Determinants of Productivity: A Two-Stage Analysis

Nazrul Islam

International Centre for the Study of East Asian Development (ICSEAD)

Abstract

Following Solow's (2001) recent advice, this paper adopts productivity as the left-hand-side variable and offers a cross-country analysis of its determinants. The analysis follows the two-stage methodology, the first of which is devoted to obtaining productivity estimates, and the second stage is devoted to analyzing these estimates. The paper classifies productivity determinants into four types, namely 'economic factors,' 'institutions,' 'social base,' and 'physical base.' The empirical productivity model presented in this paper includes productivity determinants of all these different types. The empirical relationships established in this paper should prove helpful in further development of a theory of total factor productivity (TFP). (*JEL Classification:* O4, O3, and O1; *Keywords:* Economic Growth, Productivity, and Development.)

Current Draft: September 2005

(Please do not circulate or quote)

I would like to thank Dale Jorgenson, Jesus Felipe, and Joy Mazumdar for their helpful comments. Ning Liu, Heather McCraw, and Deliana Kostova provided valuable research assistance. All remaining errors are mine. Send your comments to nislam@icsead.or.jp

Determinants of Productivity:
A Two-Stage Analysis

1. Introduction

Following Solow's (2001) recent advice, this paper adopts productivity as the dependent variable and offers a cross-country analysis of its determinants. There are two reasons why Solow suggests productivity to be taken as the dependent variable. The first is the recent finding that productivity differences are more important than differences in accumulation rates in explaining per capita income differences across countries. The second reason is methodological. In paying attention to productivity differences, many researchers have resorted to growth regressions with *extended specifications*. Under this methodology, variables supposed to proxy for productivity differences are added directly to the right hand side of standard growth regressions. Solow feels uncomfortable with this methodology, because it lumps together in the same regression variables that have very different theoretical status. He therefore suggests that these additional variables should be taken as right hand side variables of regressions that have productivity itself as the dependent variable.

Implementation of Solow's suggestion however requires a *two-stage* methodology, the first of which is devoted to obtaining productivity estimates, and the second stage is devoted to analyzing these estimates. The two-stage methodology thereby allows productivity differences occupy the center-stage of analysis. This contrasts with the 'extended specification' methodology, where productivity differences are considered as a sort of 'nuisance,' to be somehow pushed aside in order to estimate accumulation parameters such as the elasticity of output with respect to physical or human capital. Ironically, treatment of productivity as a 'nuisance' does not allow bias-free estimation of accumulation parameters either. In contrast, by separating productivity analysis from estimation of accumulation parameters, the two-stage methodology creates better conditions for bias-free estimation of the accumulation parameters.

By now researchers have offered a host of variables as determinants of productivity. These may be classified into the following four types.

(A) Economic factors, such as 'size of the government,' 'openness of the economy,' etc.

(B) Institutions, such as the 'political stability,' 'democratic rights,' etc.

(C) Social base, such as the 'ethnic and religious composition of the population,' etc.

(D) Physical base, such as the 'location of a country,' 'climate,' 'access to sea,' etc.

The Physical Base variables are often referred to as Geography variables. In recent years, there has been some spirited debate about the role of the Physical Base variables. While Sachs (2003) and his co-authors maintain that Physical Base variables exert some *direct* influence on the economic performance of an economy, other researchers such as Acemoglu et al. (2001) and Rodrik et al. (2004) have argued that the influence of Physical Base variables is mediated entirely through other, such as institutions, variables.

Separation of the productivity analysis into a different stage however does not solve the problem of indeterminacy regarding which variables to be included in the productivity model. The problem here is similar to that faced by the adherents of the extended specification approach to growth regressions. One response to this problem has been adoption of what may be called the statistical approach. This approach was initiated by Levine and Renelt (1992) who used Leamer's (1983) bound in order to find out which variables prove to be robust in growth regressions. This approach has been extended further in recent years by Sala-i-Martin (1997, 1999), who has suggested more refined statistical methods to resolve the issue.

However, many researchers have found the purely statistical approach to determining robust determinants of economic performance not very satisfactory. They have instead emphasized the causality among possible determinants. Hall and Jones (1999) for example try to explain economic performance by the single underlying variable, Social Infrastructure. Similarly, both Acemoglu et al. (2001) and Rodrik et al. (2004) try to explain economic performance by quality of institutions. The approach adopted in these papers may be termed as the causality approach.

This paper adheres to the causality approach and extends it further. It constructs an elaborate causality scheme that subsumes most of the causality schemes proposed in the literature. The paper then allows this causality scheme to guide the regression specification. In this regard, the paper follows a sequential approach rather than the encompassing approach advocated by Hendry (1988). Given the huge number of possible candidates (for inclusion in the regression) and given the explicitly laid out causality scheme, the sequential approach appears to be more suitable for the analysis of this paper.

The analysis of this paper does not lead to a mono-causal model of productivity. Instead, elements of all the four types of productivity determinants find some space in it. For example, the results support the view that Physical Base variables influence aggregate productivity primarily through their influence on other 'intermediate' variables. However, whether or not a country is landlocked seems to have some remaining direct effect. Similarly, Social Base variables (such as ethno-linguistic and religious fractionalization) seem to retain significant direct effect on productivity. Some researchers, such as Alesina et al. (2003), have earlier tried to explain away the direct effect of social fractionalization by including additional variables in the model. Their success in this regard however remains questionable. Instead of banishing fractionalization from the model by any means, this paper rather draws attention to the issue so that further research is directed to it. Among Institutions variables, prevalence of democratic rights displays some direct influence on productivity. Here again, more research may reveal more fully the validity and/or nature of this direct effect.

Of the Economic Determinants, distortions in the economy, as measured by the black market premium for a country's currency, is found to affect productivity negatively. In addition, government consumption expenditure, as measured by this expenditure as a ratio to GDP, appears to have a very robust negative association with productivity. So far as externalities from input accumulation are concerned, the results suggest that such externalities are true more of human capital accumulation rather than of physical capital accumulation.

