (1.68)	-0.23E-02 (2.36)	-0.017 (3.60)	0.27E-02 (2.95)	-0.025 (0.79)	.207	702
58-83	27	yes	no	0.147 (6.31)	-0.91E-02 (6.67)	-0.45E-02 (0.44)	0.39E-02 (1.75)	0.118 (5.40)	.757	702

Table 3. Alternative Functional Forms

Dependent Variable	Country Dummies	ln Y	(ln Y)^2	ln N	(ln N)^2	F	R^2	N
D	no	0.023 (2.40)	-0.20E-02 (3.08)	-0.43E-03 (0.14)	0.33E-03 (0.63)	0.087 (6.41)	.080	1609
D	yes	0.094 (5.81)	-0.55E-02 (5.62)	0.71E-02 (1.01)	-0.13E-02 (0.94)	0.074 (7.07)	.815	1609
		Y	Y^2	N	N^2	F	R^2	N
D	no	-0.52E-05 (5.02)	0.23E-09 (2.57)	0.21E-03 (5.10)	-0.41E-06 (5.94)	0.085 (6.39)	.100	1609
D	yes	0.87E-05 (6.06)	-0.60E-09 (6.98)	0.41E-03 (3.25)	-0.27E-06 (2.27)	0.078 (7.58)	.815	1609
		1/Y	(1/Y)^2	1/N	(1/N)^2	F	R^2	N
D	no	0.63E-02 (3.90)	-0.46E-03 (2.31)	0.32E-04 (0.20)	0.58E-06 (0.27)	0.094 (6.92)	.059	1609
D	yes	-0.015 (5.50)	0.73E-03 (3.53)	0.22E-05 (0.01)	-0.23E-05 (0.93)	0.075 (7.18)	.813	1609
		ln Y	(ln Y)^2	ln N	(ln N)^2	F	R^2	N
$\log_{10}(D/(1-D))$	no	0.144 (4.27)	-0.011 (4.79)	-0.23E-02 (0.21)	0.19E-02 (1.01)	0.497 (10.27)	.113	1609
$\log_{10}(D/(1-D))$	yes	0.421 (6.77)	-0.025 (6.58)	-0.39E-02 (0.14)	-0.31E-02 (0.61)	0.337 (8.40)	.795	1609
		Y	Y^2	N	N^2	F	R^2	N
$\log_{10}(D/(1-D))$	no	-0.15E-04 (4.00)	0.59E-09 (1.87)	0.84E-03 (5.76)	-0.16E-05 (6.42)	0.491 (10.23)	.128	1609
$\log_{10}(D/(1-D))$	yes	0.32E-04 (5.83)	-0.23E-08 (6.89)	0.15E-02 (3.10)	-0.89E-06 (1.97)	0.344 (8.72)	.795	1609
		1/Y	(1/Y)^2	1/N	(1/N)^2	F	R^2	N
$\log_{10}(D/(1-D))$	no	0.023 (3.93)	-0.24E-02 (3.44)	-0.54E-03 (0.95)	0.91E-05 (1.19)	0.523 (10.67)	.091	1609
$\log_{10}(D/(1-D))$	yes	-0.060 (5.79)	0.28E-02 (3.48)	0.13E-02 (1.01)	-0.15E-04 (1.59)	0.341 (8.44)	.793	1609

Table 4. Pairwise J-Tests of Functional Form for Level of Development

Dependent Variable	Country Dummies Included	Top Form 1%	$\hat{b}_{Y:LY}$	$\hat{b}_{LY:Y}$	$\hat{b}_{Y:IY}$	$\hat{b}_{IY:Y}$	$\hat{b}_{IY:LY}$	$\hat{b}_{LY:IY}$
D	Y versus LnY	*none*	4.49	4.63				
D	Y versus InvY	*none*			4.43	5.69		
D	InvY versus LnY	*LnY*					4.74	0.86
$\log_{10}(D/(1-D))$	Y versus LnY	*none*	4.50	4.54				
$\log_{10}(D/(1-D))$	Y versus InvY	*none*			4.44	5.74		
$\log_{10}(D/(1-D))$	InvY versus LnY	*LnY*					4.63	0.84

Double border signifies t-statistic significant at the 1% level or above.

