Is the Relationship Between Aid and Economic Growth Nonlinear?*

Andros Kourtellos
University of Cyprus

Chih Ming Tan
Tufts University

Xiaobo Zhang
IFPRI

September 22, 2006

Abstract

In this paper, we investigate the relationship between foreign aid and growth using recently developed sample splitting methods that allow us to uncover evidence for the existence of heterogeneity and nonlinearity simultaneously. We also implement a new methodology that allows us to deal with model uncertainty in the context of these methods. We find some evidence that aid may have heterogeneous effects on growth across two growth regimes defined by ethnic fractionalization. In particular, countries that belong to a growth regime characterized by levels of ethnic fractionalization above a threshold value experience a negative partial relationship between aid and growth, while those in the regime with ethnic fractionalization below the threshold experience no growth effects from aid at all. Nevertheless, there exists substantial model uncertainty so that attempts to pin down the typology of these growth regimes as being decisively characterized by ethnic fractionalization remain inconclusive. When we account for model uncertainty, we find no evidence to suggest that the relationship between aid and growth is nonlinear. Overall, our results suggest that the partial effect of aid on growth is very likely to be negative although we cannot reject the hypothesis that aid has no effect on growth. In this sense, our findings suggest that aid is potentially counterproductive to growth with outcomes not meeting the expectations of donors.

*Kourtellos: Department of Economics, University of Cyprus, P.O. Box 20537, CY-1678 Nicosia, Cyprus. Email: andros@ucy.ac.cy. Tan: Department of Economics, Braker Hall, Tufts University, 8 Upper Campus Road, Medford, MA 02155. Email: chih-ming.tan@tufts.edu. Zhang: International Food Policy Research Institute (IFPRI), 2033 K Street, NW Washington, DC 20006-1002. Email: X.Zhang@cgiar.org. We are deeply grateful to Rob McCulloch for his invaluable advice on the implementation of the Bayesian tree methods used in this paper. We thank the USAID/International Food Policy Research Institute (IFPRI) for research support. Charalambos Michael and Ioanna Stylianou provided excellent research assistance.

1 Introduction

One of the most controversial debates in the empirical growth literature with big policy implications is whether foreign aid is beneficial to a country's economic growth. In an influential paper, Burnside and Dollar (2000) examine the effect of aid, as measured by the ratio of the sum of grants and the grant equivalents of official loans in constant prices to real GDP[1], on growth. Using standard cross-country panel growth regressions that include an interaction term of aid with a policy index, they find that aid has a positive impact on growth in developing countries as long as these countries have sound macroeconomic policies. The policy implication of this finding was straightforward. Policy makers at international aid agencies could now argue that development assistance can contribute to poverty reduction in countries with good policy environments.

On the other hand this finding has sparked an industry of mainly empirical papers trying to examine the sensitivity of Burnside and Dollar's results to model specification, alternative sets of included/excluded variables, and different data series. Some of the most notable papers include Guillaumont and Chauvet (2001), Hansen and Tarp (2001), Collier and Dehn (2001), Collier and Dollar (2002, 2004), Collier and Hoeffler (2004), Easterly (2003), Easterly, Levine, and Roodman (2004), Dalgaard, Hansen and Tarp (2004), Roodman (2004), and Rajan and Subramanian (2005a,b). Some of these papers confirm the main finding of Burnside and Dollar; i.e., that aid is effective only in countries with good policies, while others find the results fragile to the addition of particular variables.

One problem that the literature on aid and growth has been dealing with is the problem of how to model heterogeneity and/or nonlinearities in growth analyses. Typically, what has been done is to treat this issue in an ad hoc way by including squares and interaction terms for aid, policy, and other growth variables. The unsystematic, ad hoc nature as to how specific choices are made over which nonlinearities/heterogeneity to include and which to leave out, however, leaves much to be desired. For instance, there is no good theoretical or statistical reason for only including an interaction term between aid and policy and not the square of aid or even both in the model. Why not also include an interaction term between policy and institutions? In fact, several new growth theories such as Azariadis and Drazen (1990) and Howitt and Mayer-Foulkes (2002) suggest that the cross-country growth process is highly nonlinear.

To make things worse, as suggested by Brock and Durlauf (2001), new growth theories are inherently open-ended. By theory open-endedness, Brock and Durlauf refer to the fact that typically the a priori statement that a particular theory of growth is relevant does not preclude other theories of growth from also being relevant. Growth models typically do not provide much guidance as to the exact specification in which growth determinants should enter the growth equation[2] as well. Brock and Durlauf point out that taken together, the com-

[1]This is known as Effective Development Assistance.

[2]Nor which particular proxy variables best represent the theoretical concepts under consideration.

bination of theory and specification uncertainty (what they refer to collectively as model uncertainty) potentially renders coefficient estimates of interest to be "fragile". The potential fragility of coefficient estimates under model uncertainty is important because it implies that findings on the relationship between aid and growth, which do not properly account for model uncertainty, may be non-robust. For instance, the finding of a nonlinear relationship between aid and growth may, in fact, be just a manifestation of some other unaccounted misspecification due to omitted variables or even due to unaccounted heterogeneity and/or nonlinearities with respect to other growth determinants. Our point is that strong a priori assumptions on the appropriate specification of growth determinants and functional form of the model are hard to justify.

Nevertheless, while there is little agreement over the exact nature of nonlinearities and heterogeneity in the growth literature, there is a growing consensus that, given that we think such nonlinearities/heterogeneity exists, they may potentially be fruitfully modeled using empirical tools that emphasize pattern recognition (see Durlauf (2003)). Sample splitting and threshold regression methods and their derivatives are important constituents of such tools. For instance, Durlauf and Johnson (1995) employed a classification and regression tree method (CART; see Breiman, Friedman, Olsen, and Stone (1984)) to sort countries based on initial per capita income and initial literacy rates. They interpret their findings as evidence in favor of the theory of poverty traps of Azariadis and Drazen (1992).

In this paper we employ recently developed sample splitting methods to systematically uncover the robust relationship between aid and growth. Sample splitting methods such as threshold regression and regression trees allow for increased flexibility in functional form and at the same time are not as susceptible to curse of dimensionality problems as nonparametric methods. Unlike parametric models with polynomial terms (squares, interactions, etc), sample splitting methods are parsimonious. More importantly, these methods are structurally interpretable as they endogenously sort the data, on the basis of some threshold determinants, into groups of countries each of which obeys the same model (i.e., multiple growth regimes). Other notable applications of sample splitting methods in growth include Tan (2006) who use an improved regression tree algorithm to CART (GUIDE; see Loh (2002)) and Masanjala and Papageorgiou (2004) who employ threshold regression (TR; see Hansen (1996, 2000) and Gonzalo and Pitarakis (2002)).

One major problem associated with the sample splitting methods that have been employed so far in the literature, however, is the sequential nature of the splitting process. By this we mean that choices of threshold variables and split values made in initial sample splits are never revised as the number of splits increases. Hence, any mistake made at the earlier stages of the process is propagated to the splits below. The result is that the classification of observations into regimes can be unstable. Small changes in the data result in large changes to the threshold or "tree" structure (see Hastie, Tibshirani, and Friedman (2001) and also Hong, Wang, and Zhang (2005)).

To be clear this is not an issue of statistical inference but rather it has to do with the qualitative nature of threshold variables. It is one thing to say that a 95% confidence interval for a (real-valued) parameter is $[0.3, 0.8]$ and quite another thing to say that a 95% confidence interval for the discrete-valued parameter associated with the choice of threshold variable includes two variables, initial per capita income and property rights. In the former case the threshold effect is consistent with theories of poverty traps and development while in the latter it says something about the importance of economic institutions in posing barriers to growth.

A contribution of this paper is to employ a simultaneous sample split method, Bayesian tree regression (BTREED; see Chipman, George, and McCulloch (1998, 2002)) to deal with this problem. BTREED is a non-sequential regression tree procedure that generates the best tree of every size. Thus, it is less likely to suffer from some of the consequences (e.g., tree instability issues) of sequential sample splitting methods such as TR or CART. Nevertheless, we compare our results with TR since this method provides formal asymptotic theory for the construction of confidence intervals for the threshold estimates.

A second key methodological contribution of this paper is to move the discussion away from model selection towards model averaging in the context of nonlinear (and, in particular, sample split or tree) models. As Cohen-Cole, Durlauf, and Rondina (2006) note, there has not so far been a systematic investigation of model uncertainty and nonlinearities in the growth context. This paper can be viewed as a first attempt towards this ambitious goal. In order to achieve this, we exploit a new statistical learning methodology, Bayesian Additive Regression Trees (BART)[3], developed by Chipman, George, and McCulloch (2002). Specifically, the idea is to generate a large number of trees, each of which is a bad fit for the data as a whole (i.e., a "weak learner"), but gives insight into a small part of the underlying data generation process, so that, taken together, the "sum-of-trees" provides a good estimate of the underlying process. Also, in contrast to single-tree methods, there is no need in BART to condition upon a particular choice of slope covariates and threshold variables. Rather inference is obtained by averaging the sum-of-tree draws from the BART posterior distribution. We view our methodological contribution in this paper as an extension of the standard model averaging exercises recently applied in the empirical growth literature (see Brock and Durlauf (2001) and Sala-i-Martin, Doppelhofer, and Miller (2004) among others).

