DEPARTMENT OF ECONOMICS WORKING PAPER SERIES

Dual Economies and International Total Factor Productivity Differences

Areendam Chanda
Louisiana State University

Carl-Johan Dalgaard
University of Copenhaven and EPRU

Working Paper 2005-11
http://www.bus.lsu.edu/economics/papers/pap05_11.pdf

Department of Economics
Louisiana State University
Baton Rouge, LA 70803-6306
http://www.bus.lsu.edu/economics/

Dual Economies and International Total Factor Productivity Differences*

Areendam Chanda
Louisiana State University

Carl-Johan Dalgaard
University of Copenhagen and EPRU

June 2005
First Draft: January 2003

Abstract

This paper shows that a significant part of measured total factor productivity (TFP) differences across countries is attributable not to technological factors that affect the entire economy neutrally, but rather, to variations in the structural composition of economies. In particular, the allocation of scarce inputs between agriculture and non-agriculture is important. We provide a framework which maps the composition of the economy to measured aggregate TFP. A decomposition analysis suggests that as much as 85 percent of the international variation in TFP can be attributed to the composition of output. Estimation exercises indicate that recent findings of the conduciveness of good institutions, and, to some extent trade, on levels of TFP, may be thus explained.

Keywords: Development Accounting, Dual Economy, Structural Change, Total Factor Productivity, Institutions, Geography, Multisector Growth Models.
JEL Classification: O41, O47, O50

*Contacts: achanda@lsu.edu and carl.johan.dalgaard@econ.ku.dk. We are grateful to Francesco Caselli and Antonio Ciccone for sharing their data and for their comments. We would also like to thank Shekhar Aiyar, Oded Galor, Doug Gollin, Tom Grennes, Halvor Mehlum, Doug Pearce, John Seater, Jon Temple, Eric Thorbecke, Dietz Vollrath and seminar participants at University of Copenhagen, University of Malaga, European Economic Association Meeting in Stockholm, Midwest Macro Meetings, The Nordic Conference in Development Economics in Copenhagen, Hebrew University, Hunter College (CUNY), University of Houston, Louisiana State University, North Carolina State University, Utah State University, the XV Villa Mondragone International Economic Seminar, 75 Years of Development Economics Conference at Cornell, and NEUDC, for their comments. We are responsible for all errors. Parts of this research was conducted while Carl-Johan Dalgaard was visiting Universitat Pompeu Fabra's Economics Department, the hospitality of which is gratefully acknowledged. The activities of EPRU (Economic Policy Research Unit) are financed through a grant from The Danish National Research Foundation.

1 Introduction

The problem of economic development is often viewed as a problem of structural change. For economists such as William Arthur Lewis, the central problem of development was to be solved by ensuring that agriculture continued to maintain its production levels while workers moved to the nascent industrial sector. In similar vein, other classical theories of economic development such as those of Stages of Economic Growth (Rostow, 1961), the Big Push (Rosenstein-Rodan, 1943) or the Critical Minimum Effort Thesis (Leibenstein, 1960) essentially viewed the problem as one of poor countries being stuck in a poverty trap characterized by a "backward" agricultural sector and the challenge being one of ensuring a transition to modern industrial production.[1]

A recent outgrowth of the new growth theory has been increasing evidence suggesting that differences in living standards can be overwhelmingly accounted for by differences in total factor productivity (TFP), and not differences in the stocks of raw labor, human capital and physical capital. Klenow and Rodriguez-Clare (1997) and Hall and Jones (1999) were the initial studies suggesting that differences in TFP might account for more than 60% of the differences in output per worker. Not surprisingly, this has led to an increasing focus on explaining differences in TFP, often taken to mean technology, rather than factor accumulation.

In this paper we attempt to build a bridge between these recent developments and the more long-standing view that aggregate productivity is intimately related to the process of structural change. To achieve this end, we begin by undertaking a novel accounting exercise. The approach builds on a relatively mild set of assumptions beyond what is standard in the literature on development accounting. Essentially we only need to invoke national income accounting identities and the assumption of a Cobb-Douglas production function for the *non*-agricultural sector of the economy. On this basis we demonstrate how aggregate TFP levels, obtained using the standard cross-country development accounting methodology, can be further decomposed into a contribution from the composition of output and a contribution from the (non-agrarian) level of technology. The key finding is that the composition effect can explain as much as 85% of the differences in aggregate TFP across countries.

Much care should be taken in interpreting this finding. For example, it should not necessarily be taken to imply that technology is unimportant for observed TFP differences. Accounting work never yields insights into what, at a deeper level, generates the observed variation in factors of production, and in this case, the composition of output. But the finding strongly suggest that determinants of the allocation of scarce resources across sectors (technological or otherwise) are likely to be important determinants of observed differences in aggregate GDP per worker, consistent with the classical approach to development economics cited above. By extension, our results also suggest that a reasonable theory of aggregate TFP differences should be based on a multi-sectoral approach.

[1] This literature spawned a sizable neoclassical growth literature on dual economies. See e.g. Jorgenson (1961), Dixit (1970) , Razin and Mass-Colell (1973). More recent contributions include Laitner (2000) and Gollin, Parente and Rogerson (2002, 2004) and Banerjee and Duflo (2004).

In the second half of this paper we investigate what might explain the empirically observed variation in composition by way of regression analysis. Inspired by the recent literature which might be put under the heading: "fundamental determinants of productivity" (e.g. Hall and Jones; 1999; Acemoglu, Johnson and Robinson, 2001 and 2002; Easterly and Levine, 2003; Alcála and Ciccone, 2004), we examine the importance of factors such as institutional quality, trade and geographic circumstances for the allocation of resources across sectors. We do not attempt to gauge which of these factors is "the more important one" in understanding productivity differences. Instead, we are solely interested in taking a first pass at studying how the variation in inter-sectoral allocations, which seem to account for the bulk of the variation in aggregate TFP levels, comes about. Our estimates, based on both ordinary least squares and also after correcting for simultaneity bias, overwhelmingly suggest that the beneficial effects both institutions and trade seem to have on aggregate TFP are channeled through sectoral allocations of inputs and technological differences across sectors.

The present paper is related to the literature on development accounting which was pioneered by Krueger (1968) and King and Levine (1994). However, in our work we attempt to move beyond the standard approach, where GDP can be seen as being generated by combining aggregate stocks of capital (human and physical), an index of technology, and a neoclassical (typically, Cobb-Douglas) production function.

In abandoning what is essentially a one-sector approach, this paper is related to an emerging literature which analyses the implications of intersectoral reallocations for long-run productivity. The work of Graham and Temple (2004) shows how externalities in the non-agrarian sector may lead to multiple steady states in terms of aggregate productivity.[2] Other contributions have in a similar vein examined the aggregate consequences of misallocation of inputs due to e.g. barriers to capital accumulation (Restuccia, 2004), Harris-Todaro type wage gaps (Temple, 2003), imperfections in factor markets (Vollrath, 2004) and the role of fixed costs (Banerjee and Duflo, 2004).

The contribution most directly related to our accounting work is that of Caselli (2004) who also provides a development accounting analysis of the importance of inter-sectoral allocations of input across agriculture and non-agricultural sectors of the economy. Caselli's methodology can be viewed as a "bottom-up" approach to the issue at hand, which consists of specifying disaggregated production functions (in agriculture and outside agriculture) upon which counterfactuals are performed so as to assess the importance of intersectoral differences in TFPs, capital-output ratios etc, for aggregate labor productivity. In contrast, our methodology can be viewed as a "top-down" approach to the issue of how intersectoral allocations matter for *TFP* differences. The key distinguishing feature in the two approaches lie in that we do *not* make assumptions about the nature of the underlying production technology in agriculture. The main advantage is that it dispenses with the need for making assumptions about e.g. inputs and factor intensities in the agricultural

[2] The different steady states can be ranked according to the size of the agricultural sector; "high" income per capita is associated with relatively few resources being devoted to agriculture. They also demonstrate, by way of calibration, how their approach can motivate a "twin peaked" distribution of aggregate TFP, a result which fits with the empirical findings of Feyrer (2002).

production function, which remains a somewhat controversial issue. Of course, the drawback of our approach is that it is not informative about disaggregated TFP differences, whereas e.g. Caselli's approach is. In this sense our work complements Caselli's.