Table A0. The Average Share of the Distribution Sector for Individual Countries

	Average D_i	Max/Min D_i	Years	Balanced Panels †	Cross-Section ‡
Argentina	14.7%	12.3/17.0	60 - 83		yes
Australia	15.2%	13.9/16.4	50 - 83	1,2	
Austria	12.4%	8.2/15.4	50 - 83 *	1,2,3	yes
Belgium	11.6%	6.7/14.1	51 - 83 *	1,2	yes
Benin	22.0%	18.0/24.7	74 - 83		
Bangladesh	9.2%	8.0/10.8	72 - 83 *		
Bolivia	15.7%	11.9/18.9	66 - 79 *		yes
Brazil	12.5%	10.1/14.3	60 - 80 *	1	yes
Burma	26.7%	23.7/31.6	50 - 83		
Canada	11.8%	9.6/14.2	50 - 83		yes
Chile	19.5%	14.1/25.2	60 - 82	1	yes
Ivory Coast	15.4%	12.9/21.1	60 - 74		yes
United Republic of Cameron	11.5%	8.9/16.0	74 - 83 *		yes
Colombia	14.2%	9.4/18.5	50 - 83	1,2	yes
Costa Rica	15.7%	11.5/19.4	60 - 77 *		
Germany, Federal Republic of	11.8%	9.2/13.7	50 - 83 *	1,2,3	yes
Denmark	13.7%	10.9/16.0	50 - 83 *	1,2,3	yes
Dominican Republic	17.4%	15.7/20.4	53 - 83	1,2	yes
Algeria	14.6%	9.2/19.8	57 - 78		
Ecuador	12.0%	10.3/13.9	50 - 83 *	1,2,3	yes
Spain	11.6%	9.2/14.3	54 - 83 *	1,2	yes
Ethiopia	8.2%	5.9/9.8	61 - 83 *		yes
Finland	9.4%	8.2/10.6	50 - 83 *	1,2,3	yes
France	11.7%	9.7/13.3	50 - 83	1,2	yes
United Kingdom	11.0%	8.6/14.9	50 - 80 *	1	yes
Ghana	14.8%	11.0/26.2	68 - 83 *		
Greece	11.1%	9.4/12.7	50 - 83 *	1,2,3	yes
Honduras	12.6%	9.7/15.1	50 - 83 *	1,2,3	yes
Upper Volta	13.3%	12.0/14.2	65 - 83		
Indonesia	16.5%	12.2/25.0	60 - 83 *	1	yes
India	10.2%	8.8/11.6	66 - 83 *		yes
Iran	7.9%	5.1/10.9	59 - 78 *	1	
Ireland	10.0%	9.2/10.9	70 - 83 *		yes
Israel	11.4%	7.7/14.0	66 - 80 *		yes
Italy	11.0%	9.0/12.9	50 - 83 *	1,2,3	yes
Jamaica	17.0%	12.1/21.2	50 - 83 *	1,2,3	
Japan	15.5%	13.3/17.1	50 - 83	1,2	yes
Kenya	8.1%	6.5/9.8	64 - 82 *		yes
Korea, Republic of	12.9%	9.0/16.2	53 - 83 *	1,2	yes
Libyan Arab Republic	4.9%	4.3/5.7	71 - 80		
SriLanka	13.1%	7.4/19.6	50 - 83 *	1,2,3	yes
Madagascar	13.1%	10.8/14.3	66 - 74 *		yes
Mexico	29.6%	27.8/31.9	65 - 77 *		
Malawi	9.6%	7.2/12.1	60 - 71 *		yes
Malaysia	14.4%	9.8/16.2	60 - 78		