The productivity model arrived at in this paper is therefore complex, perhaps reflecting aptly the complexity of the real phenomenon that we call productivity. However, the model may be more expedient from policy point of view. Mono-causal models, assuming that they are valid, are not always that helpful in this regard. For example, papers concluding that institutions explain it all, beg the question how a country having bad institutions can change their institutions. In fact papers, such as of Acemoglu et al. (2001), which point to Physical Base variables as the determinant of institutions leave these countries in a hopeless situation, because a country usually cannot alter her physical dimensions. To the extent that the productivity model presented in this paper has more elements in it, it offers more hooks for policy levers to work on. In particular, the presence of both economic and non-economic variables in the productivity points to the necessity of both economic and political policies for improving productivity.

In comparing the results of this paper with those of previous studies, one has to keep in mind that while this paper studies determinants of *productivity*, most of the previous studies have examined determinants of *income* level or growth. While a determinant of productivity will also be a determinant of income, the converse is not necessarily true. For example, it is quite possible for a determinant to influence output via investment rate and not so much via productivity.

The paper is organized as follows. Section 2 provides the background by noting empirical importance of productivity differences and examining this importance from the perspective of both the neoclassical growth theory (NCGT) and the new growth theory (NGT). Section 3 provides arguments for the two-stage methodology of productivity study. Section 4 discusses various productivity determinants suggested in the literature. Section 5 presents the main empirical results. Section 6 discusses the results further. Section 7 concludes.

2. The importance of productivity differences

Solow's (2001, p. 287) suggestion that "…TFP (or its growth rate) should be the left hand side variable" in cross-country analysis of economic performance comes in response to the Easterly and Levine (2001) study showing that TFP, rather than factor accumulation, is the main explanatory factor of per capita income differences across countries. Elaborating further, Solow advises that "…comparative growth studies should focus on understanding and analyzing the various sources of differences in TFP and the policies that might affect them." (p. 286) He therefore recommends that cross country growth studies should focus "… more directly on TFP or factor augmentation as the proper left-hand-side variables." (p. 285) Romer (2001, p. 225) expresses similar views. He laments that many recent growth researchers abandoned the lead of growth researchers of the 1960s who allowed technology/productivity to differ across countries.[1] Romer therefore welcomes the current recognition of productivity differences across countries. Klenow (2001, p.221) concurs that "… TFP should be the focus of growth research." Easterly and Levine (2001, p. 177) themselves recommend that "Economists should devote more effort toward modeling and quantifying TFP."

The fact that productivity differences are important has been pointed out earlier by a number of papers including King and Levine (1994), Islam (1995, 2003), Hall and Jones (1996, 1999), and Prescott (1998). For example, proceeding from the following production function,

[1] "When economists in the 1950s and 1960s used growth models to understand the experience of developing countries, they allowed for the possibility of technology differences between developing countries and the United States. But because they did not have a good theory for talking about the forces that determined the level of the technology – in the United States any more than in developing countries – technology factors tended to be pushed into the background in policy discussions." (Romer 2001, p. 225)

(1) $$Y_i = K_i^{\alpha} (A_i H_i)^{1-\alpha},$$

where *Y* is output, *K* is physical capital, and *H* denotes human capital, Hall and Jones (1999) find that of the 35-fold difference in per capita income between the US and Nigeria, the difference in physical capital intensity accounts for a factor of 1.5, and the difference in educational level accounts for another factor of 3.1, but the difference in *A* accounts for a factor of 7.7. Similalrly, using the aggregate production function,

(2) $$Y_i = K_i^{\alpha} (A_i L_i)^{1-\alpha},$$

Islam (1995) finds that in a sample of 96 countries, the highest estimated value of the productivity level, *A*, was 39 times greater than its lowest estimated value. Prescott (1998, p. 1) actually goes so far as to declare that in explaining income differences across countries, "saving rate differences do not matter, all that is important is total factor productivity."

The fact that the TFP term *A* proves to be so important in explaining cross-country income and growth differences should not be surprising. In fact, both NCGT and NGT allow for such a role. In NCGT, TFP is often represented by the Hicksian shift term, *A(t)*, in the production function,

(3) $$Y(t) = A(t)\, F[K(t), L(t)],$$

and is generally associated with technological progress. However, there was a parallel recognition that TFP or the *A(t)* term was not all technological progress. To emphasize the point, Solow (1988) draws a distinction between 'technological progress in the narrow sense' and 'technological progress in the broader sense.' More recently, Mankiw, Romer, and Weil (1992) note that "…the *A(0)* term reflects not just technology but resource endowments, climate, institutions, and so on; it may therefore differ across countries." (p. 6) Thus, even if technology in the narrow sense in NCGT is exogenous, many of these other components of *A* may clearly vary across countries.

Furthermore, suppose A_i and A_j represent only the 'technology in the narrow sense' of countries *i* and *j*, respectively. Going by the usual neoclassical specification, $A_i = A(0)_i\, e^{g_i t}$ and $A_j = A(0)_j\, e^{g_j t}$. NCGT *per se* does not imply that $A(0)_i = A(0)_j$ and/or $g_i = g_j$. Yet any departure from these equalities leads to differences in A_t across countries. Thus, NCGT can allow even the 'technology in the narrow sense' to differ across countries.

In most NGT models, on the other hand, the productivity term *A* does not have to represent technological progress, because these models have alternative avenues for technological progress and/or source of long run growth. In these models *A* can therefore stand entirely for non-technological components, which may obviously vary across countries.