We find some evidence in the BTREED and TR results that aid may have heterogeneous effects on growth across two growth regimes defined by ethnic fractionalization. In particular, countries that belong to a growth regime characterized by levels of ethnic fractionalization above a threshold value experience

[3]BART is closely related to so-called "ensemble" methods such as random forests (Breiman (2001)), bagging (Breiman (1996)), and, most directly, boosting (Friedman (2001)) in the machine learning literature. Ensemble methods have been shown to have extremely good out-of-sample prediction performance besting even those of neural networks (see, in particular, Friedman (2001) and Hastie, Tibshirani, and Friedman (2001)). Unlike the above mentioned machine learning methods, however, BART is not defined purely by an algorithm, but, instead, by a statistical model within the Bayesian framework.

a negative partial relationship between aid and growth, while those in the regime with ethnic fractionalization below the threshold experience no growth effects from aid at all. We also find that countries in the regime with higher levels of ethnic fractionalization experience, on average, lower growth rates than countries in the lower ethnic fractionalization regime. Nevertheless, we do find substantial tree instability in our sample split exercises so that attempts to characterize the typology of these growth regimes with a high degree of certainty remains elusive. There is evidence that the typology of these regimes may be alternatively well-characterized by property rights institutions or macroeconomic policies such as the level of inflation, and not just ethnic fractionalization. The data simply cannot be certain.

Our BART results are therefore particularly valuable given the high degree of uncertainty generated by tree instability. Here, we find very little evidence to suggest that the relationship between aid and growth is nonlinear for the set of developing countries who are aid recipients. Overall, our results suggest that the partial effect of aid on growth is very likely to be negative although we cannot reject the hypothesis that aid has no effect on growth. In this sense, our findings suggest that aid is potentially counterproductive to growth with outcomes not meeting the expectations of donors. We are therefore sympathetic to the positions of work such as Easterly, Levine, and Roodman (2004) and Rajan and Subramanian (2005a) which are generally pessimistic about the potential contributions of aid to improving economic performance.

The remainder of the paper is organized as follows. In Section 2 we briefly describe our econometric methodology, which includes Bayesian tree regression (BTREED), threshold regression (TR), and Bayesian Additive Regression Trees (BART). In Section 3 we describe our data. Section 4 presents our findings. Section 5 concludes.

2 Econometric Methodology

We conduct our analysis of the relationship between aid and growth using a generalized sample split model that can be defined as follows:

$$g_i = \alpha_j + h_i\beta_j + x_i'\gamma_j + \epsilon_i \ \text{ iff } \ z_i \in R_j(\{\lambda_s\}_{s=1}^{b-1}) \ \text{ for } \ j = 1, ..., b \tag{1}$$

such that

$$\forall j \neq l, \quad R_j \cap R_l = \emptyset \ \text{ and } \ \bigcup_{j=1}^{b} R_j = Z$$

where i indexes the observations (i.e., countries) and j indexes the b growth regimes. g_i is the average growth rate of per capita income for country i across a time period. h_i is the foreign aid proxy (i.e., the variable of interest). We distinguish between two sets of growth determinants. The $k-$dimensional vector x denotes the set of slope covariates while the $p-$dimensional vector z denotes threshold variables.

The set of slope covariates includes the usual Solow regressors, that is, the logarithms of the average rates of physical and human capital accumulation, the logarithm of average population growth rate plus 0.05, and the logarithm of initial per capita income. We also include variables from a wide range of new growth theories including macroeconomic policy, geography, ethnic fractionalization, political institutions, and property rights institutions. Most of the covariates can also be viewed as threshold variables For instance, theories of development such as Azariadis and Drazen (1992) suggest that initial per capita income may act as a threshold variable. To be as agnostic as possible a slope covariate is also a threshold variable as long as it makes sense.

To this end, we specify in our sample split exercises that, with the exception of the factors of accumulation and population growth rates (which are period averages), all slope covariates (including aid) are also threshold variables. The set of parameters is given by $\Psi = (b, \{\lambda_s\}_{s=1}^{b-1}, \Theta)$, where $\Theta = \left(\alpha_j, \beta_j, \gamma_j, \sigma_j^2\right)_{j=1}^{b}$ is the set of regression parameters, b is the number of regimes, and $\{\lambda_s\}_{s=1}^{b-1}$ is the set of threshold parameters that define the set of threshold splits. Note that $\{\lambda_s\}_{s=1}^{b-1}$, in effect, partitions the support of the threshold variables Z into b mutually exclusive regions $\{R_j\}_{j=1}^{b}$.

We can visualize an example of a tree or threshold regression estimation procedure using Figure 1 which is due to Hastie, Tibshirani, and Friedman (2001). Here, the set of observations is partitioned into five regimes, $R_1, \ldots R_5$, defined by the interaction between variables x_1 and x_2. In this example, the model in (1) is modified to be a piece-wise constant model so that a local average is estimated within each regime. The model we use to analyze the effect of aid on growth will be inkeeping with (1); i.e., it will be a piece-wise linear model. That is, we would replace each "step" in Figure 1 with a plane in each growth regime which slope is determined by the coefficients to the local augmented neoclassical growth model defined by (1).

It is worth noting the generality of (1). If we ignore the effects of z on growth; that is, if we specify, a priori, a single growth regime, then we are back to the canonical growth regressions of Mankiw, Romer, and Weil (1992) and Barro (1996). However, as pointed out by Brock and Durlauf (2001), such a formulation ignores prior knowledge regarding the existence of heterogeneity across country units. That is, it ignores the possibility that the effect of the right-hand side covariates on growth may differ systematically across groups of countries. Brock and Durlauf explore a special case of (1) to study the robust heterogeneous effects of ethnic fractionalization on growth. In their paper, the number of regimes b is trivially fixed to two as their threshold variable is a single dummy variable for Sub-Saharan Africa. Given the binary nature of the dummy variable there is no need to estimate a threshold parameter and hence the classical inference is still valid[4]. In contrast, our methodology enables us to

[4]However, this is not true anymore when the threshold variable is not binary and we need to estimate a threshold parameter because the threshold parameter is not identified under the null. Hansen (2000) shows that the inference is non-standard and develops an asymptotic theory for both the threshold parameter and the regression slopes including a method to

have multiple regimes and multiple threshold variables. This is very important in our context given the large number of growth determinants that can act as threshold variables. What is more, the number of regimes b is not pre-specified, but instead is endogenously determined.

One way to estimate (1) is to use the threshold regression methodology of Hansen (2000). At each stage of the sample splitting, we carry out Hansen's test to see whether the sample should be split. If so, we choose the best (in the sense of minimizing sum of squared errors) threshold variable, associated threshold value estimate, and the set of regression estimates for Θ. The same procedure is then applied iteratively to each of the two subsequent subsamples. This "tree growing" procedure stops when either the null of no-split fails to be rejected, or the number of observations in the (sub-)sample falls below a predetermined minimum value. It is worth noting that TR bears deep similarities to the classification and regression trees (CART) method of Breiman, Friedman, Olsen, and Stone (1984). The added advantage of using threshold regression as opposed to CART is that the statistical inference[5] for both the threshold and the regression slopes has been well developed by Hansen (2000).

Its primary weakness, however, lies in the instability of trees to small perturbations in the data as well as in the way that variables are defined. It has been well-documented that small changes in the data can lead to very different threshold variables, threshold values, and even number of regimes being selected by sample splitting methods (see, Hastie, Tibshirani, and Friedman (2001) and also Hong, Wang, and Zhang (2005)). A major reason for the instability of trees is due to the sequential nature of typical sample splitting algorithms. That is, the tree building method does not "update" the tree as it gets bigger. Therefore, it may be that as the tree gets bigger, the previously selected threshold variables and split values in the "upper" parts of the tree (i.e., the initial sample splits) are no longer optimal. We should note that Bai (1999) had suggested an alternative method for getting around the sequential nature of traditional threshold regression models. He calls this method "repartitioning". The idea is to revise upper parts of the tree once lower parts of the tree are estimated. However, we found the practical implementation of repartitioning to be computationally expensive and quickly lost computational tractability even when the tree size was only moderately large. This has led us to consider Bayesian tree regression (BTREED) developed by Chipman, George, and McCulloch (1998, 2002).

BTREED is not a sequential splitting method. Instead, what BTREED does is to search through trees of all sizes (i.e., the (final) number of regimes) and then locate the tree with the highest evidentiary weight for each size. Specifically, it employs MCMC to stochastically search over the posterior distribution of trees for high posterior probability trees. We then select the final tree using BIC. Because each of these trees (no matter the size) is generated probabilistically at every stage of tree building, we do not have the situation, as we do with

---

construct asymptotic confidence intervals for the former.

[5]It should be noted that Hansen (2000) only claims the validity of these results for the single threshold (i.e., two-regime) case, even though he has shown examples of proceeding with these tests iteratively beyond this case.

sequential splitting methods such as TR, where "upper" portions of the tree are never revised even as we vary (increase) the size of trees[6]. We refer the reader to the Technical Appendix for more details on TR and BTREED.