Table A0. The Average Share of the Distribution Sector for Individual Countries

	Average D_i	Max/Min D_i	Years	Balanced Panels †	Cross-Section ‡
Niger	13.9%	12.4/14.7	63 - 69		
Nigeria	16.0%	10.6/21.5	53 - 83		yes
Nicaragua	21.6%	19.9/24.5	63 - 78 *		
Netherlands	11.9%	10.9/13.5	50 - 83	1,2	yes
Norway	12.7%	10.0/18.5	50 - 83 *	1,2,3	yes
Pakistan	15.5%	13.5/18.6	70 - 83 *		yes
Panama (including Canal Zone)	12.7%	8.8/16.3	50 - 83	1,2	
Peru	14.0%	10.8/16.3	51 - 83		yes
Philippines	13.0%	8.0/17.2	58 - 83 *	1,2	yes
Papua New Guinea	7.7%	5.7/10.2	65 - 83		
Portugal	12.5%	7.1/19.4	52 - 83 *	1,2	yes
Saudi Arabia	4.8%	2.2/6.9	63 - 76 *		
Sudan	18.0%	13.5/21.7	66 - 83		
El Salvador	23.0%	20.6/24.8	58 - 83 *	1,2	yes
Sweden	10.1%	9.1/11.6	58 - 83 *	1,2	
Syrian Arab Republic	19.6%	17.3/25.5	66 - 77 *		
Taiwan, Province of	15.2%	12.4/17.8	51 - 67 *		
Togo	20.3%	17.0/24.0	63 - 72		
Thailand	17.9%	15.0/19.9	53 - 83 *	1,2	
Tunisia	14.0%	9.9/18.9	61 - 83		yes
Turkey	11.3%	7.4/16.8	50 - 83 *	1,2,3	
United Republic of Tanzania	10.8%	9.6/11.4	66 - 76 *		yes
Uganda	9.6%	8.0/10.7	68 - 71 *		
Uruguay	13.5%	10.7/17.2	58 - 74 *		yes
USA (including Puerto Rico)	16.5%	15.9/17.1	53 - 73		yes
Venezuela	8.7%	6.8/10.8	70 - 83 *		yes
Zaire	14.3%	11.2/22.0	68 - 77 *		
Zambia	9.7%	7.5/13.0	64 - 83 *		yes
Zimbabwe	12.6%	10.1/14.9	60 - 83 *		yes

* An asterisk indicates that the years were continuous.
† Indicates in which balanced panels the country was included: 1=1960-1978; 2=1958-1983; 3=1950-1983.
‡ Indicates whether the country was included in the cross-section regressions in Section V.

Table A1. Robustness of Results

Dependent Variable	Time Dummies	Robustness Variables	ln Y	$(\ln Y)^2$	ln N	$(\ln N)^2$	F	R^2	N
D	no	no	0.094 (5.81)	-0.55E-02 (5.62)	0.71E-02 (1.01)	-0.13E-02 (0.94)	0.074 (7.07)	.815	1609
D	yes	no	0.119 (6.82)	-0.78E-02 (6.77)	-0.019 (1.86)	-0.16E-02 (1.21)	0.070 (6.54)	.819	1609
D	no	yes	0.109 (6.60)	-0.69E-02 (6.66)	-0.012 (1.39)	-0.90E-04 (0.07)	0.071 (6.85)	.825	1609
D	yes	yes	0.119 (6.74)	-0.77E-02 (6.62)	-0.020 (1.94)	-0.22E-03 (0.16)	0.070 (6.62)	.826	1609

D represents the share of distribution in GDP; country dummies and constants were included in all regressions but are not displayed here; t-ratios (absolute value) are in parentheses.

Table 5. The Distribution Sector and Development:The Log of Per Capita Income at PPP Rates.