Thus both NCGT and NGT allow *A* to vary across countries. The question is whether these theories can support the empirical domination of *A* in explaining per capita income difference across countries. Introspection suggests that they do, because both these theories allow changes in *A* to affect the economic performance along several different routes, as follow:

Direct Effect: Under NCGT, changes in *A* lead to transitional growth *directly*, because *A* is a determinant of the steady state income *level* in this model.[2] The long-run (steady state) *growth rate*, being exogenous, remains unaffected, however. This is often summarized in the statement that improvements in *A* under NCGT have *level effect* only and no *growth effect*. While some have viewed this as a lack of potency of *A* under NCGT, that is actually not the case. As Jones (2002) emphasizes in the context of the post-war US economy, a succession of transitional growth effects induced by continual improvements in *A* may virtually be indistinguishable from a situation of permanent growth effect. So far as NGT is concerned, improvements in *A* cab have a *direct growth effect*, because the formulae of the equilibrium growth rate of most NGT models include *A* as an explicit argument.[3]

Indirect Effect: Given the wide scope of *A*, it is quite likely that investment and labor force growth rates themselves depend on it. In other words, we may have $s = s\,(A, \ldots.)$ and $n = n\,(A,$

---

[2] For example, proceeding from production function in (1), the steady state per capita income level, y^*, solves out to be:

$$y^* = A(t)\left(\frac{s}{n+g+\delta}\right)^{\frac{\alpha}{1-\alpha}} = A_0 e^{gt}\left(\frac{s}{n+g+\delta}\right)^{\frac{\alpha}{1-\alpha}},$$

showing that y^* depends on the vector [$A, s, n, g, \delta, \alpha$], where *n* is the growth rate of the work force, *s* is the rate of investment (measured as a ratio to GDP), and δ is the rate of depreciation.

[3] For example, the equilibrium growth rate, γ, for the Romer (1990) model solves out to be as follows:

$$\gamma = \frac{1}{\theta}\left[\left(\frac{L}{\eta}\right)A^{\frac{1}{1-\varphi}}\left(\frac{1-\varphi}{\varphi}\right)\varphi^{\frac{2}{1-\varphi}} - \rho\right],$$

where η is the fixed cost per invention, which Romer takes as given, and φ is the exponent with which the (intermediate) inputs enter the production function for the final goods. Clearly changes in *A* change γ.

....). That being the case, changes in *A* can induce transitional growth *indirectly* through its influence on *s* and *n*, which are also determinants of the NCGT steady state income level. Similarly, it is likely that $\rho = \rho(A, ...)$ and $\theta = \theta(A, ...)$, where ρ is the (time) discount rate, and θ is the inter-temporal elasticity of substitution. Accordingly, changes in *A* can have *indirect growth effect* under NGT too, because the equilibrium growth rate of many NGT models includes ρ and θ, as we just saw above.

Feedback Effect: Finally, it is possible that $A = A(s, n, ...)$ and $A = A(\rho, \theta, ...)$, so that improvements in *s* and *n* (or in ρ and θ) can cause improvements in *A*. Such feedback effects can lead to fresh rounds of *direct* and *indirect* effects. Thus, an initial positive shift in *A* can set an economy on to a virtuous cycle of mutually reinforcing positive effects.

In view of its both theoretical and empirical potency, it is unfortunate that growth theories do not have much to say about determinants of *A*. This gap in our understanding of growth is now prompting researchers to call for the development of a TFP theory. In offering their ideas in this regard, different researchers have emphasized different aspects of TFP, something not surprising given TFP's broad scope.[4] In order to determine which of these different aspects to focus on in building a TFP theory, it is helpful to have some stylized facts regarding the determinants of TFP. In view of these circumstances, empirical research on TFP may have to precede theoretical research on it at the current stage.

Empirical papers cited earlier have made some progress in developing TFP empirics. However, most of these studies are focused on particular TFP determinants, and hence do not reveal the overall picture.[5] This paper therefore adopts a comprehensive approach and strives to examine the productivity determinants under a unified framework. The analysis of this paper is therefore *extensive* rather than *intensive*, and the picture portrayed is in broad strokes. However,

---

[4] As Easterly and Levine (2001, p. 181) note, "The empirical importance of TFP has motivated economists to develop models of TFP. These focus variously on technological change (Aghion and Howitt 1998; Grossman and Helpman 1991; Romer 1990); impediments to adopting new technologies (Parente and Prescott 1996); externalities (Romer 1986, Lucas 1988); sectoral development (Kongsamut, Rebelo, and Xie 2000); or cost reductions (Harberger 1998)."

[5] As Easterly and Levine (2001, p. 178) note, "Evidence that confidently assesses how well these conceptions of TFP explain economic growth is lacking. Economists need to provide much more shape and substance to the amorphous term TFP, distinguishing empirically among these different theories." Further, Easterly and Levine note, "…empirical work does not yet decisively distinguish among the different theoretical conceptions of TFP growth." (ibid)

once the broad picture has emerged, and the methodology has been agreed upon, future research can focus more intensively on the individual patches.

3. The two-stage approach to productivity study

In dealing with productivity differences across countries, empirical researchers have generally followed two methodological approaches, namely (a) the 'extended specification' (of growth regressions) methodology and (b) the 'two-stage' methodology. The difference between these two approaches can best be explained by invoking the following neoclassical growth-convergence equation that has now become ubiquitous in the empirical growth literature:[6]

(4) $$\ln y_{t_2} - \ln y_{t_1} = (1-e^{-\lambda\tau})\frac{\alpha}{1-\alpha}\ln s_{t_1} - (1-e^{\lambda\tau})\frac{\alpha}{1-\alpha}\ln(n_{t_1}+g+\delta) - (1-e^{-\lambda\tau})\ln y_{t_1} + (1-e^{-\lambda\tau})\ln A(0) + g(t_2 - e^{-\lambda\tau}t_1),$$

where y is per capita income, t_1 denotes the initial period, t_2 denotes the subsequent period, $\tau = (t_2 - t_1)$, and $\lambda = (n+g+\delta)(1-\alpha)$, which is generally termed as the 'rate of convergence.' The rest of the notations are as before. Denoting

(5) $$\gamma = (1-e^{-\lambda\tau}),$$

(6) $$\beta_1 = (1-e^{-\lambda\tau})\frac{\alpha}{1-\alpha},$$

(7) $$\beta_2 = -(1-e^{-\lambda\tau})\frac{\alpha}{1-\alpha}, \text{ and}$$

(8) $$\eta_t = g(t_2 - e^{-\lambda\tau}t_1),$$

equation (5) may be alternatively expressed as

(9) $$\ln y_{t_2} - \ln y_{t_1} = \beta_1 \ln s_{t_1} - \beta_2 \ln(n_{t_1}+g+\delta) - \gamma \ln y_{t_1} + \gamma \ln A(0) + \eta_t.$$

Equation (9) presents the conventional 'growth-initial level' regression, and it clearly shows that the productivity term, *A(0)*, is a part of this equation. The difference between the 'extended specification' and 'two-stage' methodologies arises mainly from the different ways in which these try to deal with the productivity term, *A(0)*.