Overall, the present paper can be viewed as providing an assessment of the scope for inter-sectoral differences (attributable to underlying sources such as those mentioned above) to explain differences in aggregate productivity. Our findings, while not placing one of the above mentioned approaches above another, suggests that a strong case can be made in favor of theories for aggregate TFP differences that are grounded in dual economy type frameworks.

2 Accounting

We begin by invoking the national income identity which states that nominal GDP, pY, constitutes the sum of nominal value added in the agricultural sector (p_aY_a) and the non agricultural sector ($p_{na}Y_{na}$):

$$pY = p_aY_A + p_{na}Y_{na} = \left[\frac{p_aY_A}{p_{na}Y_{na}} + 1\right] p_{na}Y_{na}. \tag{1}$$

Accordingly, p is the GDP deflator (suitably defined) whereas $p_i, i = a, na$ is the price of agricultural and non-agricultural goods, respectively. Further, if we denote nominal labor productivity by lower case letters (e.g. $y_a = p_aY_a/L_a$), and the respective sectoral labor force shares by $\lambda_i, i = a, na$, then equation (1) can be used to express nominal GDP per worker as

$$\frac{pY}{L} = \left[\frac{y_a}{y_{na}}\lambda_a + 1 - \lambda_a\right] p_{na}\frac{Y_{na}}{L_{na}}. \tag{2}$$

Hence, aggregate labor productivity can be viewed as a multiple of labor productivity in non-agriculture.

In order to proceed we need to impose some additional structure. Hence, our key assumption is that the *non*-agricultural sector is characterized by a Cobb-Douglas production function:

$$\frac{Y_{na}}{L_{na}} = \left(\frac{K_{na}}{Y_{na}}\right)^{\frac{\alpha}{1-\alpha}} h_{na}A_{na}.$$

Since by definition the aggregate capital-output ratio, K/Y, and average economy-wide skill level, h, is proportional to the corresponding levels in the non-agricultural sector, i.e. $\frac{K_{na}}{Y_{na}} \equiv \kappa\frac{K}{Y}$ and $h_{na} = \eta h$, it follows that real GDP per worker can be written

$$\frac{Y}{L} = \left[\frac{y_a}{y_{na}}\lambda_a + 1 - \lambda_a\right] \frac{p_{na}}{p}\kappa^{\frac{\alpha}{1-\alpha}}\eta A_{na}\left(\frac{K}{Y}\right)^{\frac{\alpha}{1-\alpha}} h. \tag{3}$$

Obviously, the allocations κ, η are bounded from above, such that $\kappa \in (0, \overline{\kappa})$ and $\eta \in (0, \overline{\eta})$.

Now, consider the standard decomposition of *aggregate* labor productivity. Assuming the existence of an aggregate Cobb-Douglas production function such an exercise consists of decomposing GDP per worker Y/L into contributions stemming from capital input, labor input and a residual (TFP):

$$\frac{Y}{L} = TFP \cdot \left(\frac{K}{Y}\right)^{\frac{\alpha}{1-\alpha}} h. \tag{4}$$

The correspondence between our equation (3) and the aggregate TFP component is then obvious:

$$TFP = \left[\frac{y_a}{y_{na}}\lambda_a + 1 - \lambda_a\right]\left[\frac{p_{na}}{p}\right]\kappa^{\frac{\alpha}{1-\alpha}}\eta A_{na} \tag{5}$$

This suggests that the numbers calculated by e.g. Klenow and Rodriguez-Clare (1997) and Hall and Jones (1999) may mask an important contribution stemming from the intersectoral allocation of scarce inputs. Of course, at the same time, they may not. If average labor productivity is about the same in the two sectors $\frac{y_a}{y_{na}} \approx 1$, marginal costs are roughly the same (so that under competitive markets $p_a \approx p_{na}$) and the factor intensities across sectors are roughly equalized (suggesting the absence of barriers to factor mobility, so that $\kappa \approx \eta \approx 1$), then aggregate total factor productivity would trivially equal the total factor productivity in the non-agricultural sector: $TFP = A_{na}$. One could reasonable expect this to be the case in OECD countries where capital-output ratios and levels of labor productivity seems to be roughly similar across sectors.[3]

These considerations induce us to split equation (5) into two parts:

$$TFP = COMP * A_{na} \tag{6}$$

where $COMP$, or what we refer to as the "composition effect", is

$$COMP = \left[\frac{y_a}{y_{na}}\lambda_a + 1 - \lambda_a\right]\left[\frac{p_{na}}{p}\right]\kappa^{\frac{\alpha}{1-\alpha}}\eta. \tag{7}$$

Our aim is to see how much of the variation in the TFP can be accounted for by COMP and A_{na}, respectively. If a significant portion of aggregate TFP reflects variations in composition effects then clearly one can conclude that intersectoral differences are important in understanding the source of TFP differences. On the other hand if these differences seem unimportant then it appears that TFP differences are caused by factors that affect individual sectors roughly symmetrically. That is, factors that do not entail a reallocation of inputs across sector nor change relative levels of productivity in any substantive way.

While the methodology above may seem straightforward, data poses a challenge. The main concern is obviously that of international comparability. Specifically, internationally comparable data at the sectoral level is required. In addition we need intersectoral capital-output ratios, and intersectoral human capital levels. In the rest of this section we discuss how we try to overcome these obstacles and present our results.

To reconcile our decomposition with existing work on development accounting we need to produce PPP shares of agriculture value added in GDP. With such numbers in hand we can obtain PPP shares of non-agriculture value added in GDP using PPP GDP numbers from Penn World Tables (PWT) and proceed to calculate average labor productivity in the two sectors, (y_a, y_{na}) using labor share data from World Development Indicators.

[3] In their analysis of sectoral convergence of worker productivities and TFP in developed economies, Bernard and Jones (1996), find that the worker productivity in the agricultural sector lies within the same band as manufacturing and services in 1987, with the exception of Japan which has a relatively low agricultural productivity (See figure 2 in Bernard and Jones).

Fortunately, FAO constructs a set of international prices on agricultural commodities at an annual frequency. Rao (1993) used this data to construct cross-country data on agriculture value added, though only for 1985.[4] It would seem that by combining Rao's work with data from PWT we would obtain the shares we need for 1985. As Caselli (2004) observes, however, the FAO numbers cannot readily be subtracted from aggregate PWT numbers in 1985 to get PPP value added for the rest of the economy. The reason is that while PPP data for aggregate prices is normalized to 1 for the US in the PWT, PPP data for agricultural prices is normalized to 1 for the US in the FAO data. Of course, nothing guarantees that the implicit PWT PPP exchange rate for agriculture should be 1. Accordingly, we need to "renormalize" the FAO numbers.

In doing so we follow a method suggested by Caselli (2004). The key observation is that since PPP prices are quantity weighted, they tend to resemble rich country prices. Accordingly, one may expect that the PPP agricultural share of GDP should be approximately equal to the nominal GDP share in the US:

$$\frac{p_{A,US}^{PPP}Y_{A,US}^{PPP}}{Y_{US}^{PPP}} \approx \frac{p_{A,US}Y_{A,US}}{Y_{US}}$$

In this expression, the right hand side fraction (nominal share of agriculture) is available from World Development Indicators. The denominator on the left hand side is PPP GDP from the Penn World Tables. Accordingly, we can solve for $p_{A,US}^{PPP}Y_{A,US}^{PPP}$ and derive the the scaling factor:

$$s = \frac{p_{A,US}^{PPP}Y_{A,US}^{PPP}}{p_{A,US}^{FAO}Y_{A,US}^{FAO}}$$

where the superscript FAO denotes the FAO numbers. Multiplying all FAO PPP numbers on agriculture value added by s should make them comparable to PWT's GDP numbers for 1985. With comparable PPP numbers on GDP and GDP in agriculture in hand we can calculate the non agricultural GDP shares as well. Using employment shares $(\lambda_a, \lambda_{na})$ from World Development Indicators we can compute PPP numbers on labor productivity in the two sectors, y_a and y_{na}.[5]

Another input in our decomposition analysis is aggregate TFP numbers. Since value added for agriculture is available only for 1985, our accounting exercise will also be for 1985. In performing the basic development accounting analysis, we follow Hall and Jones (1999) (henceforth, HJ). Since HJ considered 1988 we have to redo their calculations. As HJ used 1985 numbers for human capital this bit of input data will be the same. The capital-output ratios for 1985 were calculated by replicating the HJ methodology exactly. Finally, invoking data on real GDP per worker from PWT version 5.6 we calculate TFP for 1985 as the residual in equation (4).[6] Below, we refer to these numbers as HJ-TFP.[7]

Turning next to the relative price p_{na}/p. Once we move to PPP numbers, rather than nominal values as assumed in the derivations above, this term becomes irrelevant for our variance

[4] Recently these numbers have been used by Caselli (2004) and Restuccia, Yang and Zhu (2004) in their work on calculating agricultural sector TFP's.