Method	ln(ICP)	$(\ln(ICP))^2$	ERDI	MEDIA	URBAN	WOMEN	WOMENAFRICA	RULEOFLAW	AVGF	AVGN	R^2
WLS (1)	0.074 (3.20)	-0.012 (3.31)									.203
WLS (2)	0.064 (2.09)	-0.014 (2.59)	-0.26E-02 (0.43)	0.120 (2.79)	0.10E-03 (0.36)	-0.90E-03 (1.52)		-0.34E-02 (0.99)	-0.048 (0.52)	0.17E-07 (0.00)	.375
WLS (3)	0.065 (1.74)	-0.014 (2.37)	-0.23E-02 (0.32)	0.121 (2.59)	0.10E-03 (0.36)	-0.92E-03 (1.27)	0.39E-04 (0.05)	-0.34E-02 (0.98)	-0.048 (0.50)	0.15E-05 (0.03)	.375
TSLOGIT (4)	0.686 (3.32)	-0.113 (3.45)									.217
TSLOGIT (5)	0.595 (2.16)	-0.124 (2.61)	-0.015 (0.29)	0.973 (2.69)	0.94E-03 (0.39)	-0.75E-02 (1.51)		-0.030 (1.03)	-0.446 (0.55)	0.653E-04 (0.16)	.379
TSLOGIT (6)	0.656 (1.96)	-0.132 (2.50)	-0.42E-02 (0.07)	1.026 (2.60)	0.91E-03 (0.37)	-0.87E-02 (1.41)	0.19E-02 (0.32)	-0.031 (1.05)	-0.437 (0.53)	0.14E-03 (0.29)	.381
WLS C-hat (7)	0.109 (4.40)	-0.020 (4.81)									.375
WLS C-hat (8)	0.086 (2.51)	-0.021 (3.24)	-0.24E-03 (0.04)	0.128 (2.38)	0.33E-03 (0.92)	-0.89E-03 (1.45)		-0.56E-03 (0.15)	-0.061 (0.64)	-0.29E-03 (1.57)	.531
WLS C-hat (9)	0.078 (2.00)	-0.020 (2.93)	-0.22E-02 (0.30)	0.114 (1.80)	0.33E-03 (0.92)	-0.59E-03 (0.65)	-0.36E-03 (0.44)	-0.31E-03 (0.08)	-0.066 (0.68)	-0.31E-03 (1.60)	.534

General Notes: 47 observations; Constants were included in all regressions but are not displayed here; t-ratios (absolute value) are in parentheses; R^2 based on transformed data; for definitions of right hand side variables see the Data Appendix.

Table 6. J-Tests of Cross-Sectional Specifications.

	Dominant Spec. at ...								
Equation	0.1%	1%	5%	lnICP	ICP	lnICP	ln(AvgY)	lnICP	1/ICP
(1)	lnICP	lnICP	—	4.28	2.19				
(1)	—	—	lnICP			2.33	0.73		
(1)	—	—	—					1.94	0.60
(2)	—	—	lnICP	2.59	1.84				
(2)	—	—	lnICP			2.54	1.07		
(2)	—	—	lnICP					2.32	0.27
(3)	—	—	—	2.19	2.17				
(3)	—	—	lnICP			2.49	1.04		
(3)	—	—	lnICP					2.27	0.24
(4)	lnICP	—	—	4.58	2.96				
(4)	—	—	lnICP			2.58	0.83		
(4)	—	—	—					1.96	0.81
(5)	—	lnICP	—	2.92	2.06				
(5)	—	—	lnICP			2.58	1.16		
(5)	—	—	lnICP					2.23	0.46
(6)	—	—	—	2.57	2.39				
(6)	—	—	lnICP			2.55	1.15		
(6)	—	—	lnICP					2.22	0.38
(7)	lnICP	lnICP	lnICP	4.47	1.90				
(7)	—	lnICP	lnICP			2.82	0.68		
(7)	lnICP	lnICP	lnICP					3.94	0.29
(8)	—	—	—	2.48	2.20				
(8)	—	lnICP	lnICP			2.73	0.62		
(8)	—	lnICP	lnICP					3.27	0.24
(9)	—	—	—	2.03	2.65				
(9)	—	lnICP	lnICP			2.78	0.70		
(9)	—	lnICP	lnICP					3.18	0.17