[6] The derivation of this equation is available in several papers including Mankiw, Romer, and Weil (1992), Barro and Sala-i-Martin (1992), and Islam (2003)

The 'extended specification' methodology

Under the 'extended specification' approach, variables that are assumed to proxy for $A(0)$ are added directly to the right hand side of equation (9). In the recent period, this approach begins with Barro (1991), where he includes many novel (from the viewpoint of mainstream growth economics at the time) variables, such as the 'number of political assassinations and coups' etc. Barro (1991) does not provide 'theoretical' arguments for such inclusion.[7] However, in a subsequent paper, Barro (1997) explains how in growth regressions it is necessary to control for cross-country differences in the steady-state income level y^*. To the extent that the productivity shift term A is a part of y^*, this provides a theoretical justification for inclusion of variables proxying for A in growth regressions. Other studies that have followed the 'extended specification' route include Caseli et al. (1996), Sachs and Warner (1997), Sala-i-Martin (1997, 1999), and Barro and Sala-i-Martin (1995, 2004).

There are however several problems with the 'extended specification' methodology. The first is that this approach leads to sets of explanatory variables of very heterogeneous character, so that 'conjectural' variables such as the 'number of coups' attain the same status as 'strictly model-based' variables such as the investment rate. It is this situation that Solow (2001, p. 285) found 'haphazard' and that prompted him to argue for inclusion of the 'conjectural' variables only in regressions that have productivity as the dependent variable.[8] Second, the extended regression approach quickly leads to a disconnection between the underlying growth model and the regression specification, causing researchers often to lose sight of the values of the growth model parameters implied by the estimated coefficients. Third, the 'extended specification' approach cannot focus on productivity differences, which are regarded rather as a 'nuisance' to be somehow pushed aside in order to estimate such input accumulation parameters as the elasticity of output with respect to physical or human capital. Accordingly, the extended regression approach, in most cases, does not even attempt to produce productivity estimates. Fourth, despite extension of the specification, it is usually not possible to include in a growth-

[7] He just mentions that "I interpret the REV and ASSASS variables as adverse influences on property rights, and thereby as negative influences on investment and growth." (Barro 1991, p. 432) REV and ASSASS refer to the number of revolutions and political assassinations, respectively.

[8] Noting all the 'exotic' variables that have appeared in growth regressions of extended specifications, Solow (2001, p. 285) observes that "if they mean anything at all, those many RHS variables in growth regressions are determinants of TFP. But then they should be selected with that function in mind, and TFP (or its growth rate) should be the LHS variable."

regression all the variables that influence productivity. Yet these left out variables impart omitted variable bias (OVB) to the estimated accumulation parameters because of their correlation with the included (accumulation) variables.[9] In view of these problems, many researchers are now resorting to the 'two-stage' approach.

The two-stage approach

The two-stage approach to the study of productivity differences accords well with Solow's suggestion of keeping the basic growth regression uncluttered and of including 'conjectural' variables only in regressions that have 'productivity' as the dependent variable. Implementation of Solow's suggestion however requires a two-stage exercise, the first of which is devoted to estimation of productivity, and the second stage focuses on analyzing the productivity estimates obtained in the first stage.

The first stage of the exercise can, in turn, be conducted in different ways. Hall and Jones (1999), for example, conduct a cross-section growth accounting exercise to compute values of $A(t)$. Islam (1995), on the other hand, adopts the panel regression approach, under which the equation (9) is reformulated as a dynamic panel data model with $A(0)$ as the 'individual country effect.' The estimated values of these 'country effects' are then used to recover the estimated values of $A(0)$. Some researchers have even used residuals from the cross-section growth regression (of appropriate specification) as the first stage estimates of TFP.[10]

The two-stage methodology has some advantages. First and foremost is of course that it makes it possible to retain the essential distinction between the basic variables of growth and the conjectural variables that are thought to proxy for productivity. Second, this methodology brings productivity to the forefront of the analysis rather than relegating them to a 'nuisance' status. Third, paradoxically the 'two-stage' methodology allows getting better estimates of the

[9] For example, Sachs and Warner observe that "An important step in estimating (2) (the growth convergence equation –ni) is to include a reasonably comprehensive set of exogenous variables in *Z*. Many empirical studies suffer from the fact that the authors include only a small subset of appropriate variables. For example, if the author is studying the effects of income inequality on growth, then only a measure of income inequality is included in (2). Without a comprehensive set of Z variables, cross-country growth studies are plagued by left-out-variable errors of great importance." See Islam (2003) for more discussion of this issue.

[10] Young (1994) provides a recent example of this particular way of obtaining first stage TFP estimates. Earlier, Feder (1975) also used this route.

accumulation parameters. For example, the panel regression approach makes it possible to obtain these parameters' estimates that are free from omitted variable bias and endogeneity bias.[11]

In view of these advantages, more researchers are gravitating to the two-stage approach with time. This paper represents a further step in developing and applying the two-stage approach.[12]

4. Determinants of Productivity

By now, numerous variables have been suggested as productivity determinants. Introspection suggests that these may be classified into the following four types.

(A) *Economic factors*, such as 'openness of an economy,' 'size of the government consumption expenditure,' 'extent of distortions in the economy,' etc.

(B) *Institutions*, such as the 'nature of the legal system,' 'nature of the political system,' etc.

(C) *Social base*, such as the 'ethnic and religious composition of the population,' 'colonial heritage,' etc.