[5] While we have followed the methodology of Caselli (2004) we have used PWT 5.6 whereas he uses PWT 6.1. Thus the implied shares will still be different as the two versions have different base years. Version 5.6 uses 1985 as the base year, which is particularly useful for us since all our calculations are done for 1985.

[6] Capital's share α is put at 1/3.

[7] The correlation between our TFP numbers for 1985 and HJ's own estimates for 1988 is 0.98.

decomposition. To see this, consider equation (2). Measured in international prices the left hand side would be PPP GDP per worker, $\frac{Y^{ppp}}{L}$. We have already described how we obtain non-agrarian labor productivity in comparable PPP terms; call it $\frac{p_{na}^{ppp}Y_{na}^{ppp}}{L_{na}}$. Finally, assuming $\frac{Y_{na}^{ppp}}{L_{na}} = (K_{na}/Y_{na})^{\alpha/(1-\alpha)} h_{na}A_{na}$ and recognizing that p_{na}^{ppp} does not vary across countries, it follows that the relative price will not matter for the variance decompositions which follow below.

This still leaves us with $\kappa^{\frac{\alpha}{1-\alpha}}$ and η- the intersectoral allocations of capital and human capital. Below we take these allocations into account. But adding information on capital allocations reduces the size of the data set considerably. In addition, the numbers are probably a little more contentious than those invoked up until now. As a result we begin by sweeping the allocations $\left(\kappa^{\frac{\alpha}{1-\alpha}}, \eta\right)$ under the carpet, as a first pass.

Hence, as a first exercise we simply define[8]

$$COMP1 \equiv \left[\frac{y_a}{y_{na}}\lambda_a + 1 - \lambda_a\right] \tag{8}$$

and the residual as

$$RES1 = \frac{HJ - TFP}{COMP} \equiv \kappa^{\frac{\alpha}{1-\alpha}}\eta A_{na}. \tag{9}$$

This implies that RES1 (short for residual) picks up some of the variance in TFP which should be attributed to the actual composition effect. Whether this should increase or decrease the variation in TFP attributable to composition effects is not quite clear. To see this, note that the variance in TFP is the sum of the variances in COMP1, $\kappa^{\frac{\alpha}{1-\alpha}}, \eta$ and A_{na} and twice their covariances:

$$Var(\text{HJ-TFP}) = \sum_v Var(v) + \sum_v \sum_u 2Cov(v, u),$$

where $v, u =$ COMP1, $\kappa^{\frac{\alpha}{1-\alpha}}$, η, A_{na} (logarithms of all these variables) and $v \neq u$. On the one hand, attributing all of the variance of $\kappa^{\frac{\alpha}{1-\alpha}}$ and η to the residual would underestimate the share explained by composition terms. On the other hand it is not clear exactly in which directions the covariances will go. The covariance between η and A_{na} is probably likely to be positive, but there is little that one can say about the covariance of the remaining two combinations. Despite these limitations, it is instructive to see what the decomposition exercise between COMP1 and the rest looks like.

Table I reports some summary statistics for HJ-TFP, COMP and RES and also agricultural share in output (ASHARE) and worker productivity (RPROD=y_a/y_{na}). The thing to notice from the table is the enormous variation in relative worker productivity, ranging from 0.004 (Cameroon) to 1.44 (New Zealand). New Zealand is actually the only country that records a number higher than 1. Most OECD countries have relative productivity at the upper end of the distribution, but less than 1. Figure 1 plots relative productivity in ascending order. A fairly large number of countries have agricultural productivity that are less than 10% of the rest of the economy.relative

<Table 1 Here>

<Figure 1 Here>

[8] Since p_{na}^{ppp} is constant across countries, we supress it from now on.

Table II reports the correlations between the various variables. A few of the interesting features of this table are that HJ-TFP is strongly negatively correlated with agricultural share of output (Ashare) and has a strong positive correlation with COMP1, but a substantially lower correlation with RES1. Another interesting observation one can make is the very weak correlation between Ashare and RES1. This suggests that RES1 may in fact be picking up "across the board" factors that cause aggregate TFP to be low in contrast to factors that should move inputs (sector specific barriers, relative productivity levels and so forth).

<Table 2 Here>

The decomposition of TFP is undertaken the same way as done in Klenow and Rodriguez-Clare (1997) – taking logs of the levels of HJ-TFP, COMP and RES and then using the fact that:

$$1 = \frac{Var(\ln COMP1)}{Var(\ln \text{HJ-TFP})} + \frac{Var(\ln RES1)}{Var(\ln \text{HJ-TFP})} + \frac{2Cov(\ln COMP1, \ln RES1)}{Var(\ln \text{HJ-TFP})}$$

Table 3, column (1) below lists each of the terms in the above expression. Attributing half of the covariance to each of the two components, COMP1 can easily account for about 85% of the variation in TFP differences while RES1 explains only 15%. The most pessimistic allocation would be to contribute all of the negative movements in the covariance to COMP1. Even then, the composition effect explains as much as 56% of the total variation in aggregate TFP. To check for the sensitivity of the results, we removed Cameroon and New Zealand (the two countries with extreme relative productivity). Column (2) shows that there is really no change in the role of composition effects. As an additional sensitivity check, we dropped all OECD countries. Needless to say this does reduce some of the variation observed in the data but, as displayed in Column (3), we find only a modest drop in the share that is accounted for by COMP: to about 78%. Finally, we limited the sample to just OECD countries. The variation now motivated by the composition term drops considerably to 46%. This is of course to be expected. These are all countries that have low agricultural shares to begin with.[9] Nevertheless, these results clearly suggest an important role for the output structure of the economy.

<Table 3 Here>

Despite these encouraging results, one might still be concerned with the treatment of κ and η. We tackle these issues next.

There has been some progress towards estimating the stock of capital in the agricultural sector. In particular Crego, Larson, Butzer and Mundlak (1997) have estimated the fixed capital stock in agriculture for 62 countries for various years covering the period 1967-92. In addition to fixed capital stocks in agriculture, they also estimated fixed capital stocks in manufacturing and the

[9] Within this group Turkey has the highest share in agriculture at 18%. This is almost twice that of the next country, Greece (10%). Turkey also has the lowest relative productivity at 0.25. The sample correlation between agricultural share of output and relative agricultural productivity is -0.53 (22 observations).

entire economy.[10] The estimates of the latter are independent from those of ours (and hence also of HJ) and therefore it is easy to compare the reliability of the former's data, at least for the economy-wide measures. A simple correlation between the two data sets for the year 1985 produces a correlation of 0.95 for a sample of 53 countries. A regression (with the constant suppressed) of the Crego *et al* numbers on our numbers yield a coefficient of 0.93. Figure 2 plots the logarithm of fixed capital per worker for both the series. The strong correlation is quite obvious.

<Figure 2 Here>

As a result we proceeded to do a second decomposition where $\kappa^{\frac{\alpha}{1-\alpha}}$ is moved to the COMP term. We now have a more accurate composition effect, which we label "COMP2", and a correspondingly revised residual called RES2:

$$RES2 \equiv \frac{\text{HJ-TFP}}{COMP1 \cdot \kappa^{\frac{\alpha}{1-\alpha}}}$$

The decomposition results are presented in Table 4. We are now limited to a much smaller number of countries with the truncation taking place mainly at the lower end of the income distribution. Out of the 46 countries for which we have data, 19 (or 40%) are OECD countries. Despite that, the table suggests that the variance in COMP2 can explain around 60% of the variance in aggregate TFP. Columns (1) through (4) replicate the sample classification of Table 3. In Column (2) we drop New Zealand (data is no longer available for Cameroon) which hardly alters the results. In Column (3) we have only non OECD countries and in Column (4) we have only OECD countries. As before, once we restrict the sample to OECD countries the role of the composition effect is reduced but it remains sizeable.