Nevertheless, we should note that BTREED, like TR, is still ultimately a model selection algorithm. Both sample split methods seek to present one tree as the best estimate for the relationship between growth and the set of growth determinants out of the forest of possible trees. While engaging in such model selection has advantages — for instance, it allows us to present a structurally interpretable typology (i.e., tree diagram) for relating aid to growth — this strategy ignores the evidentiary weight associated with alternative trees. Cohen-Cole, Durlauf, and Rondina (2006) have suggested that, even in the context of nonlinear models, researchers should still attempt to report robust estimates of relationships that take into account alternatives to the chosen or benchmark model. We pursue this suggestion in this paper. That is, we attempt to combine the evidentiary weight on the effect of aid on growth across a large number of tree models.

To do so, we employ a new methodology due to Chipman, George, and McCulloch (2005) known as Bayesian Additive Regression Trees (BART). More precisely, we do not condition on a particular choice of slope covariates and threshold variables but rather inference is obtained by pooling information from a large number of tree models in order to flexibly estimate the average effect of a variable of interest on the dependent variable.

Formally, if we define $w_i = (h_i, x_i, z_i)$, then we can write the growth model (1) as

$$g_i = f(w_i) + \varepsilon_i \tag{2}$$

where $\varepsilon_i|w_i \sim N\left(0, \sigma^2\right)$ and $f(w_i) = E(g_i|w_i)$. Then BART provides a way to estimate (2) by combining information across tree models using posterior tree model weights,

$$\widehat{f}(w_i) = \sum_{m=1}^{M} \widehat{f}_m\left(w_i, T_m, \Theta_m\right) \mu\left(m|w_i\right) \tag{3}$$

Here, the j^{th} regime for each of the M trees T_m, $m = 1, ..., M$, is associated with a real parameter θ_j. Hence, any w_i is associated with one of the θ_j within each tree. Letting $\Theta = (\theta_1, \theta_2, ..., \theta_b)$ where b is the number of regimes in T, a single tree model may be denoted by the pair (T, Θ). Let $f(w_i, T_m, \Theta_m)$ denote the θ_j associated with w_i in the m-th tree. The model weights $\mu\left(m|w_i\right)$ are given by Bayes rule,

$$\mu\left(m|w_i\right) \propto \mu\left(w_i|m\right) \mu\left(m\right) \tag{4}$$

so that each weight is the product of the likelihood of the data given a model,

[6]In fact, key steps in BTREED's stochastic tree building algorithm; i.e., "swap" and "change"split decisions (see Technical Appendix), are in the spirit of Bai's "repartitioning".

$\mu(w_i|m)$, and the prior probability for a model, $\mu(m)$. The latter is given by,

$$\begin{aligned} \mu(m) &= \mu\left(\left(T_1,\Theta_1\right),\left(T_1,\Theta_1\right),...,\left(T_1,\Theta_1\right),\sigma\right) \\ &= \mu\left(T_1,T_2,...,T_m\right)\mu\left(\Theta_1,\Theta_2,...,\Theta_m|T_1,T_2,...,T_m\right)\mu(\sigma) \end{aligned} \tag{5}$$

For computational reasons we follow Chipman et al and assume independence so that,

$$\begin{aligned} \mu(m) &= \mu(\sigma)\prod\mu\left(T_j\right)\prod\mu\left(\Theta_j|T_j\right) \\ &= \mu(\sigma)\prod\mu\left(T_j\right)\prod\prod\mu\left(\theta_{l,j}|T_j\right) \end{aligned} \tag{6}$$

where $\theta_{l,j}$ is the l^{th} component of Θ_j.

BART samples from the above posterior using a Markov Chain Monte Carlo (MCMC) algorithm. The construction of each tree T_j for $j = 1, ..., m$ employs precisely the tree building algorithm of BTREED. However, each tree is constrained to be small by appropriately setting the tree priors. The choice of parameter priors are also essentially similar to those of BTREED. Specifically, they are the normal-inverse gamma conjugate priors for the special case where the growth model is constrained to just estimating a constant term θ_j. We refer the reader to the Technical Appendix for more details about BART.

For better approximations, we would want to set M to be relatively large. In our exercises, we follow Chipman et. al. and set $M = 200$. Notice that BART is greatly more flexible than (1). To see this consider first the case of $M = 1$, then $f_1(w_i, T_1, \Theta_1)$ is the conditional mean of g given w. However, when $M > 1$, the terminal node parameters are merely components of the conditional mean of g given w. Furthermore, these terminal node parameters will represent direct and indirect effects (interaction terms) depending on sizes of the trees. In the special case where every terminal node assignment depends on just a single component of w, the sum-of-trees model reduces to a simple additive function of splits on the individual components of w.

To assess the effect of each of the determinants on growth we use Friedman's (2001) partial dependence plot. To do so, first rewrite $f(w)$ as $f(h, h_c)$ where h_c is the complement of h in the set w. To estimate the (partial) effect of h on growth, Friedman suggests that we average out the effect of h_c on growth; i.e.,

$$\begin{aligned} E(g|h) &= E_{\tilde{h}_c}\left[E(g|h)\right] \\ &= E_{\tilde{h}_c}\left[f(h, h_c)\right] \\ &= \int f(h, h_c)\,p(h_c)\,dh_c \end{aligned} \tag{7}$$

However, if the data is $i.i.d$, then, we can approximate (7) with,

$$\widehat{f}_h(h) = \frac{1}{N}\sum_{i=1}^{N} f(h, h_{c,i}) \tag{8}$$

where each $h_{c,i}$ for $i = 1, ..., n$ is an observation in the data. The above is the prediction by BART of the partial dependence of growth rates on h at each level in its support. The pointwise 95% confidence intervals for $\widehat{f}_h(h)$ can also be easily obtain from its posterior distribution using the 2.5^{th} and 97.5^{th} percentiles of the MCMC draws.

3 Data

We use an unbalanced panel dataset (see Table 4) over two periods 1965-79 (42 countries) and 1979-94 (56 countries) based on a broad set of cross-country growth variables. As discussed in the previous section, the dependent variable in (1) is the average growth rate of real per capita GDP corresponding to the two periods. The set of explanatory variables x includes a time dummy for the time period 1979-94 and the canonical Solow variables; i.e., the logarithm of the sum of average population growth plus 0.05 for net depreciation (GPOP), the logarithm of the average proportion of real investments (including government) to real GDP (INV), the logarithm of years of male secondary and higher school attainment (UYRM), and the logarithm of real per capita GDP for the initial year of the time period (Y0). The national accounting data used to construct these data series are obtained form Penn World Table 6.1 (see, Heston, Summers, and Aten (2002)), while schooling data comes from Barro and Lee (2001).

To proxy foreign aid we use data on Effective Development Assistance (EDA) as a share of real GDP constructed by Easterly, Levine, and Roodman (2004) and revised by Roodman (2004). This is the most current version of the panel data used in much of the aid-growth literature (see, for instance, Burnside and Dollar (2000), Hansen and Tarp (2001), Dalgaard, Hansen and Tarp (2004)). This panel data set is available in 5-year periods from 1970-1999. Previously, aid data were only available every 4 years. We use the 5-year panel data set to construct average measures of EDA for the two sample periods 1965-79 and 1980-94.

Following the literature on growth and aid we include four macroeconomic policy variables. We include the logarithm of inflation rate plus one (INF), the ratio of budget surplus to GDP (BS), money supply (M2), and the Sachs-Warner (1995) variable measuring openness to trade (SACW). It is worth noting that we deviate from Burnside and Dollar who include a single measure of economic policies. Burnside and Dollar first estimate a growth regression without aid but with all the covariates and three indicators of macroeconomic policy — log (1+inflation), budget balance to GDP, and the Sachs-Warner (1995) variable. Then, they construct their policy measure by forming a linear combination of the three using the coefficients as weights. We believe that the inclusion of generated regressors in the analysis will result in unnecessary biases so we include all four variables, instead.

Additionally, we expand the Solow space with fundamental determinants of economic growth that include proxies for geography, ethnic fractionalization, political institutions, and property rights institutions. Following Rodrik, Subramanian, and Trebbi (2004) and Sachs (2003) we proxy geography using a climate variable that measures the percentage of a country's land area classified as tropical and subtropical via the Koeppen-Geiger system (KGATRSTR) and a variable that measures the percentage of a country's land area that lies within the geographic tropics (TROPICAR; Gallup, Sachs, and Mellinger (1999)). We also include a variable of geographic isolation that measures the percentage of a country's land area within 100km of an ice-free coast (LCR100KM). To proxy the effect of ethnic fractionalization we use two a measures due to Alesina et al (2003). We include a variable of racial and linguistic characteristics (ETHNIC) and a measure of linguistic fractionalization (LANG). Furthermore, we proxy political institutions using the average of Freedom House index of political rights (POLRIGHTS; see Barro (1991)) while for property rights we use the ratio of assassinations to GDP (ASSAS; see Banks (2002)), a measure of the risk of expropriation of private investments (EXPRSK; see Acemoglu, Johnson, and Robinson (2001), executive constraints (EXCON; Polity IV), and a composite governance index (KKZ96; see Kaufmann, Kraay and Mastruzzi (2005)). Finally, we include time dummies and regional dummies for East Asia (EASIA), Sub-Saharan (SSAFR) Africa, and Latin America (LATINCAR) to account for time and regional heterogeneity, respectively. Please refer to Table 1 for a detailed description of variables. Table 2 provides some summary statistics.