(D) *Physical base*, such as the 'physical location,' 'climate,' 'disease environment,' etc.[13]

Not all researchers follow the same classification of productivity determinants. For example, Rodrik et al. (2004) (henceforth RST) classify productivity determinants into three groups,

---

[11] Apart from being drawn by these advantages, some researchers are arriving at the two-stage methodology via their analysis of the role of human capital in growth. As is known, while initial human capital *level* proves significant in cross-country growth regressions, *change* in human capital generally does not. Krueger and Lindahl (2001) point to measurement errors in human capital data as the main reason for this anomalous result. The two-stage methodology can help get around this measurement error problem. To the extent that the second stage of the two-stage analysis can be of cross-sectional nature, human capital variable can be examined fruitfully in this framework, because the cross-section dimension of the human capital data has more signal-to-noise ratio than its time series dimension does. There are other ways in which researchers are arriving at the two-stage approach. For example, Canova and Marcet (1995, p. iv) propose a two stage approach, the first of which is devoted to estimation of the steady states. "Once the steady state estimates are obtained for each unit," the authors suggest, "we can test, in a second step, what variables determine the cross-sectional distribution of steady states."

[12] The two-stage approach has some problems too. First, the task of separating productivity from input accumulation is fraught with many pitfalls. For example, Abramovitz (1956), Abramovitz and David (1973), Nelson (1969), Wolff (1991), and others have emphasized the complex interaction between productivity and input accumulation and have emphasized the difficulty of separating the two. In the more recent period, Hulten (1979, 2001) has emphasized capital accumulation induced by productivity increases. In implementing the 'two-stage' methodology, it will be necessary to take note of this interaction. For a fuller comparison between the 'extended specification' and the 'two-stage' methodology, see Islam (2003).

[13] It may be noted that, three of the four types above are 'non-economic' in nature. By putting forward these variables as determinants of economic performance, growth economists are therefore conceding an important role to non-economic factors. This should help growth economics and development economics come closer, because

namely ‘Policies,’ ‘Institutions,’ and ‘Geography.’ This agrees with the classification above to a great extent. However there are two important differences. First, the RST classification does not allow a formal place for ‘Social Base’ variables.[14] Yet recent research shows that Social Base variables play an important role in a country’s economic performance. For example, Easterly (1997) finds that ethnic division is an important factor behind Africa’s poor growth performance. Alesina et al. (2003) offer similar conclusion based on a more comprehensive analysis. Hence neglect of Social Base as a determinant of economic performance may not be satisfactory. Second, by ‘Policies’ RST mainly refer to *economic policies*, such as reduction of tariff rates, etc. However, policies can be *non-economic* too, such as reform of the political system. The classification proposed in this paper therefore allows the possibility of both types of policies, with economic policies influencing economic determinants of productivity and non-economic policies influencing non-economic determinants of productivity. Paradoxically, such a treatment accords well with RST’s own suggestion that ‘institutions’ are ‘stock’ variables, while ‘policies’ are ‘flow’ variables, and that ‘policy’ variables should therefore be left out of regressions that take stock variables, such per capita income level, as the dependent variable.[15] To the extent that this paper’s dependent variable, productivity level, is a stock variable, it is rather appropriate that policies be left out of the specification.

The classification of productivity determinants also brings up the issue of interrelationship and chain of causality among them. Researchers initially did not pay much attention to these issues and instead included different variables in extended regressions in a rather kitchen sink style. However, they soon started to worry about the ‘hierarchy’ among these variables and about different channels of their influence. Different researchers have presented different views in this

---

emphasis on non-economic determinants (such as institutions) by development economics was historically one of the main reasons why the two fields drifted apart.

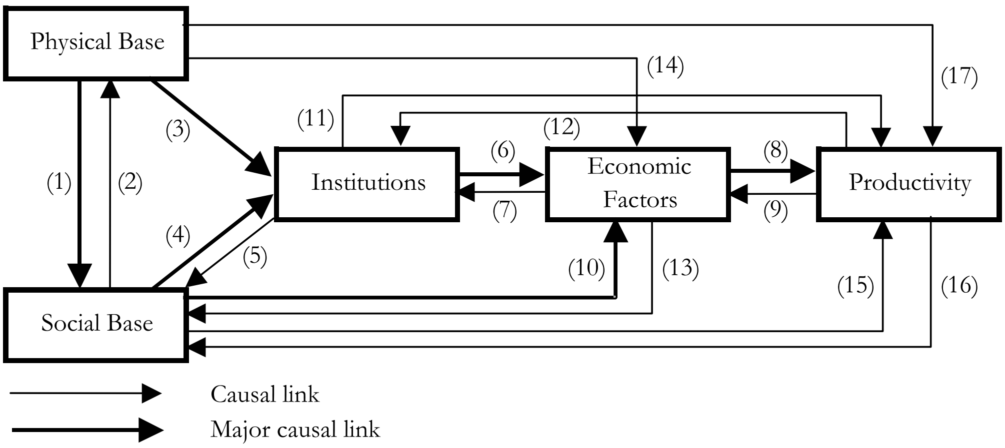
[14] It should be noted that even though RST leave out Social Base variables from their main specification, they use some of these variables as *additional* controls for robustness check of specifications comprised of other variables. For example, in regressions presented in their Table 4, RST include as additional controls such variables as ‘legal origin,’ ‘identity of colonizer,’ and ‘religion.’

[15] However, RST’s treatment of the division between ‘policy’ and ‘institution’ is not entirely clear. First, contrary to their argumentation, RST do include ‘policies’ in their regressions even though their dependent variable is per capita income *level*. Second, it is unclear why such variables as ‘extent of openness’ should not be viewed as a ‘stock,’ rather than a ‘flow’ (such as ‘reduction in tariff’). Third, in their algebraic formulation of this relationship, RST treat institutions as just a single index, while allowing for many ‘dimensions’ of ‘policies.’ However, ‘institutions’ have many dimensions/components too, and once this is recognized, it becomes difficult to postulate an exact correspondence between different dimensions of ‘institutions’ on the one hand and different dimensions of ‘policies’ on the other. Take for example ‘legal system.’ If this is stock, what is the corresponding flow? The same question may be posed for other components of institutions.

regard.[16] Figure 1 presents a scheme that subsumes many of the causality schemes proposed in the literature. Each arrow in Figure 1 represents a causal *link*, and these links have been numbered in order to facilitate the discussion below. A combination of causal links is referred to as a (causal) *channel*. An important consideration in deciding about the direction of causality is the time-horizon in mind.