When comparing Table 3 and 4 it is clear that once κ is taken into account we are forced to reduce the underlying country coverage by almost half. How much of the change in results are simply due to this reduction? To check this we conducted a decomposition in terms of COMP1 and RES1 (rather than COMP2 and RES2) but for the smaller sample where data on κ is available. The results are reported in Column 5, and they are virtually identical to the results reported in Column 1. Hence the reduction in the share accounted for by COMP is actually not produced by correcting for the allocation of capital. It is almost entirely driven by the reduction in the overall sample size. As a result one is still in a position to conclude that the composition effect is of substantial importance in accounting for aggregate TFP differences.

<Table 4 here>

As a final exercise we attempt to correct for human capital. Regrettably, we have been unable to find any research that provides intersectoral data on schooling (average years of schooling, schooling completion rates or even literacy rates). Not surprisingly, some of the recent work on intersectoral differences tend to ignore human capital completely (for example, Gollin, Parente

[10] In addition to fixed agricultural capital they also estimate a broader measure of capital stock which includes livestock and orchards (treestock). The aggregate fixed capital stock in their estimates is greater than the sum of fixed capital stocks in agriculture and manufacturing, leaving room for other sectors. That is, the three series are independent estimates.

and Rogerson, 2004; Graham and Temple, 2004; Restuccia, Yang and Zhu, 2004; Vollrath, 2004). However, in Caselli (2004) a simple correction is proposed which is outlined below. Given the lack of any other precedent, we redo our estimates following basically his approach.

In the development accounting literature, the widely adopted practice is now to use microeconomic based Mincerian returns and combine them with Barro-Lee (1999) estimates of schooling. What this implies is that the average human capital per worker in the economy can be calculated using a formula such as

$$h = \exp\left(\phi_p u_p + \phi_s u_s + \phi_\tau u_\tau\right),$$

where ϕ_p, ϕ_s and ϕ_τ represent the returns to an additional year of schooling at primary, secondary and higher levels, whereas u_p, u_s and u_τ represent the average years of schooling for an economy at each of these levels. This is the approach followed by HJ. Since we have applied their human capital numbers to calculate aggregate TFP, we stick with this approach.[11]

In following HJ we assume $\phi_p = 0.139$ (i.e. 13.9%), $\phi_s = 0.101$, $\phi_\tau = 0.068$ for all countries. We initially follow Caselli (2004) by assuming that labor in agriculture is completely uneducated; zero years of schooling at any level. Hence, if average human capital in the agricultural sector is denoted by h_a, then clearly $h_a = 1$ given this assumption. The average human capital in the non-agricultural sector is then easy to derive. First note that the average human capital in the labor force, h, must by definition be:

$$h = h_a \lambda_a + h_{na}\left(1 - \lambda_a\right),$$

which implies (since $h_a = 1$ by assumption)

$$h_{na} = \frac{h - \lambda_a}{1 - \lambda_a}.$$

This delivers a straightforward expression for η which can be calculated with the data in hand,

$$\eta = \frac{h_{na}}{h} = \frac{h - \lambda_a}{\left(1 - \lambda_a\right) h},$$

where h is the average stock of human capital from Hall and Jones.

What can we expect for η? In countries with very low agricultural labor shares $\eta \approx 1$, whereas η will tend to be larger than 1 in economies with very large agricultural labor shares (λ_a close to 1).[12] Hence, obviously η will be highly, and negatively, correlated with GDP per worker (and by extension aggregate TFP).

Table 5A lists the results of this new decomposition. In doing this decomposition we have ignored differences in κ, given their observed lack of importance.

<Table 5A Here>

[11] In contrast, Caselli (2004) uses a slightly different approach by assuming $h = \exp\{\phi u\}$, where u is average years of schooling.

[12] In our calculations the US and UK have the lowest values for η – a few decimal points higher than 1.

As can be seen from the table the roles of the composition term and the residual have been switched. It is now the residual which explains more than 60% of the variation in aggregate Total Factor Productivity. This result remains true even when we drop all OECD countries (column 3) and of course gets further accentuated when we focus only on OECD countries.

Why do the results change so much? The most probable reason is that COMP1 and η are negatively correlated (recall COMP3=COMP1$\times\eta$). As we saw from Table 2, COMP1 and aggregate TFP are positively correlated. As just explained, η and aggregate TFP are negatively correlated (cf Table 2: the correlation between TFP and agriculture's share is about -0.8). Hence, η and COMP1 will tend to move in opposite directions. Indeed, the correlation between them is -0.2 (for the 74 countries in Column 1). Not surprisingly then the variation in COMP3 will be much lower than the variation in COMP1 and thus the former will have a relatively more limited role in explaining aggregate TFP differences.

One take on these results would be that the burgeoning literature which argues that intersectoral factors play an important role in explaining TFP differences would do well to consider the role of human capital differences across sectors. Another take is that this approach to measuring the allocation of human capital is too crude. Certainly it is hard to escape a lurking feeling that average years of schooling in agriculture may well be greater than zero, even in the poorer countries of the world.

As a final robustness check we therefore tried another round of decompositions, but simply raised the average years of schooling in agriculture to 2 years (half the duration of primary schooling). Given that the rate of return for the first four years is 13.9%, this means that human capital in the agricultural sector is

$$h_A = \exp(0.139 * 2) = 1.32.$$

As a result, we can now redo the earlier calculation so as to obtained a revised calibration of η:

$$\eta = \frac{h - (1.32\lambda_a)}{h\,(1 - \lambda_a)}$$

Table 5B lists the decomposition results when this human capital allocation is applied. It seems that a small adjustment has had a major impact. Once again the composition effect seems to be key; except in the case of OECD countries (as expected). Unfortunately the covariance is negative and much higher in numerical magnitude, making any conclusion sensitive to its division between the two variance terms. Hence its difficult not to be left with the clear impression that further work on examining the role of human capital, in an intersectoral context, would be worthwhile.

<Table 5B Here>

The overall conclusion from the decompositions is that the allocation of inputs across sectors seems to of first order importance. In a broad sample of countries the contribution of the COMP term in accounting for aggregate TFP is at the very least 40% and possibly as high as 85%. This finding naturally fuels an interest in trying to discern what might be driving the composition in the first place. In the next section we take a first pass at examining this question.

3 Regression Analysis

Having shown that a substantial fraction of the variation in measured aggregate total factor productivity can be attributed to composition effects, we next undertake some econometric exercises to investigate how it is determined. In explaining the components of aggregate TFP we implement a very parsimonious specification, inspired by previous research. Hall and Jones (1999) where the first study which attempted to explain empirically observed variation in GDP per worker, *and* its components: Capital-output ratios, human capital stocks and TFP.

Their key finding is that of a highly significant impact from "Social Infrastructure". The variable is the mean of the Sachs and Warner index of openness to international trade (YrsOpen) and a measure of "Government Anti-Diversionary Policy" (GADP) – a composite average of five variables published in the International Country Risk Guide that measure country risk for international investors. The latter group of five variables were introduced into the literature by Knack and Keefer (1995). Hall and Jones were attempting to construct a variable that could adequately reflect "institutions and government policies that determine the economic environment within which economic individuals accumulate skills, [and] firms accumulate capital and produce output." Of course the average of the GADP variable and the Sachs-Warner variable is only a proxy for what constitutes SOCINF. We would like to examine whether the composition of output in the economy is driven by SOCINF. That is, essentially perform a simple channeling analysis of the HJ finding.