4 Results

4.1 Multiple Regimes and Foreign Aid

We first turn to our sample splitting (TR and BTREED) results. These methods require us to pre-specify which growth variables should be treated as slope covariates, which as potential threshold variables, and finally which as both. We carried out exercises for many alternative specifications. Our aim in carrying out these different exercises is to observe two forms of robustness. Firstly, we want to see if the trees obtained by TR and BTREED are stable. That is, we investigate whether the uncovered tree structures do not vary dramatically across specifications when we (1) vary the set of covariates, (2) given a set of covariates, vary the choices on which variables should be threshold variables, split variables, or both, and (3) vary the number of observations in the data due to the inclusion or exclusion of countries because of variations in missing values across specifications. And secondly, we want to see the extent to which the results obtained by these different sample splitting methods — TR (sequential) and BTREED (non-sequential) — are in agreement.

We report results for three specifications — Baseline, Solow, and Parsimonious — that turned out to be most interesting. The Baseline specification is meant to reflect closely the cross-country growth equation in the aid literature (see Burnside and Dollar (2000)). The set of slope covariates includes the Solow variables (i.e., GPOP, INV, UYRM, and Y0), aid (EDA), macroeconomic policy variables (i.e., SACW, INF, BS, M2), geography (i.e., TROPICAR), ethnic fractionalization (LANG), regional dummies (EASIA, SSAFR, LATINCAR), political institutions (POLRIGHTS), and property rights (ASSAS, EXPRSK, EXCON, KKZ). The set of threshold variables includes most of the slope variables. We do not include in this set the rates of human and physical capital accumulation and population growth rates because these are period averages and not initial conditions.

The Solow specification differs from the Baseline specification in that the set of covariates only includes the Solow and the Aid variables. Following Durlauf, Kourtellos, and Minkin (2001), the idea behind the Solow specification is to examine local generalizations of the Solow model in the sense that a Solow model applies to each country within a growth regime, but the model's parameters vary across regimes.

Finally, the Parsimonious specification aims to maximize the number of observations by excluding the macroeconomic policy variables. Specifically, the set of slope covariates includes GPOP, INV, UYRM, Y0, EDA, TROPICAR, LANG, PRIGHTS, KKZ96, EASIA, SSAFR, and LATINCAR. The set of threshold variables for the Parsimonious specification comprises TROPICAR, LANG, PRIGHTS, KKZ96, AID, and Y0.

Figures 2(a)-(c) show the tree diagrams for BTREED for, respectively, the Baseline, Solow, and Parsimonious specifications, while Figures 3(a)-(c) show the corresponding diagrams for TR. We also report Hansen's 95% confidence bounds in Figures 4(a)-(c); these correspond to the TR threshold value estimates in Figures 3(a)-(c). While the tree structures generated by TR in Figures 3(a)-(c) offer us an interpretable relationship between various growth determinants and economic growth, the confidence bounds provide us with a measure of the uncertainty over the classification of particular countries into each growth regime. The classification of countries into regimes is given, for both BTREED and TR and for all three specifications, in Table 4. Where applicable (i.e., in the TR cases), a superscript "c" denotes countries within Hansen's 95% confidence bounds for the first threshold split as given in Figures 4(a)-(c). Finally, the coefficient estimates and standard errors for each of the BTREED growth regimes are given in Table 5. The corresponding numbers for the TR growth regimes are given in Table 6.

4.1.1 Analysis of Baseline Tree Diagrams

Our Baseline results for BTREED and TR are essentially in agreement. In terms of the tree structures, comparing Figure 2(a) with Figure 3(a), we find that both BTREED and TR identify two growth regimes defined by ethnic heterogeneity (LANG). The size of the regimes are roughly equal. We also note that the regime with ethnic heterogeneity falling below the threshold value (regime (1)) is initially richer and has a faster rate of per capita income growth on average than the regime where ethnic heterogeneity falls above the threshold value (regime (2)). If we look at the country breakdowns for the regimes; please refer to columns 1 and 4 of Table 4, we find that the breakdowns are also very similar for both BTREED and TR. Those countries for which the two are not in agreement — i.e., Algeria and Zimbabwe — fall within Hansen's 95% confidence bounds.

The countries in the high ethnic fractionalization growth regime are predominantly Sub-Saharan African countries (with the key exception of Botswana which is classified as belonging to the other regime). On the other hand, the low ethnic fractionalization growth regime is composed mostly of Latin American and Caribbean countries (with the exception of Paraguay and possibly Guatemala). The countries in Asia, Europe, North Africa, and the Middle East have more heterogeneous predicted growth experiences. While most countries in Asia appear to fall in the worse performing (high ethnic fractionalization) regime, some such as Bangladesh, China, South Korea, and Papua New Guinea are predicted to fall in the better performing (low ethnic fractionalization) group. Similarly, while most countries in the set that we label for convenience as Europe, North Africa, and the Middle East are classified as belonging to the better performing (low ethnic fractionalization) regime, there are notable exceptions such as Iran and Israel that get placed into the worse performing (high ethnic fractionalization) regime.

The finding that ethnic fractionalization is an important driver of heterogeneity in growth is consistent with work by Easterly and Levine (1997) and Alesina et. al. (2003). Easterly and Levine, in particular, argue that ethnic fractionalization is critically important in accounting for Sub-Saharan Africa's underdevelopment. Given that the set of countries in this study are necessarily confined to the set of developing countries (aid recipients), the fact that almost all Sub-Saharan African countries (with the lone and well-documented exception of Botswana (see, for instance, Acemoglu, Johnson, and Robinson (2003))) are separated out in this way and classified under the worse performing regime would appear to provide especially strong support for Easterly and Levine's hypothesis.

4.1.2 Baseline parameter estimates for multiple growth regimes

The evidence on the nature of the growth regimes has important implications for the recent debates over the effect of aid on growth. In contrast to the current literature, our Baseline results suggest that the effect of aid on growth (if any) does not depend on policy variables but rather depends on the fundamental determinant, ethnic fractionalization. Specifically, columns 1 and 2 of Table 5 (for BTREED) and Table 6 (for TR) provide the results for the two growth regimes for the respective sample split methods. We find that aid has no significant effect for countries in the regime with low ethnic fractionalization, but, has a negative and highly significant (at the 1% level) effect for countries in the regime with high ethnic fractionalization. Since the countries in the latter regime are, on average, initially poorer to begin with, our results suggest that aid is in fact strongly counter-productive for this set of countries. Our results therefore are consistent with and support the findings by work such as Easterly, Levine, and Roodman (2004) and Roodman (2004).

In terms of the coefficient estimates and standard errors for growth determinants, the results in Tables 5 and 6 are revealing. For both BTREED and TR, we find that the coefficients to initial per capita income for countries in both the high and low ethnic fractionalization growth regimes are highly significant at the 1% level and negative. A negative coefficient on log initial income per capita is typically taken as evidence in the literature that poorer countries within the regime are catching up with richer countries in the same regime after controlling for other growth factors. Our findings are therefore consistent with the interpretation in the literature of "conditional convergence" within each of the two growth regimes. In this sense, the findings appear to suggest the existence of two convergence clubs defined by ethnic fractionalization, where countries within each club are converging to a different steady state.

Both BTREED and TR find that climate (as measured by TROPICAR) has a significant negative effect on growth for countries in both regimes, while property rights institutions (as measured by EXPRSK) exhibit a significant positive relationship for both regimes. Macroeconomic policies also appear to be important for countries in the worse performing (high ethnic fractionalization) regime. For instance, conditional on the other growth determinants, countries with higher rates of inflation (INF) experience significantly lower growth rates in this regime. Finally, the Solow variables; i.e., population growth (GPOP), investment (INV), and schooling (UYRM), are all significant and have the correct signs; that is, negative, positive, and positive, respectively, for countries in the worse performing (high ethnic fractionalization) regime, although they are insignificant for countries in the low ethnic fractionalization regime.

In sum, the findings from the Baseline specification, which is meant to reflect the literature at large, would appear so far to be stable — in the sense that both BTREED and TR are in agreement — and reflect the consensus of the recent work on the relationship between aid and growth. Nevertheless, we would like to go a step further in order to investigate whether the results we obtained for the Baseline specification holds when we perturb the exercises a bit. We turn now, therefore, to the results for the Solow and Parsimonious specifications.

4.1.3 Results from alternative specifications

Figures 2(b) and 3(b) show the tree diagrams for BTREED and TR, respectively, for the Solow specification. Recall that the only difference between the Solow and Baseline specifications is that, except for aid and the canonical Solow variables, all other variables that were pre-assigned to be slope covariates in the Baseline setup are now assigned to be solely potential threshold variables. As can be seen, the tree diagrams for the Solow specification are dramatically different from those obtained for the Baseline specification. The BTREED tree for the Solow specification is split into two regimes. But, now, the threshold variable selected is no longer ethnic fractionalization, but inflation (INF). Furthermore, the set of countries within each regime also differs dramatically from what we obtained before. There are now 85 observations in one regime (the low inflation regime) and 13 in the other (the high inflation regime) as opposed to 49 for both under the Baseline specification. Also, as far as the breakdown of countries into regimes is concerned (see column 2 of Table 4), there does not appear to be such a strong separation according to geographic regions as we obtained before. Essentially, a few countries from each regional grouping with particularly high levels of inflation are picked out to form the high inflation regime. Nevertheless, if we look at the estimates for the relationship between aid and growth (see columns 4 and 5 of Table 5), we see that it is (negative but) insignificant from zero for both regimes. These results, therefore, should not be taken as evidence to support the position that aid may be beneficial to those developing countries who are made to implement desirable macroeconomic policies as precondition to receiving aid (policy conditionality).