Figure 1

Causality among Productivity Determinants



Physical Base: In the long sweep of history, even national boundaries, and hence the Physical Base of a country, are endogenous. However, the recent empirical growth research focuses on the post World War-II decades. For this rather short period, the Physical Base of an economy may indeed be deemed exogenous. What is however contentious is whether Physical Base variables exert their influence on productivity directly (as represented by link (17)) or through their influence on other, 'intermediate' productivity determinants only, as represented by channels (3)+(6)+(8), (14)+(8), or (3)+(11), etc. Physical Base variables are sometimes also called Geography variables, and recent research has seen some spirited debate about their role. As noted above, Acemoglu et al. (2001) argue that while Geography had a historic role in determining institutions, it is the latter that have the primary role in explaining economic performance. This agrees with channel (3)+(6)+(8) or (3)+(11). By contrast, other researchers, such as Sachs (2003) and Sachs and McArthur (2001), argue that Geography has an important

[16] One of the first studies to emphasize the issue of 'hierarchy' among productivity determinants is Hall and Jones (1999), who argue that the underlying determinant of economic performance is 'social infrastructure.'

direct influence in addition to their indirect influence via institutions. This view therefore emphasizes the link (17).

Social Base: Given the time horizon mentioned above, many of the Social Base variables may also be deemed exogenous.[17] For example, the ethnic and religious composition of most of the developing countries was determined by the way their political boundaries were demarcated at the time of their independence from colonial rule. The Physical Base and Social Base variables of these countries may therefore be thought to have been determined jointly, as reflected by the inclusion of link (2) in addition to (1) in Figure 1. Some researchers suggest that Social Base variables (at least some of them) can also be endogenous. Alesina et al. (2003), for example, note that the political atmosphere, which may in turn depend on the productivity performance of a country, can lead people to rethink about their ethnic or religious identities.[18] Such possibilities do exist, and channels (9)+(7)+(5), and (9)+(13) of Figure 1 recognize them.

Institutions: The literature emphasizing the role of institutions in the growth and productivity performance of a country is now quite large.[19] Among recent papers emphasizing the role of institutions are Acemoglu et al. (2001) and Rodrik et al. (2004), as already mentioned.[20] It has been generally thought that good institutions help a country adopt and implement good economic policies, thus enabling it to have good economic determinants of productivity. Figure 1 reflects this chain of causality through the channel (6)+(8). In addition, institutions may be thought to have some direct influence on productivity, as reflected by the link (11). However, unlike the Physical Base and Social Base variables of a country, Institutions are very likely to be endogenous, because higher levels of productivity and income can conduce to better institution

[17] Mauro (1995) also notes the exogeneity of Social Base variables for the analysis of post-war growth and productivity performance.

[18] See also Acemoglu et al. (2001) on this point.

[19] Apart from the ones mentioned earlier, other papers emphasizing the role of institutions include Knack and Keefer (1995, 1997), Keefer and Knack (1997), Temple (1999), Temple and Johnson (1998), Easterly and Levine (1997), Engerman and Sokoloff (2000), Landes (1998), and Rodrik (1999). Earlier works emphasizing the role of institutions include North and Thomas (1973), Jones (1981), and North (1981, 1990).

[20] Rodrik at al. (2004) however emphasize the mutability of institutions. Their paper (pp. 153-5) also makes the point that Acemoglu et al. (2001) sometimes lapse into reading theory in an instrument. This refers to the fact that Acemoglu et al. (2001) use early European settler mortality rate as the instrument for the institutional quality variable, RISK. However, their discussion sometimes tends to suggest that colonial past itself has been the determinant of subsequent economic performance. Rodrik et al. (2004) refute this (mis)interpretation of Acemoglu et al.'s (2001) findings by restricting the sample to countries that were never colonized and showing that institutional quality proves equally important for this subset of countries too. As Rodrik et al. (2004, p. 154) put it: "…if colonial experience were the key determinant of income levels, how would we account for the variation in incomes among countries that had never been colonized by the Europeans?"

building in a country. Link (12) and channel (9)+(7) capture these feedback processes.[21] Endogeneity is therefore an important issue in identifying the role of institutions.

Economic factors: Economic Factors are the most immediate determinant of productivity. This is represented by the relationship (8). However, Economic determinants are also endogenous. Higher levels of productivity can help a country have better magnitudes of Economic Factors, either directly or through their positive influence on institutions and the Social Base. These feedback processes are captured by the link (9) and through the channels (12)+(6) and (16)+(4)+(6).

Some economic determinants of productivity such as the 'size of the government' or 'the degree of openness of an economy' are rather straightforward in nature. However, economic determinants also include in their scope possible externalities emanating from input accumulation. In Section 2, we alluded to this possibility as the 'feedback effect,' and expressed it through the relationship $A = A\ (s, \ldots)$, with s denoting the rate of investment, which can be in either physical or human capital. In order to capture these potential externalities, physical and human capital can appear as explanatory variables of the productivity regressions even though these were already taken into account in the computation of productivity. The following general production function framework can help explain the pertinent issues:

$$(10) \qquad Y = F[A(K, H, \ldots), K, H, L]$$

The first stage of the exercise quantifies A, and in doing so it takes into account of K, H, and L, which are independent arguments of the production function $F[.]$. However, one has to confront K again in the second stage in order to analyze $A(K, H, \ldots)$. As this expression suggests, both K and H can be viewed as arguments of A, recognizing the possibility of externality arising from accumulation of both physical and human capital. As is known, many NGT models have emphasized H as an argument of A.[22] The role of H can also be argued from the viewpoint of technological diffusion across countries. According to this view, human capital helps a country benefit more from the 'advantage of backwardness,' thus yielding a higher value of A. However, some researchers have suggested that externalities can emanate from physical capital

[21]The channel represented by (16)+(4) can be another route, albeit very remote and weak, of such feedback influence

[22] The first generation new growth models, such as of Romer (1986) and Lucas (1988), can be interpreted to be of this spirit.

accumulation too.[23] In a series of articles, DeLong and Summers claim to have found empirical evidence for this proposition.