More recently Alcalá and Ciccone (2004) (AC) have extended the analysis by HJ, with particular attention paid to the importance of trade for differences in GDP per worker, and its underlying components. The particular new element in AC's analysis consists of introducing a novel measure of trade: "Real Trade". In previous research the trade variable of choice has been nominal imports plus exports, as a fraction of nominal GDP (Frankel and Romer, 1999). AC argues that a problem with this measure is that it is affected by relative price changes between tradeables and non-tradeables. Simply put, if increasing specialization leads to an increase in the relative price of non-tradeables, through the well-known Balassa-Samuelson effect, nominal trade to GDP may in fact decline; even though trade has increased and extended the degree of specialization. As a result, they suggest an alternative measure of trade which should be resistant towards this problem: nominal trade to PPP GDP or REAL TRADE. They find that real trade, once properly instrumented, explains GDP per worker as well as aggregate TFP. Consequently, we also include this measure of trade as an alternative to the Sachs and Warner index.[13] Throughout we use the GADP variable to proxy for institutions. Finally, inspired by AC – who in turn cite Ades and Glaesar (1999) and Alesina *et al* (2000) – we include controls for the size of the market measured by log population size (logPoP) and country size (logArea).

In sum, our baseline specification is on the form

$$\log Z = \gamma_0 + \gamma_1 GADP + \gamma_2 Trade + \gamma_3 \log POP + \gamma_4 \log Area + \varepsilon, \tag{10}$$

[13] Rodriguez and Rodrik (2001) criticizes the latter index as as a measure of trade. In the original HJ analysis, however, the authors do seem to rather think of the Sachs-Warner index as a measure of distortionary policies, rather than as a measure of trade per se.

where Z =TFP in 1985, the composition effect and the residual, "Trade" is either YrsOpen or REALTRADE; and ε is noise. In some specifications we will include HJ's SOCINF instead of the corresponding elements (GADP and YrsOpen). Also, inspired by AC we will, in some specifications, add continent controls.[14] In order to get as large a country coverage as possible we use COMP1 and RES1 as our dependent variables. Below we also comment on the results from examining COMP2 and RES2.

Table 6 reports correlations between key variables included in the analysis. A noteworthy feature of the correlation matrix is the relatively high positive correlation between COMP and Real trade, YrsOpen and GADP. In contrast, the correlation between our measure of institutions and the residual is in fact negative. The same goes for the association between RES and the trade variables.

$>$Table 6 $<$

In Table 7A and 7B we report the results from estimating equation (10) by OLS. The first three columns in Table 7A reexamine the impact of SOCINF on aggregate TFP in 1985.

$>$Table 7A and 7B$<$

As in HJ's study the variable is strongly and positively related to TFP. Splitting SOCINF into GADP and YrsOpen reveal that only the former is significantly related to TFP. In column 3 we substitute YrsOpen for AC's trade measure. The required data for REAL TRADE leads to the loss of 3 observations. As seen, REAL TRADE is significant in explaining TFP. It is also interesting to note that the size of population turns significant once the alternative trade variable is included. This is consistent with theories arguing in favor of significant scale effects on the level of output per worker. The change in significance is likely due to the fact that large countries tend to trade less (cf. Table 6). Hence, without controlling for trade, the scale variable may simultaneously be picking up a detrimental impact on TFP (less trade) and a positive impact (size of the market); thereby rendering it insignificant in column 1 and 2. In the remaining three columns of Table 7A and in Table 7B, we repeat this batch of regressions though substituting the dependent variable for COMP and RES. Overall, it seems that institutions and trade are more strongly related to COMP rather than RES. Taken at face value this would suggest that institutions and trade work so as to lower inefficiencies in the allocation of inputs across sectors, which in turn stimulates aggregate TFP.

A potential problem with the regressions above is that they are likely to suffer from simultaneity bias. HJ and AC stress this issue in the context of social infrastructure and trade, respectively. Therefore we need instruments for both social infrastructure and trade. Since our hypothesis is that institutions and trade determine the composition of output, there is no reason why we cannot

[14] In general the continent controls are not all significant. Hence we select the controls following a method suggested by AC (see footnote 9 in their paper). The method consists of eliminating the most insignificant controls sequentially until each remaining continent control is significant at 10%.

use the same variables as instruments as HJ for SOCINF. Similarly, we adopt the instrument for REALTRADE constructed by AC.

The instruments we selected include a) the fraction of the population speaking one of Western Europe's five main languages including English (EURFRAC), b) the absolute value of the latitude (ABSLAT) c) the logarithm of predicted trade share of an economy based on a gravity model that only uses a country's geographical and population figures (LOGFR and LogAC, respectively) and d) A measure of long run existence of formal governments (STATEHIST).[15] The first three variables come from Hall and Jones and were used to instrument SOCINF (except for LogAC, which is the fitted trade share constructed by AC, used to instrument for REAL TRADE). STATEHIST comes from Bockstette, Chanda and Putterman (2002).[16] This variable measures the length and coverage of formal states in current geographical borders over the past 2 millennia. The motivation is that a long experience with formal bureaucracies can lead to a greater stock of institutional capital which might position some countries more favorably than others in framing appropriate legal systems, property rights etc. Since the variable is based on actual histories from 1-1950 CE it is free of problems of reverse causality. Further Bockstette *et al.* show that it is a better instrument for social infrastructure than most of the instruments suggested by Hall and Jones. This is confirmed in Table 8, where STATEHIST is significant in explaining both SOCINF and YrsOpen (as well as REAL TRADE).

<Table 8 Here>

It is less successful in explaining the institutional variable GADP, however[17]. Another noteworthy feature is that logAC is strongly related to REAL TRADE. In all cases our instruments are jointly highly significant determinants of the endogenous variables.

Table 9A and 9B shows the results from estimating equation (10) by 2SLS. The first three columns recover results similar to those of HJ- SOCINF and its underlying elements are significant in explaining aggregate TFP. The table also lists the p-values for Hansen's test statistic for overidentifying restrictions. The values imply that the null hypothesis of orthogonality cannot be rejected.

The 3rd specification is inspired by AC's set of controls. We also find trade and population to be significant determinants of TFP. In magnitudes we find a smaller impact from REAL TRADE (.45 compared to 1.3), and a smaller impact from population (.17 compared with .34).[18] This difference

[15] Hall and Jones also used another variable, ENGFRAC, the fraction of the population that speaks English, as an instrumental variable. However, we ultimately dropped this variable, since, in our initial investigations, keeping this variable in the set of instruments led to poor identification results. Further, the variable fared poorly in terms of predictive power in the first stage regressions.

[16] Bockstette et al create different values for STATEHIST using different rates for "discounting the past". The variable here uses a 5% rate of discounting- the same that is used for all the econometric exercises in their paper.

[17] In another paper, Chanda and Putterman (2004) show that STATEHIST might have a negative effect on recent measures of institutional quality but a positive effect on post war economic growth. They speculate that this might be because high STATEHIST countries (mainly within LDC's) experienced a "reversal of fortune" (Acemoglu, Johnson and Robinson, 2002) but have probably been more successful in catching up after gaining independence compared to low STATEHIST countries.

[18] Cf Alcalá and Ciccone (2004), Table 6.

is likely caused by (a) smaller data coverage in the present context, (b) we use a different measure of institutions compared with AC's specification. In fact, whereas AC find their institutional measure to be insignificant, GADP turns out to be significant in column 3.

Turning to the key results of this section, which are reported in columns 4-6 and in Table 9b. The overriding theme of the results is that both institutions and trade seem to be significant determinants of COMP. The YrsOpen policy variable is not significant however. In contrast we find little evidence of either institutions or trade having an impact on the remaining residual. These findings suggest that a critical manifestation of institutions and trade may be that they matter for the allocation of inputs across sectors, and that they thereby affect aggregate TFP. In addition the size of the market seems also to matter for the composition effect, as reflected by the highly significant influence from the size of population.

In order to test the robustness of these findings we also ran the same set of regressions using COMP2 and RES2, whereby capital is also part of the COMP term (available upon request). In reduced form 2SLS we continue to find a significant effect of institutions and trade on aggregate TFP in 1985, though the data set is limited to only 45 countries. However, while GADP continues to work its way through the composition effect, REAL TRADE now turns significant in the regression where RES is the left hand side variable. Since, in this sample, almost half the countries are OECD where the size of the agricultural sector is modest, this new finding may simply be consistent with the notion that trade matters for the adoption of technologies – thus stimulating TFP. Still, even in this smaller sample GADP remains significant in explaining, not RES, but COMP.