The situation for TR is worse. As can be seen from Figure 3(b), TR now splits the set of countries into five growth regimes according to institutions and geography. These are the low-quality institutions regime (regime (1)), the medium-quality institutions/less tropical regime (regime (2)), medium-quality institutions/more tropical regime (regime (3)), the high-quality institutions/less geographically accessible regime (regime (4)), and the high-quality institutions/more geographically accessible regime ((regime (5)). The classification of countries into regimes is therefore not at all similar to what was achieved before under the Baseline specification. However, if we consider the classification of observations for the Solow specification according to just the first split; i.e., according to whether or not EXPRSK for countries are above or below the threshold value of 0.455, then, the sample splits obtained under TR are somewhat similar to those obtained under BTREED. For instance, if we compare the country breakdown for the first regime in TR with the second regime in BTREED (i.e., compare columns 2 and 4 of Table 4), we see that these are largely similar. Therefore, at least at some level, we find that BTREED and TR do agree on the classification of countries into regimes. However, even if we are willing to concede that, we cannot escape from the fact that BTREED and TR do not agree on the exact driver of heterogeneity. Given the same choices for possible threshold variables, BTREED chooses macroeconomic policies (i.e., inflation (INF)) while TR chooses institutions (EXPRSK). It is therefore very difficult to assign a consistent structural interpretation to these findings.

The tree diagrams for the Parsimonious specification bear somewhat better news. Recall that the difference between the Parsimonious specification and the Baseline and Solow specifications is that for the Parsimonious specification, we drop the set of policy variables (except for aid). The reason we did so was to attempt to maximize the number of observations in the sample. If we compare the TR tree for the Parsimonious specification (Figure 3(c)) with that for the Baseline specification (Figure 3(a)), we see that they are identical. However, when we carry out the analogous comparison for BTREED (i.e., cf. Figure 2(c) with Figure 2(a)), we find that BTREED has selected a single regime (no heterogeneity) model for the Parsimonious specification. Hence, yet again, there is no clear message from our tree diagrams.

In other unreported exercises where we consider alternative choices for designating variables as threshold, slope, or both for these three specifications, we find very little evidence of tree stability. As represented by the trees in Figures 2(a)-(c) and 3(a)-(c), we find that the trees we obtain tend to (1) vary in size, (2) classify countries quite differently, and (3) choose different threshold variables; occasionally by fundamental determinants (such as geography, institutions, or ethnic fractionalization) and other times by policy variables (such as aid, inflation, or government budget surplus). The instability of the trees obtained under both BTREED and TR renders attempts to interpret them structurally to be, unfortunately, precarious. We are forced to conclude that there is very little evidence of a robust/reliable typology that would relate aid to growth. Another way of putting this is that we are severely limited in our ability to engage in tree (model) selection in any sensible way.

Nevertheless, there are some strong regularities in the results across specifications (please refer to Tables 5 and 6). We find that the relationship between aid and growth tends to be negative with most cases being significant. The exception is to be found in the high-quality institutions/less geographically accessible regime (regime (4)) for the Solow specification where the relationship between aid and growth appears to be positive and highly significant. Also, consistent with the larger debate in the growth literature over the importance of institutions versus geography to economic performance, we find that, at least for the set of developing countries in our sample, both these fundamental determinants are important to growth. Climate (as measured by TROPICAR) has a significant negative effect on growth for countries across specifications and regimes with the sole exception of the high ethnic fractionalization regime (regime (2)) for the Parsimonious specification for which its effect is also negative but insignificant. Similarly, property rights institutions (as measured by EXPRSK and KKZ96) have a significant positive effect on growth for countries in all regimes and for all specifications. We also find that conditional convergence holds strongly in the growth regression. For almost all regimes across all specifications (the exception being regime (2) of the Solow specification), we find the coefficient to initial per capita income to be negative and highly significant.

4.2 Robust Relationship between Aid and Growth

These regularities are encouraging because they suggest that even though the instability of the trees we obtained implies that finding one that would be robust enough to tell a structurally interpretable story about the relationship between aid and growth may be difficult, there may be a way for us nevertheless to give policymakers some sense of a "robust" relationship between growth determinants of interest, such as aid, and growth. As described in the Econometric Methods section above, we attempt to uncover such robust relationships using partial dependency plots generated using the BART algorithm.

Figure 5(a) shows the partial (i.e., conditioning upon heterogeneity in terms of the other covariates) dependency plot of growth on aid for the Baseline/Solow set of variables. Similarly, the top left-hand graph in Figure 6 shows the partial dependency plot of growth on aid for the Parsimonious set of variables. We also show the corresponding MCMC confidence bounds around the point estimates in both figures. We find that the (partial) relationship between growth and international aid is probably not nonlinear, and very likely negative. Nevertheless, the confidence bounds do not allow us to reject the possibility that the relationship is flat.

The rest of Figures 5 and 6 show the partial dependence plots for the other growth variables and growth for the respective sets of variables (i.e., Baseline/Solow and Parsimonious). While some of these partial dependence plots — notably those for ethnic fractionalization (LANG) — are suggestive of possible nonlinear relationships, the large confidence bounds make it difficult for us to find conclusively in favor of this outcome. Taken together with the sample split (i.e., TR and BTREED) results, the evidence for a nonlinear relationship between ethnic fractionalization and growth appears to be the strongest amongst the set of regressors.

Like Figures 2(a) and 3(a) for the Baseline model, the partial dependence plots for ethnic fractionalization suggest that there exists a positive relationship between growth and ethnic fractionalization when the degree of fractionalization is low (below approximately 0.45), and a negative relationship when the degree of fractionalization is high (above 0.45).

The plots also show the correct relationships, as suggested by the neoclassical growth model, between the Solow variables and growth; i.e., negative for population growth, positive for investment and schooling, and negative for initial per capita income. They confirm the regularities from the TR and BTREED findings that property rights institutions (as measured by EXPRSK and KKZ96) have strong positive relationships with growth while climate (measured by TROPICAR) has a strong negative relationship. Finally, policies such as trade openness and inflation also appear to have (positive and negative, respectively) consequences for growth.

5 Conclusion

In this paper, we attempt to characterize the relationship between aid and growth using recently developed sample splitting methods such as Bayesian tree regression (BTREED) and threshold regression (TR). Our aim is to uncover the factors that cause divergent effects, if any, of aid on growth for particular subsets of countries. We also sought evidence of a nonlinear relationship between aid and growth. While our results are suggestive of an interaction effect between ethnic fractionalization and aid — so that countries with levels of ethnic fractionalization above a threshold value experience a negative relationship between aid and growth, while those with ethnic fractionalization below the threshold experience no growth effects — our efforts are severely complicated by the high degree of tree instability, and therefore model uncertainty, associated with these sample splitting methods.

A key methodological contribution of our paper therefore is to implement in the growth context a strategy for obtaining robust characterizations of the aid/growth nexus using model averaging methods such as Bayesian Additive Regression Trees (BART). When we do so, we find no evidence of a nonlinear relationship between aid and growth. The relationship between aid and growth is, in fact, likely to be negative. Our findings therefore leave us skeptical as to any potential positive contributions to growth from increasing foreign aid to developing countries. Nevertheless, the evidence from the data is noisy (as seen from the large confidence bounds we obtained), and we therefore expect the debate over the role of foreign aid in promoting growth to continue.

References

1. Acemoglu, D., J. A. Robinson, and S. Johnson (2001), "The Colonial Origins of Comparative Development: An Empirical Investigation," *ican Economic Review*, 91, 1369-1401.

2. Acemoglu, D., J. A. Robinson, and S. Johnson (2003), "An African Success Story: Botswana," in In Search of Prosperity: Analytical Narrative on Economic Growth, edited by Dani Rodrik, Princeton University Press.

3. Alesina, A., A. Devleeschauwer, W. Easterly, S. Kurlat, and R. Wacziarg (2003), "Fractionalization," *Journal of Economic Growth*, 8(2), 155-94.

4. Azariadis, C. and A. Drazen (1990), "Threshold Externalities in Economic Development", *Quarterly Journal of Economics*, 105, 501-526.

5. Bai, J., (1997), "Estimating Multiple Breaks One at a Time," *Econometric Theory* 13, 315-352.

6. Banks, A. (2002), Cross-National Time-Series Data Archive. Bronx, NY: Databanks International.

7. Barro, R. J. (1991), "Economic Growth in a Cross-Section of Countries," *Quarterly Journal of Economics*, 106(2), 407-43.

8. Barro, R. J. and J-W Lee (2001), "International Data on Educational Attainment: Updates and Implications," *Oxford Economic Papers*, 53(3), 541-63.

9. Breiman, L. (1996), "Bagging Predictors," *Machine Learning*, 26, 123-140.

10. Breiman, L. (2001), "Random Forests," *Machine Learning*, 45, 5-32.

11. Breiman, L., J. H. Friedman, R. A. Olsen, and C. J. Stone (1984), Classification and Regression Trees, Wadsworth, Belmont.