Barro (2001) works out the implications of these externality propositions for the conventionally computed TFP. He shows that proceeding from the production function:

(11) $$Y_i = AK_i^{\alpha} K^{\beta} L_i^{1-\alpha},$$

where Y_i, K_i, and L_i are output, capital, and labor of the individual producer *i*, while *K* is the aggregate of capital in the economy, it is possible to derive

(12) $$\frac{\dot{Y}}{Y} - \alpha\frac{\dot{K}}{K} - (1-\alpha)\frac{\dot{L}}{L} = \frac{\dot{A}}{A} + \beta\frac{\dot{K}}{K}.$$

This equation shows that the conventionally computed TFP growth rate (the left hand side of equation (12)) comprises in part of the aggregate capital accumulation rate. Barro (2001) further shows that equations for TFP growth similar to (12) can be derived proceeding from production functions used by second generation new growth models too. The standard TFP growth rate is again found to be related with input accumulation rate expressed in the form of rate of either expansion of the number of inputs or rise along the quality rung, depending on the production function chosen. In short, externality or feedback effect of input accumulation is an important factor to be considered among the Economic Determinants of productivity.

Having discussed the classification of productivity determinants and the interrelationship among them, we can now move to the discussion of the empirical results.

5. Empirical results

In implementing the two-stage methodology, we first note that several previous studies have already presented productivity estimates for large samples of countries. As noted earlier, Hall and Jones (1996) present relative productivity *levels* for a sample of 118 countries for 1988. Similarly, Islam (1995) presents estimates of relative productivity *levels* for a sample of 96 countries for the period of 1960-85. It is therefore possible to use these productivity estimates and skip redoing the first stage of the two-stage methodology. In fact there are some advantages

[23] Arrow (1963) can be viewed in this light.

in using these productivity estimates, because these have already undergone the profession's scrutiny and have been available to other researchers for use.

Endogeneity and identification strategies: From the discussion above, it is clear that many of the productivity determinants, particularly the Economic Factors and Institutions, are endogenous, requiring instruments for identification of their coefficients. However, finding satisfactory instruments is not easy, as is often the case. In dealing with similar endogeneity problems of his growth regressions, Barro (1997) advocates the use of pre-determined variables as instruments. The validity of this approach depends in part on the absence of serial correlation in the error term. Based on residuals from several consecutive decade specific growth regressions, Barro concludes that any such serial correlation is very low, validating the instruments.[24] Following Barro, this paper also makes use of the 'instrumentation through predetermination' approach. However, the paper in addition uses a number of other instrumental variables that have recently surfaced in the literature. Details about instruments, over-identification tests, and other technical aspects of estimation are provided at appropriate points in the paper.

In the following, we refer for brevity to productivity indices of Hall and Jones (1996) and Islam (1995) as HJ and IS indices, and the regressions with them as dependent variable as HJ and IS regressions, respectively. The results obtained from these two sets of regressions proved to be similar. In the paper we present mainly the results from HJ regressions, alluding to IS regressions only when necessary. One reason for preferring HJ regressions is that these allow for more lagged variables to serve as instruments and thus for identification tests to be conducted.[25] Also, the sample size of HJ regressions is in general larger.

The basic results from the HJ regressions are presented in Tables 1a, 1b, and 1c. Following the sequential approach, we begin developing the model first with the Economic Determinants, and then move progressively to more 'underlying' determinants of productivity. Such a progression follows from the causality scheme presented in Figure 1, and it helps to find out whether the influence of underlying determinants is entirely mediated through intermediate determinants or they have independent, direct influence too.

[24] Barro and Sala-i-Martin (2004) also uses this identification strategy.
[25] Recall that the HJ index pertains to 1988, whereas the IS index pertains to 1960-85 period as a whole.

5.1 Economic determinants of productivity

Externality from input accumulation

We begin the analysis by examining the potential role of externality arising from input accumulation. Column (1) of Table 1a presents the results from regressing the HJ productivity index on human capital level, as measured by the Barro-Lee (1994) variable (denoted by *TYR15*), that gives average total (i.e., inclusive of primary, secondary, and tertiary) years of education in the population of age 15 and above. We use the 1985 value of *TYR15* as the explanatory variable, allowing thus a lag in the influence of human capital on productivity. Data availability is another reason for using the 1985 value. The lag also makes the right hand side variable pre-determined with respect to the dependent variable and thus help alleviate the problem of endogeneity. Thus under standard assumptions regarding the error term, even OLS estimation of the model would be valid. In fact, in many cases we present OLS estimates. This is because of two reasons. First, the results often do not differ that much when estimated using instrumental variables.[26] Second, the IV results differ among themselves to some extent depending on how many and which particular lagged values are used as instruments. This causes some indeterminacy as to which particular IV results are to present. The OLS results by contrast do not suffer from this ambiguity.

The results show a significant positive coefficient, indicating that the direct impact of H may not exhaust its influence on output, and it may have an externality effect, via its positive influence on the total factor productivity. Column (1) shows the OLS results. The model was also estimated with further lagged values of *TYR15* as instruments. The estimated value of the human capital coefficient remains remarkably stable, showing only slight sensitivity with respect to the choice of instrument set.[27] Though not presented in the paper, similar strong association between human capital and productivity is found from IS-regressions too.[28]

---

[26] One reason why instrumentation does not markedly change the OLS results is the fact that the explanatory variable itself is lagged.

[27] The *p*-value of the over-identification test, when all the lagged values are used as instrument, proves to be 0.16. The usefulness of the test is however somewhat moot given the fact that it pertains to the joint hypothesis of validity of the instruments as well as validity of the specification.