These findings can be viewed as supportive of a number of different theories on productivity differences. For example, Galor, Moav and Vollrath (2005) present a theory of how institutional innovation (public schooling) may be blocked by the landed elite thus hampering industrialization and thereby over-all productivity. Since institutions matter for the sectoral composition of the economy, according to this theory, it would implicitly stipulate a link between aggregate TFP and the quality of (human capital promoting) institutions. Likewise, the finding of a conducive effect of scale provides some evidence in favor of theories which leaves a major role for the size of the market in spurring industrialization. Of course, it is at this stage not possible to distinguish between different mechanisms, such as coordination failures (Rosenstein-Rodan, 1943; Murphy et al, 1989), externalities (e.g Temple and Graham, 2004), fixed costs (e.g. Banerjee and Duflo, 2004) and so on.

4 Concluding Remarks

Prescott's (1998) call for a theory of total factor productivity has been accompanied by a large body of research which attributes differences in output per worker to technological differences (often assumed to be the same as TFP) generated by institutional barriers and, not unrelated, geographical factors, which hamper the adoption of socially profitable innovations. However, arguing that aggregate TFP is *solely* determined by technological factors is almost certainly wrong.

As Robert Solow (2001, p.285, 287) recently put it:

> It is certainly unwise to suggest that all economies are equally efficient at reallocating inputs across sectors. This difference will be reflected in $A(t)$, and maybe not only there [...] the non-technological sources of differences in TFP may be more important than the technological ones. Indeed, they may control the technological ones, especially in developing countries.

In this paper we have tried to take this observation one step further, by asking whether this influence is of any quantitative importance. We believe it is. Specifically we have demonstrated that a significant fraction of the observed variation in measured TFP is attributable to the allocation of inputs and differences in technologies across sectors, and furthermore, that the efficiency with which inputs are channeled to high productivity sectors is strongly affected by the institutional environment of individual economies, the extent of trade and size of the market (proxied by the size of population). In sum, it seems that in order to provide a rigorous theory of cross-country total factor productivity differences, a theory of output's structural composition will be an important component.

References

[1] Acemoglu, D., Johnson, S., Robinson, J. A., 2001. The colonial origins of comparative development: an empirical investigation. *American Economic Review*, 91, 1369-1401.

[2] Acemoglu, D., Johnson, S., Robinson, J. A., 2002, "Reversal of Fortune: Geography and Institutions in the Making of the Modern World Income Distribution," *Quarterly Journal of Economics,* 117 (4)

[3] Ades, Alberto F. and Edward L. Glaeser, 1999, "Evidence on Growth, Increasing Returns and the Extent of the Market," *Quarterly Journal of Economics*, 114(3):1025-1046.

[4] Alesina A., E. Spolaore and R. Wacziarg, 2000. Economic Integration and Political Disintegration. *American Economic Review*, Vol. XC, 1276-96.

[5] Alcala, Francisco, and Antonio Ciccone, 2004, Trade and Productivity, *Quarterly Journal of Economics,* Vol. CXIX, No. 2 (May), 613-46.

[6] Banerjee, Abhijit and Esther Duflo, 2004, Growth Theory through the Lens of Development Economics, forthcoming in Philippe Aghion and Steve Durlauf ed., *Handbook of Economic Growth.*

[7] Barro, Robert J. and Jong-Wha Lee, 1999, International Data on Educational Attainment: Updates and Implications, CID Working Paper no. 42, Harvard University.

[8] Bernard A. B., and C. Jones, 1996, Comparing Apples to Oranges: Productivity, Convergence and Measurement Across Industries and Countries, *American Economic Review*, 86:5, 1216-1238.

[9] Bockstette, V., A. Chanda and L. Putterman, 2002. States and Markets: The advantage of an early start. *Journal of Economic Growth*, December, 7 (4), 347-369.

[10] Caselli, Francesco, 2004, Accounting for Cross-Country Income Differences, Forthcoming in Philippe Aghion and Steve Durlauf ed., *Handbook of Economic Growth*.

[11] Chanda, Areendam and Louis Putterman, 2004, "State Effectiveness, Economic Growth, and the Age of States," forthcoming in Matthew Lange and Dietrich Rueschemeyer ed., *States and Development: Historical Antecedents of Stagnation and Advance,* Palgrave/Macmillan Press

[12] Crego, A., D. Larson, R. Butzer and Y. Mundlak, 1997. A New Database on Investment and Capital for Agriculture and Manufacturing. Working Paper, World Bank.

[13] Dixit, A., 1970. Growth Patterns in a Dual Economy. *Oxford Economic Papers,* 22, 229-34.

[14] Easterly, W. and R. Levine, 2003, "Tropics, germs, and crops: the role of endowments in economic development", *Journal of Monetary Economics*, 50:1, January 2003

[15] Feyrer, James, 2002, Convergence By Parts, mimeo, Dartmouth College.

[16] Frankel, Jeffrey A. and David Romer, 1999, "Does Trade Cause Growth?" *American Economic Review*, 89(3): 379-99

[17] Galor O., O. Moav and D. Vollrath, 2005. Land Inequality and the Emergence of Human Capital Promoting Institutions. Working Paper (Brown University)

[18] Gollin D., S. Parente and R. Rogerson, 2002. The role of Agriculture in Development. *American Economic Review, Papers and Proceedings*, 92 (2), pp.160-64.

[19] Gollin D., S. Parente and R. Rogerson, 2004, Farm Work, Home Work and International Productivity Differences, *Review of Economic Dynamics*, 7: 827-850..

[20] Graham, Bryan, and J. Temple, 2004. Rich Nations, Poor Nations: How Much can Multiple Equilibria Explain? IIIS Discussion Paper No. 17

[21] Hall, R. and C. I. Jones, 1999. Why do Some Countries Produce So Much More output per worker? *Quarterly Journal of Economics*, 114, s. 83-116.

[22] Jorgenson, D. W., 1961. The Development of a Dual Economy. *Economic Journal*, 71, pp. 309-34..

[23] King, R and R. Levine, 1994. Capital Fundamentalism, Economic Development and Economic Growth. Carnegie-Rochester Conference Series on Public Policy, 40, p.259-92

[24] Klenow, P. J and A. Rodriguez Clare, 1997. 'The Neoclassical Revival in Growth Economics: Has It Gone Too Far?', in Bernanke, B. and J-J. Rotemberg (eds.), *NBER Macroeconomics Annual*, Cambridge and London: MIT Press, p. 73-103.

[25] Knack, Stephen and P. Keefer, 1995, "Institutions and Economic Performance: Cross Country Tests Using Alternative Institutional Measures," *Economics and Politics*, 7:207-27.

[26] Krueger, A.O., 1968. Factor Endowments and Per Capita Income Differences among Countries. *Economic Journal*, 78, pp. 641-659

[27] Laitner J., 2000. Structural Change and Economic Growth. *Review of Economic Studies,* 67, p. 545-61.

[28] Leibenstein, H., 1960. *Economic backwardness and economic growth: studies in the theory of economic development*, John Wiley, New York

[29] Mas-Colell, Andreu and Assaf Razin, 1973, "A Model of Intersectoral Migration and Growth," *Oxford Economic Papers*, v25, n1 (March): 72-79.

[30] Murphy, K., A. Shleifer and R.W. Vishny, 1989." Industrialization and the Big Push". *Journal of Political Economy,* Vol. 97, pp. 1003-1026

[31] Prescott Edward, 1998. Needed: A Theory of Total Factor Productivity. *International Economic Review*, 39, 525-51.

[32] Rao D.S.P., 1993. Intercountry Comparisons of agricultural output and productivity. FAO Economic and Social Development Paper.

[33] Restuccia, D, D. Yang and X. Zhu, 2004. Agriculture and Aggregate Productivity: A Quantitative Cross-Country Analysis. Working Paper (Univerity of Toronto)

[34] Restucia, Diego, 2004, "Barriers to Capital Accumulation and Aggregate Total Factor Productivity," *International Economic Review*, Vol.45 No.1, February 2004, pp. 225-238.