Double border indicates t statistic significant at the 0.1% level; Single border indicates t-statistic significant at the 1% level; Dashed border indicates t-statistics significant at the 5% level. In the results columns, a “ — ” indicates that the test is inconclusive.

Table A2. The Distribution Sector and Development: Per Capita Income at PPP Rates.

Method	ICP	ICP2	ERDI	MEDIA	URBAN	WOMEN	WOMENAFRICA	RULEOFLAW	AVGF	AVGN	R^2
WLS (1)	-0.18E-03 (0.32)	0.28E-06 (0.05)									.033
WLS (2)	-0.17E-02 (1.70)	0.82E-05 (1.07)	-0.36E-02 (0.59)	0.122 (2.73)	0.23E-03 (0.72)	-0.14E-02 (2.78)		-0.33E-02 (0.88)	-0.024 (0.23)	-0.12E-04 (0.26)	.338
WLS (3)	-0.27E-02 (2.22)	0.16E-04 (1.72)	-0.90E-02 (1.25)	0.086 (1.68)	0.33E-03 (1.02)	-0.58E-03 (0.74)	-0.96E-03 (1.41)	-0.13E-02 (0.34)	-0.20E-02 (0.02)	-0.47E-04 (0.88)	.373
TSLOGIT (4)	-0.70E-03 (0.14)	-0.39E-05 (0.08)									.023
TSLOGIT (5)	-0.014 (1.61)	0.69E-04 (1.03)	-0.023 (0.44)	1.042 (2.73)	0.23E-02 (0.81)	-0.012 (2.85)		-0.029 (0.91)	-0.138 (0.15)	-0.66E-04 (0.15)	.343
TSLOGIT (6)	-0.023 (2.15)	0.14E-03 (1.70)	-0.069 (1.12)	0.696 (1.58)	0.32E-02 (1.12)	-0.49E-02 (0.74)	-0.85E-02 (1.44)	-0.010 (0.30)	0.086 (0.09)	-0.37E-03 (0.79)	.378
WLS C-hat (7)	-0.14E-03 (0.20)	-0.43E-05 (0.58)									.170
WLS C-hat (8)	-0.26E-02 (2.24)	0.13E-04 (1.30)	-0.44E-03 (0.07)	0.125 (2.15)	0.49E-03 (1.18)	-0.15E-02 (2.96)		-0.58E-03 (0.14)	-0.046 (0.42)	-0.30E-03 (1.57)	.500
WLS C-hat (9)	-0.42E-02 (2.94)	0.25E-04 (2.16)	-0.89E-02 (1.20)	0.046 (0.64)	0.68E-03 (1.63)	-0.28E-04 (0.03)	-0.15E-02 (1.82)	0.22E-02 (0.52)	-0.020 (0.19)	-0.35E-03 (1.83)	.542

General Notes: 47 observations; Constants were included in all regressions but are not displayed here; t-ratios (absolute value) are in parentheses; R^2 based on transformed data; for definitions of right hand side variables see the Data Appendix.