12. Brock, W. and S. Durlauf, (2001), "Growth Empirics and Reality," *World Bank Economic Review*, 15, 2, 229-272.

13. Burnside, C. and D. Dollar (2000), "Aid, Policies, and Growth," *American Economic Review*, 90(4), 847-868.

14. Cohen-Cole, E., S. N. Durlauf, and G. Rondina (2005), "Nonlinearities in Growth: From Evidence to Policy; Foreign Aid and Macroeconomic Policy," University of Wisconsin SSRI Working Paper 2005-09.

15. Chipman, H., E. George, and R. McCullough (1998), "Bayesian CART Model Search," *Journal of the American Statistical Society*, 93, 935-960.

16. Chipman, H., E. George, and R. McCullough (2002), "Bayesian Treed Models," *Machine Learning*, 48, 299-320.

17. Chipman, H., E. George, and R. McCullough (2005), "BART: Bayesian Additive Regression Trees," University of Chicago GSB mimeo.

18. Collier, P. and J. Dehn (2001), "Aid, Shocks, and Growth," World Bank Policy Research Working Paper No. 2688.

19. Collier, P. and D. Dollar (2002), "Aid Allocation and Poverty Reduction," *European Economic Review*, 46 (8), 1475-1500.

20. Collier, P. and D. Dollar (2004), "Development Effectiveness: What Have We Learnt?" *Economic Journal*, 114 (496), F244-F271.

21. Collier, P. and A. Hoeffler (2004), "Aid, Policy and Growth in Post-Conflict Countries," *European Economic Review*, 48, 1125-1145.

22. Dalgaard, C-J, H. Hansen, and F. Tarp (2004), "On The Empirics of Foreign Aid and Growth," *Economic Journal*, 114(496), F191-F216.

23. Durlauf, S. N. and P. Johnson, (1995), "Multiple Regimes and Cross-Country Growth Behavior", *Journal of Applied Econometrics*, 10, 4, 363-384.

24. Durlauf, S. N., A. Kourtellos, and A. Minkin (2001), "The Local Solow Growth Model," *European Economic Review*, 45(4-6), 928-940.

25. Durlauf, S. N. (2003), "The Convergence Hypothesis After 10 Years," University of Wisconsin SSRI Working Paper #2003-06.

26. Easterly, W. and R. Levine (1997), "Africa's Growth Tragedy: Policies and Ethnic Divisions", *Quarterly Journal of Economics*, 112 (4), p.1203-50.

27. Easterly, W. (2003), "Can Foreign Aid Buy Growth?" *Journal of Economic Perspectives*, 17(3), 23-48.

28. Easterly, W., R. Levine, and D. Roodman (2004), "New Data, New Doubts: A Comment on Burnside and Dollar's "Aid, Policies, and Growth"," *American Economic Review*, 94(3), 774-780.

29. Friedman, J. H. (2001), "Greedy Function Approximation: A Gradient Boosting Machine," *Annals of Statistics*, 29, 1189-1232.

30. Gallup, J., J. D. Sachs, and A. Mellinger (1999), "Geography and Economic Development," *International Regional Science Review*, 22(2), 179-232.

31. Guillaumont, P. and L. Chauvet (2001), "Aid and Performance: A Reassessment," *Journal of Development Studies*, 37(6), 66-92.

32. Gonzalo, J and J. Pitarakis (2002), "Estimation and Model Selection based Inference in Single and Multiple Threshold Models," *Journal of Econometrics*, 110, 319-352.

33. Hansen, B. E. (1996), "Inference When a Nuisance Parameter Is Not Identified under the Null Hypothesis," *Econometrica*, 64(2), 413-30.

34. Hansen, B. E. (2000), "Sample Splitting and Threshold Estimation," *Econometrica*, 68(3), 575-604.

35. Hansen, H. and F. Tarp (2001), "Aid and Growth Regressions," *Journal of Development Economics*, 64(2), 547-570.

36. Hastie, T., R. Tibshirani, and J. H. Friedman (2001), The Elements of Statistical Learning, Springer-Verlag.

37. Heston A., R. Summers and B. Aten (2002), Penn World Table Version 6.1, Center for International Comparisons at the University of Pennsylvania (CICUP).

38. Hong, Y., D. Wang, and X. Zhang (2005), "Identifying Threshold Effects and Typologies in Economic Growth: A Panel Approach," IFPRI Working Paper.

39. Howitt, P., and D. Mayer-Foulkes (2002), "R&D, Implementation and Stagnation: A Schumpeterian Theory of Convergence Clubs," NBER Working Paper No. 9104.

40. Kaufmann, D., A. Kraay and M. Mastruzzi (2005), "Governance Matters IV: Governance Indicators for 1996-2004," The World Bank, mimeo.

41. Loh, W.-Y. (2002), "Regression Trees with Unbiased Variable Selection and Interaction Detection", *Statistica Sinica*, 12, p. 361-386.

42. Mankiw, N. G., D. Romer, and D. Weil (1992), "A Contribution to the Empirics of Economic Growth," *Quarterly Journal of Economics*, 107, 407-437.

43. Masanjala, W. and C. Papageorgiou (2004), "The Solow Model with CES Technology: Nonlinearities and Parameter Heterogeneity," *Journal of Applied Econometrics*, 19, 2, 171-202.

44. Rajan, R. G. and A. Subramanian (2005a), “”Aid and Growth: What Does the Cross-Country Evidence Really Show?," NBER Working Paper No. 11513.

45. Rajan, R. G. and A. Subramanian (2005b), "What Undermines Aid's Impact on Growth?" NBER Working Paper No. 11657.

46. Rodrik, D., A. Subramanian, and F. Trebbi (2004), "Institutions Rule: The Primacy of Institutions Over Geography and Integration in Economic Development," *Journal of Economic Growth*, 9(2), 131-165.

47. Roodman, D. (2004), "The Anarchy of Numbers: Aid, Development, and Cross-country Empirics", Center for Global Development Working Paper No. 32.

48. Sachs, J. D. and A. Warner (1995), "Economic Reform and the Process of Global Integration," *Brookings Papers on Economic Activity*, 1-118.

49. Sachs, J. D. (2003), "Institutions Don't Rule: Direct Effects of Geography on Per Capita Income," NBER Working Paper No. 9490.

50. Sala-i-Martin, X., G. Doppelhofer, and R. Miller (2004), "Determinants of Long-term Growth: a Bayesian Averaging of Classical Estimates (BACE) Approach," *American Economic Review*, 94, 4, 813-835.

51. Tan, C. M., (2005), "No One True Path: Uncovering the Interplay between Geography, Institutions, and Fractionalization in Economic Development", Tufts University, Dept. of Economics Working Paper No. 2005-12.

Technical Appendix

1 Bayesian Tree Regression (BTREED)

A Bayesian tree regression (BTREED) is defined as a parameter-tree set (Θ, T). BTREED starts by defining priors over unknown quantities, which in this case consists of priors over the parameters Θ and priors over the structure of the tree T. Defining each (Θ, T) set as a tree model and using Bayes rule, the posterior probability of each tree model is derived. That is,

$$p(\Theta, T|Y, X) \propto p(Y|X, \Theta, T)\, p(\Theta, T)$$

$$p(\Theta, T) = p(\Theta|T)\, p(T)$$

where the tree prior and priors over parameters are given respectively by $p(T)$ and $p(\Theta|T)$.

1.1 Priors over parameters $P(\Theta|T)$

The specific choice of priors over parameters will depend on the choice of the model for the likelihood. In the context of this paper, it is assumed that within each of the $j = 1, ..., b$ regimes, growth rates are Gaussian distributed. Formally, growth rates are distributed according to

$$g|h, x, z, T, \Theta \sim N\left(\alpha_j + h\beta_j + x'\gamma_j, \sigma_j^2\right) \quad \text{if and only if} \quad z \in R_j \tag{A1}$$

where $T = \left(b, \{R_j\}_{j=1}^{b}\right)$ denotes a tree structure with b regimes, and Θ is as defined above. The priors over the parameters are then chosen to be of the normal-gamma conjugate form, which is standard in this literature. If we define $w = (h, x, z)$ and $\tilde{\beta} = (\alpha, \beta, \gamma)$, then these priors are specified in the following manner,

$$g|\Theta, T, w \sim N\left(w'\tilde{\beta}, \sigma^2\right)$$

$$\tilde{\beta}|\sigma \sim N\left(\bar{\beta}, \sigma^2 A^{-1}\right)$$

$$\sigma^2 \sim \frac{\nu\lambda}{\chi_\nu^2}$$

where the hyperparameters $\left(\nu, \lambda, \bar{\beta}, A\right)$ are to be chosen. The idea is then to choose values which reflect, as much as possible, prior noninformativeness. Following Chipman, George, and McCulloch (2002), we choose $\nu = 3$, which is interpreted as giving prior information about σ equivalent to that which is contained in 3 observations. Letting s be the classical unbiased estimate of σ based on a linear regression fit for the data, we wish to choose λ to reflect the idea that for each terminal node model, the σ associated with these models should be smaller than s but perhaps not too much smaller. One way to do this, is to choose a quantile q such that $\Pr(\sigma < s) = q$, and then to use the implied value of λ since,

$$\lambda = \frac{s^2 \Phi_v^{-1}(1-q)}{\nu}$$

where Φ_v is the cumulative distribution function for the chi-squared distribution with ν degrees of freedom. It should be noted that the BTREED software initially transforms the data variables into mean 0 and range 1 variables. Hence, noninformativeness would mean that we select $\bar{\beta} = 0$. Finally, to choose A, first assume $A = aI$, where I is the identity matrix. Note that the marginal distribution for β is given by $t_v\sqrt{\frac{\lambda}{a}}$ where t_v is the t distribution with ν degrees of freedom. Hence, we may choose a by choosing a c such that $\Pr(-c < \beta < c) = 0.95$ since the marginal distribution for β yields $a = \frac{\lambda 3.18^2}{c^2}$. Following Chipman et. al., we ran trees for $c = 1$ and 3, and for $\lambda = \left(0.404s^2\right)$ which corresponds to $q = (0.75)$ and $\lambda = \left(0.1173s^2\right)$ which corresponds to $q = (0.95)$.