[35] Rodríguez, Francisco and Dani Rodrik, 2000, Trade Policy and Economic Growth: A Skeptic's Guide to the Cross-National Evidence, Published in NBER Macroeconomics Annual 2000, eds. Ben Bernanke and Kenneth S. Rogoff, MIT Press for NBER, Cambridge, MA.

[36] Rosenstein-Rodan, P., 1943. Problems of Industrialisation of Eastern and South-Eastern Europe, Economic Journal, 53, p. 202-11.

[37] Rostow, W., 1961. *Stages of Economic Growth: A Non-Communist Manifesto.* Cambridge University Press.

[38] Solow, R., 2001. What have we learned from a decade of empirical research on growth? Applying Growth Theory across Countries. *World Bank Economic Review*, 15, pp. 283-288

[39] Temple, J., 2003, "Dualism and Aggregate Productivity," University of Bristol Discussion Paper no. 02/532.

[40] Vollrath, Dietrich, 2004, Factor Market Imperfections and Aggregate Productivity in Dual Economies, mimeo, Brown University.

5 Tables

TABLE 1: SUMMARY STATISTICS[19]

Variable	N	Mean	Std Dev	Min	Max
Agricultural Share	**74**	**0.09**	**0.07**	**0.007**	**0.279**
Relative Productivity	**74**	**0.21**	**0.26**	**0.004**	**1.44**
HJ-TFP (rel to USA)	**74**	**0.57**	**0.29**	**0.10**	**1.24**
COMP1(rel to USA)	**74**	**0.62**	**0.30**	**0.09**	**1.05**
RES1 (rel to USA)	**74**	**1.01**	**0.51**	**0.35**	**3.73**

TABLE 2: CORRELATIONS[20]

(n=74)

	AShare	RPROD	HJ–TFP	COMP1	RES1
AShare	**1**	**...**	**...**	**...**	**...**
RPROD	**-0.17**	**1**	**...**	**...**	**...**
HJ-TFP	**-0.78**	**0.42**	**1**	**...**	**...**
COMP1	**-0.72**	**0.55**	**0.79**	**1**	**...**
RES1	**-0.005**	**-0.25**	**0.22**	**-0.41**	**1**

TABLE 3: VARIANCE DECOMPOSITION[21]

Shares	1	2	3	4
Var(COMP1) share	**1.14**	**1.12**	**1.15**	**0.47**
Var(RES1) share	**0.44**	**0.42**	**0.59**	**0.54**
Cov(COMP1, RES1) share	**-0.29**	**-0.27**	**-0.37**	**-0.01**
Implied Share of COMP1	**0.85**	**0.85**	**0.78**	**0.46**
Implied Share of RES1	**0.15**	**0.15**	**0.22**	**0.53**
N	**74**	**72**	**52**	**22**

[19] The sample of 74 countries exclude countries in transition (formerly communist) and those with a mining sector greater than 15% of GDP.

[20] For HJ-TFP, COMP and RES1 we used logged values in calculating correlations. This is meaningful since the variance decomposition exercise undertaken later also requires the use of logs.

[21] Column (1) excludes all transition countries and those with mining shares greater than 15% of GDP. Column(2) additionally drops Cameroon and New Zealand. Column (3) restricts the sample to non OECD countries with mining shares less than 15% and not in transition.. Column (4) uses only OECD countries with mining shares less than 15% and not in transition.

Implied Shares are calculated by allocating the covariance equally to COMP1 and RES1.

Table 4: Variance Decomposition[22]

Shares	**1**	**2**	**3**	**4**	**5**
Var(COMP2) share	**0.60**	**0.59**	**0.58**	**0.51**	**0.64**
Var(RES2) share	**0.33**	**0.33**	**0.41**	**0.53**	**0.31**
Cov(COMP2, RES2) share	**0.03**	**0.04**	**0.01**	**-0.02**	**0.02**
Implied Share of COMP2	**0.63**	**0.63**	**0.59**	**0.49**	**0.66**
Implied Share of RES2	**0.36**	**0.37**	**0.42**	**0.51**	**0.33**
N	**46**	**45**	**27**	**19**	**46**

Table 5A: Variance Decomposition[23]
(With Human Capital)

Shares	**1**	**2**	**3**	**4**
Var(COMP3) share	**0.36**	**0.35**	**0.39**	**0.14**
Var(RES3) share	**0.58**	**0.57**	**0.72**	**0.72**
Cov(COMP3, RES3) share	**0.03**	**0.04**	**-0.06**	**0.07**
Implied Share of COMP3	**0.39**	**0.39**	**0.33**	**0.21**
Implied Share of RES3	**0.61**	**0.61**	**0.66**	**0.79**
N	**74**	**72**	**52**	**22**

Table 5B: Variance Decomposition[24]
(With Human Capital)

Shares	**1**	**2**	**3**	**4**
Var(COMP3) share	**0.97**	**0.95**	**1.06**	**0.29**
Var(RES3) share	**0.64**	**0.62**	**0.87**	**0.66**
Cov(COMP3, RES3) share	**-0.30**	**-0.28**	**-0.47**	**0.02**
Implied Share of COMP3	**0.67**	**0.67**	**0.59**	**0.31**
Implied Share of RES3	**0.34**	**0.34**	**0.40**	**0.68**
N	**71**	**69**	**49**	**22**

[22] Column (1) excludes all transition countries and those with mining shares greater than 15% of GDP. Column(2) additionally drops New Zealand. Column (3) restricts the sample to non OECD countries with mining shares less than 15% and not in transition.. Column (4) uses only OECD countries with mining shares less than 15% and not in transition.

Implied Shares are calculated by allocating the covariance equally to COMP2 and RES2.

[23] Column (1) excludes all transition countries and those with mining shares greater than 15% of GDP. Column(2) additionally drops New Zealand. Column (3) restricts the sample to non OECD countries with mining shares less than 15% and not in transition.. Column (4) uses only OECD countries with mining shares less than 15% and not in transition.

Implied Shares are calculated by allocating the covariance equally to COMP3 and RES3.

[24] Column (1) excludes all transition countries and those with mining shares greater than 15% of GDP. Column(2) additionally drops New Zealand. Column (3) restricts the sample to non OECD countries with mining shares less than 15% and not in transition.. Column (4) uses only OECD countries with mining shares less than 15% and not in transition.

Implied Shares are calculated by allocating the covariance equally to COMP3 and RES3.

Notice that the sample is smaller than in Table 5A. This is because assuming average years of schooling of 2 in agriculture is incompatible with the very low average years of schooling in aggregate for some countries.

Table 6: Correlation matrix for selected variables[25]

	1	2	3	4	5	6	7	8
	logA85	logComp	logRes	GADP	YrsOpen	LogRealTrade	logPop	logArea
logA85	1.00							
logComp	0.80	1.00						
logRes	0.20	-0.42	1.00					
GADP	0.53	0.61	-0.20	1.00				
YrsOpen	0.48	0.57	-0.21	0.69	1.00			
LogRealTrade	0.39	0.42	-0.11	0.66	0.57	1.00		
logPop	0.19	0.13	0.07	0.08	0.08	-0.32	1.00	
logArea	0.10	0.09	0.01	0.07	-0.02	-0.27	0.57	1.00

[25] The definition and sources of the variables are given in the text. The correlations are calculated for 69 countries where all necessary data are available.