Figure 1. Pooled Sample Regressions

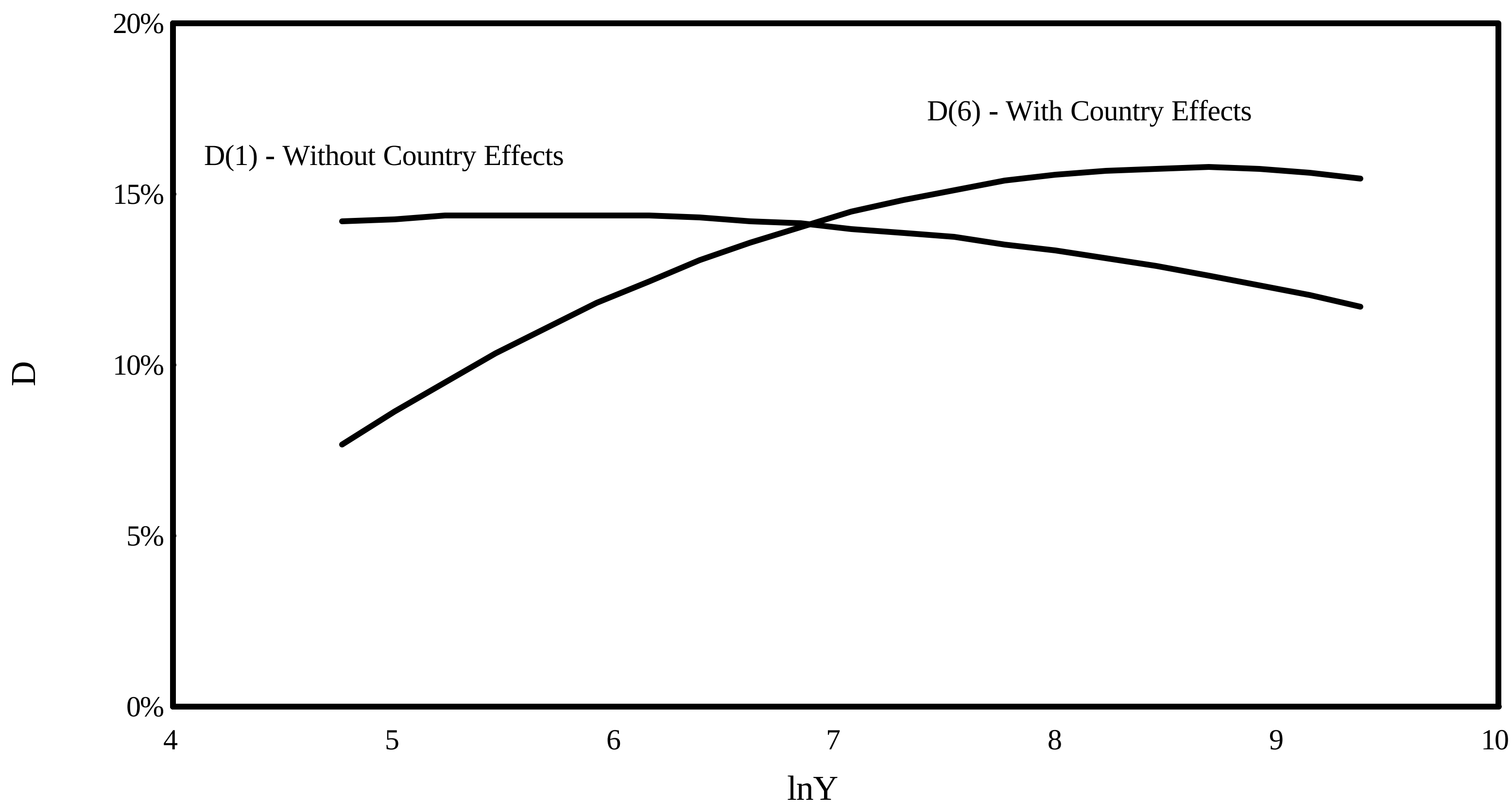

Figure 2. Balanced Panel Regressions