Note also that since the posterior for $\tilde{\beta}$ is a noncentered t distribution (see Zellner (1971)), exact expressions exist for the posterior mean and variance-covariance of $\tilde{\beta}$. The Bayesian tree regression procedure exploits these to obtain the estimation results.

1.2 Tree Prior $P(T)$

Furthermore, because of the above parametric assumptions, it will be possible to obtain an analytical form for the marginal posterior tree distribution, $p(T|Y,X)$ by integrating across the model parameters. Stochastic search methods can then be employed to locate trees with high posterior probability. In effect, the tree prior is arrived at implicitly through a stochastic tree generation process. Structurally, a (binary) tree consists of nodes which are either terminal, or split into left and right children nodes. At each of these splits, a splitting variable has to be decided upon, and some split value assigned for the chosen variable so as to define the left and right nodes. Therefore, a tree can be generated using the following algorithm (see, Chipman, George, and McCulloch (1998)):

1. Begin by setting T to be the trivial tree consisting of a single root (and terminal) node denoted η.
2. Split the terminal node η with probability $p_{split}(\eta, T)$.
3. If the node splits, assign it a splitting rule ρ according to the distribution $p_{rule}(\rho|\eta, T)$, and create the left and right children nodes.
4. Let T denote the newly created tree, and apply steps 2 and 3 with η equal to the new left and right children nodes.

The splitting probability $p_{split}(\eta, T)$ is modeled as follows,

$$p_{split}(\eta, T) = \phi (1 + d_\eta)^{-\varsigma}$$

where d_η is the depth of the node η (i.e., the number of splits above η). Intuitively, if the term in the RH brackets were taken out so that the probability of a node splitting was set to a constant ϕ, then tuning the hyperparameter ϕ would control the probability of obtaining either larger or smaller size trees (that is, trees with more or fewer terminal nodes). Including the term in the brackets, we see that tuning ς essentially penalizes for more complex trees with deep splits. The idea is to penalize overfitting.

The splitting rule $p_{rule}(\rho|\eta, T)$ which assigns the split value (for the chosen threshold variable) that defines the left and right children nodes is modeled as follows. At every split, a threshold variable is chosen randomly (uniformly so) from the set of all potential threshold variables. If the chosen threshold variable is ordinal, then the split value is chosen uniformly from the available observed values of the threshold variable. If the chosen threshold variable is categorical, then the split value is chosen uniformly from the available categories that define the threshold variable[1].

Therefore, in effect, what BTREED does is to search through trees of all sizes and then locate the tree with highest evidentiary weight. Because each of these trees (no matter the size) is generated probabilistically at every stage of tree building using the algorithm above, we do not have the situation, as we do with sequential splitting methods such as TR, where "upper" portions of the tree are never revised even as we vary (increase) the size of trees.

In terms of the actual hyperparameter values chosen, we follow Chipman, George, and McCulloch and carry out our exercises for a wide variety of prior specifications (see, Figure 3, in particular of Chipman, George, and McCulloch (2002)). We find that our results are robust to changes in tree priors. The trees reported in this paper reflect the choice of $(\phi, \varsigma) = (0.5, 0.5)$. This prior distribution is conservative in that it puts the largest amount of mass on the size 1 tree (simple linear regression/one-regime model) and then tapers downwards with increasing tree sizes. Hence, this prior puts less weight on more complex nonlinear regression structures and puts more weight on a simple linear one. With this set of hyperparameter values, the prior mean size of trees is given to be about 2. This prior reflects a conservative assumption that there should only be a small number of growth regimes given that the set of countries have already been pre-selected to be largely developing countries (international aid recipients). Nevertheless, we emphasize once again that these results are, in fact, representative and are robust to alternative tree prior specifications.

[1] Chipman, George, and McCulloch (1998) refer to this specification for $p_{rule}(\rho|\eta, T)$ as the uniform specification of p_{rule}.

1.3 Metropolis-Hasting Algorithm

The model in (A1) yields an analytical form for the marginal likelihood obtained by integrating across the parameter space,

$$p(g|w,T) = \int p(g|w,\Theta,T)\, p(\Theta|T)\, d\Theta$$

Combining the above with the tree priors, we get the posterior probability for trees,

$$p(T|w,g) \propto p(g|w,T)\, p(T)$$

The idea is now to use a Metropolis-Hasting algorithm to simulate a sequence of trees $T^0, T^1, T^2, ...$ which converge in distribution to the posterior $p(T|w,g)$. The algorithm is as follows. Start with some initial tree T^0 and simulate the transition from any current tree T^i to T^{i+1} in the following manner:.

1. Generate a candidate tree T^* with probability distribution $q\left(T^i, T^*\right)$.
2. Set $T^{i+1} = T^*$ with probability

$$\alpha\left(T^i, T^*\right) = \min\left\{\frac{q\left(T^*, T^i\right)}{q\left(T^i, T^*\right)} \frac{p\left(g|w, T^*\right) p\left(T^*\right)}{p\left(g|w, T^i\right) p\left(T^i\right)}, 1\right\}$$

   Otherwise, set $T^{i+1} = T^i$.

Under weak conditions, the sequence generated by the above algorithm will be a Markov chain with limiting distribution $p(T|w,g)$.

The specification for the transition kernel $q\left(T^i, T^*\right)$ is obtained by randomly choosing among four steps:

1. **GROW** : Randomly pick a terminal node. Split it into two new ones by randomly assigning it a splitting rule according to p_{rule} used in the tree prior.
2. **PRUNE** : Randomly pick a parent of two terminal nodes and turn it into a terminal node by collapsing the nodes below it.
3. **CHANGE** : Randomly pick an internal node, and randomly reassign it a splitting rule according to p_{rule} used in the tree prior.
4. **SWAP** : Randomly pick a parent-child pair which are both internal nodes. Swap their splitting rules unless the other child has an identical rule. In that case, swap the splitting rule of the parent with that of both children.

How do we then go about choosing the good trees generated by this process? One way to do this would be to compare the (unnormalized) posterior probabilities of trees, $p\,(g|w,T)\,p\,(T)$. However, as pointed out by Chipman, George, and McCulloch, there is a subtle problem to using this approach. The problem is that two trees of equal size can have very different prior probabilities, and consequently different posterior probabilities. For example, suppose we had a categorical variable x_1 which took values of either 1 or 2, and another categorical variables x_2 which took values $1, 2, 3, ..., 100$. Then, a binary tree that splits on x_1 has prior probability $\frac{1}{2} \times \frac{1}{1}$ since there are two variables to split on, and given that x_1 is chosen, there is only one assignment possibility given the values x_1 takes. If the tree splits on x_2, then a specific tree will have prior probability $\frac{1}{2} \times \frac{1}{99}$ of occurring, since there are 99 unique split values for this variable. This means that the posterior probability of a given tree can be "diluted" by the prior depending on what variables are split on.

Comparing individual trees using posterior probabilities are therefore misleading. Using posterior probabilities for comparisons only make sense if we are looking at collections of trees. In this example, for instance, it may be possible that there are a dozen different trees that split on x_2 that are all "good". Each might have small posterior probability, but when taken together, they might have greater posterior probability than a single tree splitting on x_2. Chipman, George, and McCulloch suggest that the marginal likelihood $p\,(g|w,T)$ be used instead for locating good trees.

In practice, however, gains in likelihood from picking a larger tree are typically marginal while the reduction in degrees of freedom is proportional to the number of terminal node parameters. Trees are picked, therefore, on the basis of the Schwartz criterion (BIC). The trees reported in this paper are those with the highest such values for runs with $5,000$ iterations per chain and for 1000 restarts. We find that the trees obtained are robust to tree prior specifications as well as to choices of c.