Table 7A: OLS Regressions[26]

	1	2	3	4	5	6
	Dependent variable					
	logA85	logA85	logA85	logComp1	logComp1	logComp1
SocInf	1.23^{***} (7.25)			1.64^{***} (7.89)		
GADP		1.06^{***} (3.63)	1.61^{***} (4.94)		1.26^{***} (4.02)	2.23^{***} (7.94)
YrsOpen		0.36 (1.57)			0.57^{**} (2.37)	
Real Trade			0.27^{**} (2.35)			0.25^{**} (2.09)
LogPop	0.07 (1.12)	0.07 (1.17)	0.14^{***} (3.24)	0.04 (0.84)	0.04 (0.94)	0.14^{**} (2.62)
LogArea	0.01 (0.13)	−0.002 (0.05)	−0.01 (0.13)	0.01 (0.22)	0.004 (0.07)	−0.01 (0.32)
Continent controls	No	No	Yes	No	No	Yes
N	72	72	69	72	72	69
R^2	0.32	0.34	0.61	0.41	0.41	0.74

Table 7B: OLS Regressions (Contd)

	1	2	3
	Dependent variable		
	logRes1	logRes1	logRes1
SocInf	-0.41^{**} (2.34)		
GADP		−0.21 (0.71)	-0.62^{*} (1.84)
YrsOpen		−0.20 (1.23)	
Real Trade			0.02 (0.12)
LogPop	0.03 (0.5)	0.03 (0.5)	0.01 (0.11)
LogArea	−0.01 (0.14)	−0.01 (0.14)	0.01 (0.16)
Continent controls	No	No	Yes
N	72	72	69
R^2	0.07	0.07	0.37

[26] All regressions contain a constant term. Absolute t-values in parentheses, based on robust standard errors. Continent controls included: Latin America, South Asia and Middle East/North Africa. Note: *,**,*** means significant and 10%, 5% and 1%, respectively.

Table 8: First stage regressions[27]

	1	2	3	4	5
	Dependent variable				
	SocInf	GADP	GADP	YrsOpen	logReal Trade
STATEHIST	0.22* (1.69)	−0.08 (0.90)	−0.03 (0.40)	0.52** (2.61)	0.41* (1.74)
EurFrac	0.18*** (3.88)	0.096*** (2.69)	0.19*** (4.95)	0.26*** (3.83)	0.37*** (3.04)
LogFR	−0.013 (0.17)	0.000 (0.00)		−0.025 (0.25)	
LogAC			0.07 (0.98)		0.62*** (3.26)
Abs.Lattitude	0.01*** (3.93)	0.01*** (7.47)	0.07*** (5.44)	0.01* (1.98)	0.003 (0.61)
LogPop	−0.007 (0.27)	0.01 (0.50)	0.01 (0.46)	−0.03 (0.61)	−0.22*** (4.18)
LogArea	−0.001 (0.004)	0.000 (0.00)	0.01 0.56	−0.02 (0.06)	0.17** (2.42)
Continent controls	No	No	Yes	No	Yes
N	72	72	69	72	69
R^2	0.49	0.60	0.76	0.38	0.64
F-stat (P-value)	29.94*** (0.000)	31.24*** (0.000)	41.07*** (0.000)	28.36*** (0.000)	20.76*** (0.000)

[27] All regressions contain a constant term. Absolute t-values in parantheses, based on robust standard errors. Continent controls included: Latin America, South Asia and Middle East/North Africa. Note: *,**,*** means significant and 10%, 5% and 1%, respectively. R^2 refers to the centered R^2.

Table 9A: 2SLS regressions[28]

	1	2	3	4	5	6
	Dependent variable					
	logA85	logA85	logA85	logComp1	logComp1	logComp1
SocInf	2.42*** (7.68)			2.73*** (7.72)		
GADP		1.22* (1.75)	1.96*** (4.40)		1.72** (2.55)	2.27*** (4.72)
YrsOpen		1.20** (2.27)			1.09** (2.18)	
logReal Trade			0.43* (1.91)			0.45** (2.18)
LogPop	0.03 (0.48)	0.04 (0.47)	0.17*** (3.03)	0.01 (0.17)	0.02 (0.28)	0.17*** (2.70)
LogArea	0.02 (0.38)	0.02 (0.35)	0.002 (0.07)	0.03 (0.44)	0.02 (0.29)	−0.003 (0.09)
Continent controls	No	No	Yes	No	No	Yes
N	72	72	69	72	72	69
Hansen's J-Static (P-value)	2.85 (0.41)	2.82 (0.24)	0.23 (0.89)	2.98 (0.34)	3.14 (0.21)	4.41 (0.11)

Table 9B: 2SLS regressions[29]

	1	2	3
	Dependent variable		
	logRes1	logRes1	logRes1
SocInf	−0.3 (1.14)		
GADP		−0.49 (0.96)	−0.31 (0.74)
YrsOpen		0.10 (0.23)	
log Real Trade			−0.02 0.08
LogPop	0.02 (0.44)	0.02 (0.3)	−0.01 (0.10)
LogArea	−0.01 (0.11)	0.03 (0.07)	0.005 (0.14)
Cont. Dum	No	No	Yes
N	72	72	69
Hansen's J-static (p-value)	10.8 (0.01)	11.3 (0.003)	2.23 (0.32)

[28] All regressions contain a constant term. Absolute t-values in parentheses, based on robust standard errors. Continent controls included: Latin America, South Asia and Middle East/North Africa. Instruments are: State history, Eurfrac, , absolute lattitude and fitted trade share. In column 1, 2, 4 and 5 it derives from Frankel and Romer (1999), in Column 3 and 6 the fitted trade share is from Alcalá and Ciccone (2004). Note: *,**,*** means significant and 10%, 5% and 1%, respectively.

[29] All regressions contain a constant term. Absolute t-values in parentheses, based on robust standard errors. Continent controls included: Latin America, South Asia and Middle East/North Africa. Instruments are: State history, Eurfrac, , absolute lattitude and fitted trade share. In column 1, 2 it derives from Frankel and Romer (1999), in Column 3 the fitted trade share is from Alcalá and Ciccone (2004). Note: *,**,*** means significant and 10%, 5% and 1%, respectively.

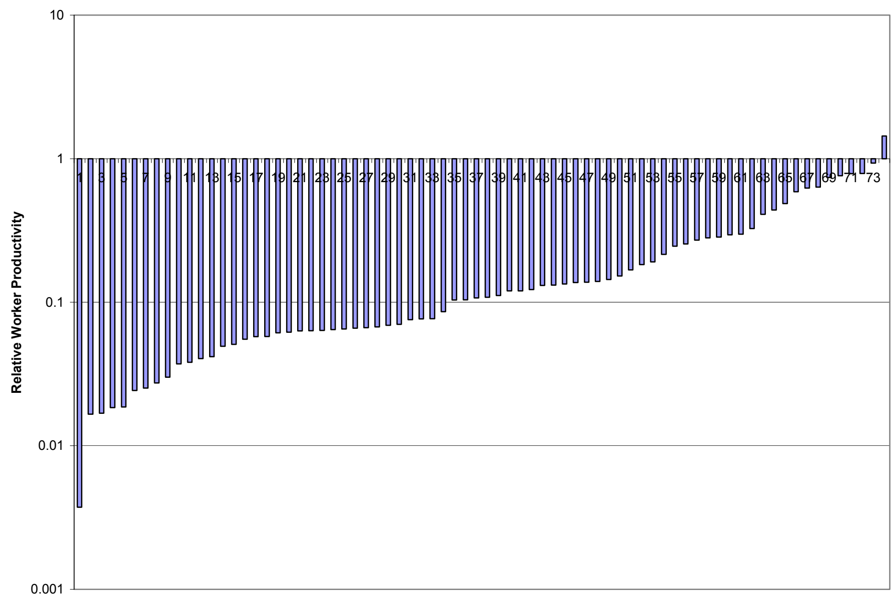

Figure 1: Countries Arranged in Order of Relative Productivity (Labor Productivity in Agriculture Relative to Rest of the Economy) (1985)

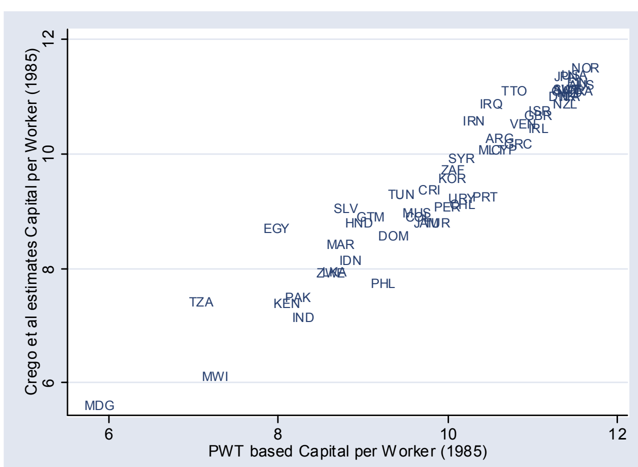

Figure 2: Comparing Capital per Worker (1985) Estimates