Country Effects Models

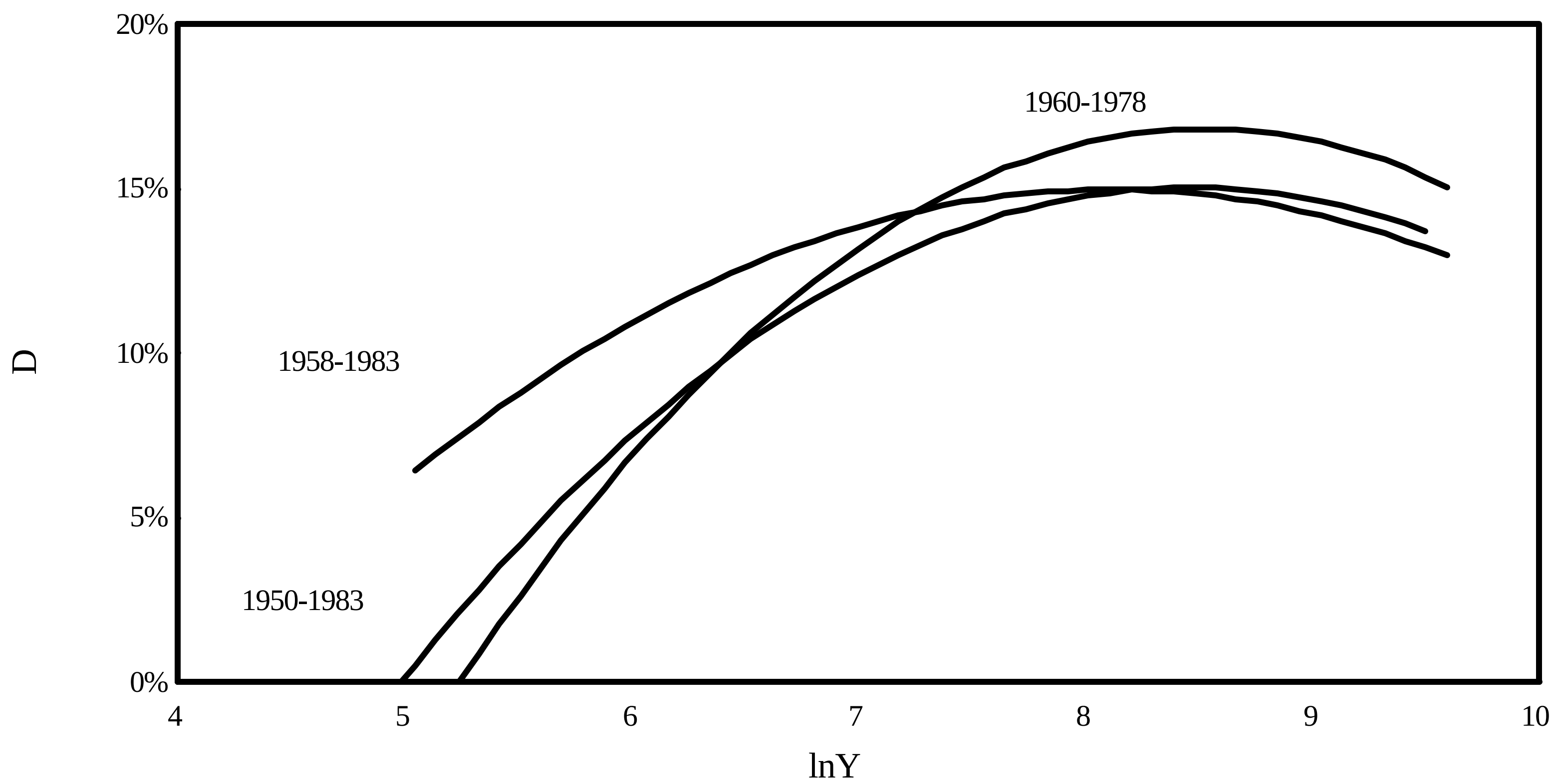

Figure 3a. Long-Run Pattern