[28] The numerical magnitude of the human capital coefficient is however much higher in the IS regression than it is in the HJ regression. This is not unexpected. If H of equation (2) is specified as $e^{\phi(E)}L$, as Hall and Jones (1996) do, with E being the average years of education for L, the relationship between IS index (A^{IS}) and HJ index (A^{HJ}), obtained on the basis of production functions given in equation (1) and (2), respectively, can be shown to be as follows:

$$\ln A^{IS} = \ln A^{HJ} + \phi(E)\,.$$

To check whether similar externality is true of physical capital accumulation, we next include per capita physical capital stock in the regression. The capital stock values are computed using perpetual inventory method based on the investment data provided in Penn World Tables. Again, for reasons mentioned earlier, the capital stock value of 1985 is used as the explanatory variable. The OLS results are in column (2) of Table 1a. We see that the physical capital variable proves significant too. The model is also estimated using further lagged values as instruments.[29] The results prove similar to those from OLS.

This exploration suggests that productivity may benefit from externality arising from both human and physical capital. As of now, therefore, the aggregate productivity model may have the following tentative specification:

(13) $$A = f(H, K\ldots).$$

These results however may not mean much as they are obtained from simple bi-variate and almost bi-variate regressions, which obviously suffer from considerable omitted variable bias. Pursuing the sequential approach we therefore proceed to include other variables in the regression, taking equation (13) as the maintained model.

Size of the government

One of the economic determinants of productivity that has been widely mentioned in the literature is the size of the government, as measured by government expenditure as a ratio of GDP.[30] Government expenditure in turn may be classified broadly into two types, namely consumption and investment. Conventional wisdom suggests that government investment expenditure should be conducive to productivity, while government consumption expenditure, being a drain on potential investment resources, is harmful to growth and productivity. The literature, both theoretical and empirical, however remains controversial on the issue.

This shows that the IS index contains a human capital component even if human capital does not affect aggregate productivity via externality. On the other hand, human capital can prove significant in the HJ regression only if human capital influences aggregate productivity via externality, in a similar way as expressed by equation (12). This explains why the H coefficient in the IS regression proves much higher in numerical magnitude than it is in the HJ regression.

[29] The p-value of the over-identification test, when all the lagged values are used as instrument, proves to be 0.20.

[30] Note that the size of the labor force employed by the government relative to the total employment can be another indicator/measure of the government size. However, the literature has mainly used the government size as measured by government expenditure relative to the GDP.

Barro (1991) introduced to the empirical study of growth several government expenditure variables that have since proved popular in research. One of these variables is *GVXDXE*, which stands for the ratio to GDP of 'government consumption expenditure minus expenditure on defense and the recurrent expenditure on education.'[31] This gives a measure of government expenditure on administrative and other regulatory purposes. It may therefore give an indirect measure of the extent of government intervention in the functioning of the economy. Depending on their type and how they are carried out, government interventions can be beneficial for the economic performance. However, generally government intervention has been thought to introduce distortions in the economy and thereby hamper its productivity. To find out an empirical answer to this question, we next include *GVXDXE* as an additional explanatory variable in the HJ regression. We use the 1985 value of GVXDE as the explanatory variable, for reasons explained earlier. Column (3) of Table 1a presents the results. As we can see, there is a significant negative association between productivity and *GVXDXE*. The outcome does not change when the model is estimated with further lagged values as instruments.

To inquire into the role of government investment expenditure in productivity, we explore with two Barro-Lee (1994) variables, namely *INVPUB*, which represents the 'ratio of nominal public domestic investment (fixed capital formation) to nominal GDP,' and *GGCFD*, which represents the 'ratio of real public domestic investment to real GDP.' Columns (4) and (5) show the results from inclusion of these variables in similar manner as was the case with earlier variables. We see that neither *INVPUB* nor *CCCFD* prove significant, suggesting that government investment expenditure does not have particularly beneficial effect on a country's productivity. To some extent this may reflect the inefficiency with which public investment projects are often undertaken.

In view of the results above, we extend the productivity model to include only *GVXDXE*, so that we now have

$$A = f(H, K, GVXDXE, \ldots\ldots). \qquad (13)$$

[31] A question may be raised regarding the appropriateness of leaving out defense expenditure in constructing the government consumption variable. The likely argument for leaving out is that expenditure on defense helps fortify property protection, which in turn should be conducive to growth. Experience of many developing countries however does not bear out this argument. This is an issue that can be investigated further. However, we do not want to push this point any further in this paper.

Openness of the economy

Another economic determinant of productivity that has received wide attention in the literature is the 'openness' of an economy. It has been generally thought that 'openness' is conducive to growth and productivity, though a full-blown theoretical model showing such a relationship has been lacking.[32] Despite this theoretical lacking, many have claimed that empirical research confirms a positive relationship between openness and economic growth. Sachs and Warner (1995) for example have made a strong empirical case for openness. However, other researchers, such as Rodriguez and Rodrik (2000), have taken a skeptical view of the role of openness.

There are two major problems regarding empirical identification of the role of openness. The first concerns measurement of "openness" itself. It is difficult to get a satisfactory measure or indicator of openness. The second concerns finding satisfactory instruments for whatever openness variable is agreed upon, as because it is surely to be endogenous.

In view of these problems, we use in this paper two different measures of openness. The first is the most widely used variable, *OPEN*, which is the ratio of imports plus exports to GDP. The data on *OPEN* is readily available either in the Penn World Tables or in Barro-Lee (1994). The variable *OPEN* however does not prove significant in the model. In the OLS regression, *OPEN* (pertaining to 1985) proves insignificant. The situation does not change when it is instrumented using further lagged values.

With regard to 'openness,' it is possible to use other instruments than just lagged values. For example, Frankel and Romer (1999) present an instrument for openness that is constructed entirely on the basis of exogenous (geographical) characteristics of a country. However, the use of this instrument, *LOGFRANKROM*, does not change the results, as shown in column (6) of Table 1a. Alcala and Ciccone (2004) make the point that instead of taking the ratio of nominal values of export plus import to the nominal GDP as the measure of openness, it is necessary to take the ratio of the corresponding *real* values. Responding to the criticism, Rodrik et al. (2004) compute a real counterpart of *OPEN* naming the logarithm of it as *LNOPEN*, and also a nominal counterpart of the Franker and Romer instrument, naming it *LOGFRANKROMR*. However, the use of these openness variables and instruments defined in real terms does not change the results.

[32] See Islam (2004) for a fuller discussion of this issue.