Weighted Least Squares

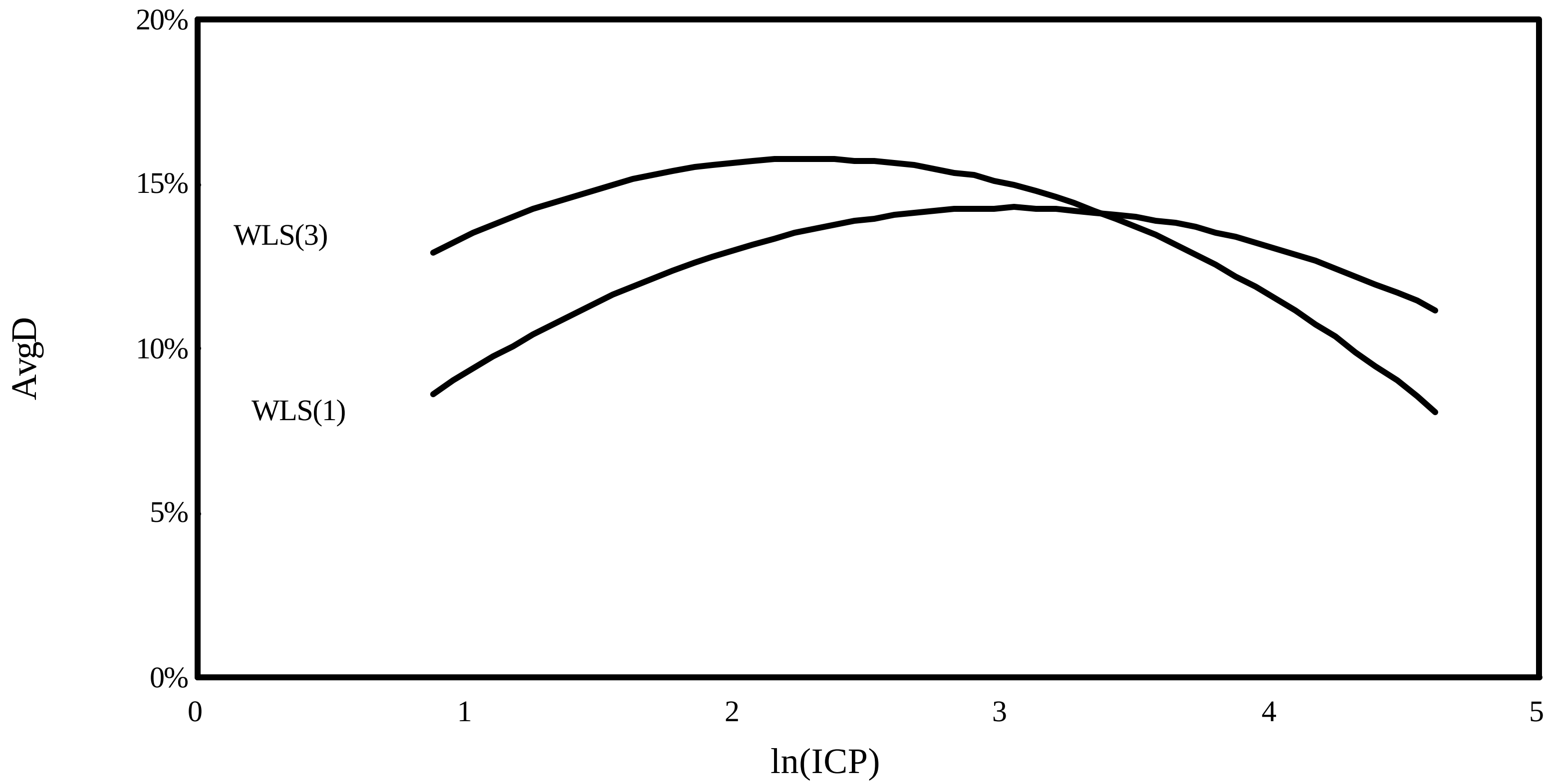

Figure 3. The Long-Run Pattern