2 Threshold Regression

To motivate our discussion on threshold regressions (TR), consider a simple version of equation (1) in the text where we merge all the growth covariates into x. Suppose there were only two growth regimes and just one threshold variable z so that we can write the modified equation (1) as,

$$g_i = x_i^T \beta_1 + x_i^T I(z_i \leq \lambda)\theta + u_i \tag{1}$$

where, $I(.)$ denotes the indicator function and $\theta = \beta_2 - \beta_1$ is the coefficient heterogeneity between the two regimes. Hansen (1999) pointed out that the threshold parameter λ in (1) is unidentified under the null of one growth regime, so that standard Wald tests performed badly. He offered instead a test with a non-standard distribution for the null of one against two growth regimes. When there are multiple candidate threshold variables for z, Hansen suggests that we test each one and choose the one with the highest evidence (lowest $p-$value) for threshold nonlinearity. Once we have selected a threshold variable z for the pre-specified set of potential threshold variables, Hansen (2000) provides an algorithm for estimating λ using concentrated least squares (CLS) regression based on a sequential search over all $\lambda \in \{ z_1, z_2, ..., z_n\}$. Specifically, notice first that conditional on λ the problem is simple and the LS estimators of β_1 and θ are given by $\widehat{\beta}_1(\lambda)$ and $\hat{\theta}(\lambda)$, respectively. Estimation of λ is then based on minimizing a CLS sum of squared errors criterion S_n defined by

$$S_n(\widehat{\beta}_1(\lambda), \hat{\theta}(\lambda), \lambda) = \sum_{i=1}^{n}(g_i - x_i^T\beta_1 - x_i^T I(z_i \leq \lambda)\theta)^2 \tag{2}$$

which yields the estimator $\widehat{\lambda}$. In turn, the slope estimates can be computed as $\widehat{\beta}(\widehat{\lambda})$ and $\hat{\theta}\left(\widehat{\lambda}\right)$. Importantly, Hansen (2000) also developed an asymptotic distribution theory for both the threshold value estimate and the regression slope coefficients. He proposed a likelihood ratio test for the threshold parameter and showed how to construct asymptotically valid confidence intervals by inverting the likelihood ratio statistic. For the case where we have more than two growth regimes, Hansen suggests that we iteratively apply his methodology to each of the (initial) two regimes above, and carry on this process iteratively.

3 BART

The key reference for this section is Chipman, George, and McCulloch (2005). As discussed in the text, the construction of each tree T_j for $j = 1, ..., m$ in the BART model (see Econometric Methodology section in text) uses the same tree building algorithm as BTREED (as described in the BTREED section of the appendix). Therefore, the BART tree prior or $\mu(Tj)$ in the text is similar to the one used in BTREED. However, each tree is constrained to be small; i.e., a "weak learner". This is done by choosing $\phi = 0.95$ and $\varsigma = 2.$. With this choice, trees with 1, 2, 3, 4, and 5 terminal nodes receive prior probability of 0.05, 0.55, 0.28, 0.09, and 0.03, respectively.

The choice of parameter priors are also almost identical to those for BTREED above. Specifically, they are the normal-inverse gamma conjugate priors for the special case where the growth model is constrained to just estimating a constant term θ_j. θ_j is assumed to be the normally distributed. Then the idea is to choose the prior mean and standard deviation so that a sum of m independent realizations gives a reasonable range for $E(g|w)$. For convenience we start by simply shifting and rescaling the dependent variable g so that we believe the prior probability that $E(g|w) \in (-0.5, 0.5)$ is very high. We then center the prior at zero and choose the standard deviation so that the mean of g falls in the interval $(-0.5, 0.5)$ with "high" probability. In practice, we use the observed

g values, shifting and rescaling so that the observed g range from $(-0.5, 0.5)$. Then, our prior for each θ_j is simply given by

$$\theta_j \sim N(0, \sigma_\theta^2), \quad \text{where } \sigma_\theta = 1/2k\sqrt{m}$$

Notice that the role of the prior variance σ_θ^2 is to constrain each tree so that it plays a smaller role to the overall fit. For larger k, and as the number of trees m increases, this prior distribution will apply greater shrinkage to the θ_j's in each tree. Here, we follow the recommendation of Chipman, George, and McCulloch (2005) and set $k = 2$.

As with BTREED, the conjugate prior for σ^2 here is the inverse chi-square distribution $\sigma^2 \sim \nu\xi/\chi^2(\nu)$. For the hyperparameter choice of ν and ξ, we proceed as follows. We obtain a rough estimate $\widehat{\sigma}^2$ of σ^2 based on the residual sum of squares over n from the least squares estimator. This choice reflects the belief that BART can provide better fit than a linear model. We then pick a degrees of freedom value ν between 3 and 10. Finally, we pick a value of q such as 0.75, 0.90 or 0.99, and set ξ so that the q^{th} quantile of the prior on σ is $P(\sigma < \widehat{\sigma}) = q$.

BART then samples from the posterior (as defined in the text),

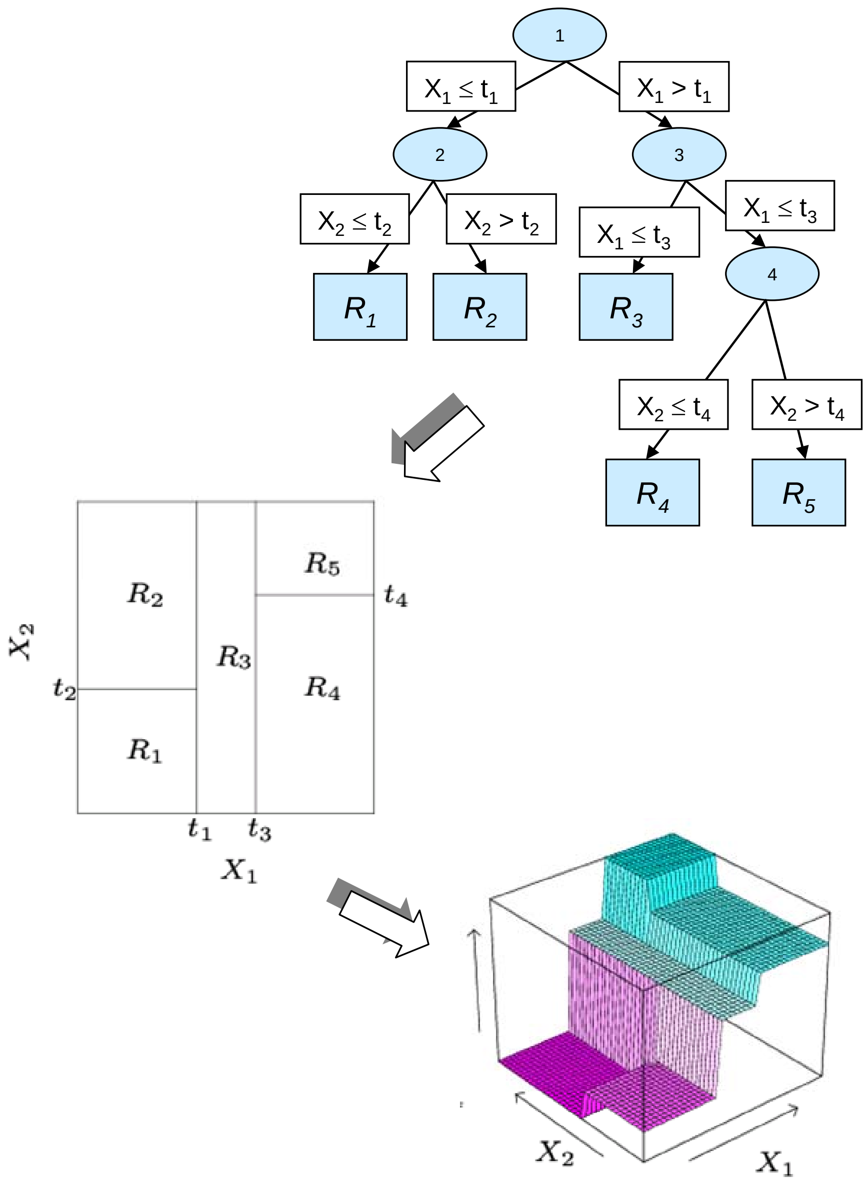
$$\mu(m|w_i) \propto \mu(w_i|m)\,\mu(m)$$

using a Markov Chain Monte Carlo (MCMC) algorithm (we refer the reader to Chipman, George, and McCulloch (2005) for more details). Heuristically, the MCMC sampler is as follows,

1. Repeat for $k = 1, ..., 1000$.
2. Repeat for $m = 1, ..., M$ times for each k.
3. Sample T_j conditional on $w, T_1, ...T_{m-1}, T_{m+1}, ...T_M, \sigma^2$.
4. Sample σ^2 given w and all other tree parameters.
5. Next m.
6. Next k.

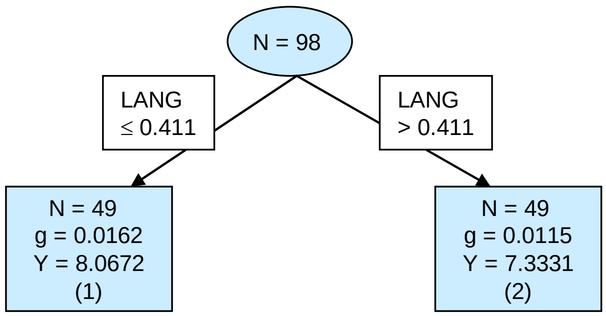
Each sweep of the above algorithm gives one estimate of $f(w)$. That is, each sweep represents a draw from the posterior distribution of $f(w_i)$. The BART estimator $\widehat{f}(w)$ of $f(w)$ is the posterior mean of $f(w)$ which is obtained simply by averaging the MCMC draws. We can also easily obtain pointwise 95% confidence intervals for $\widehat{f}(w)$ from its posterior distribution using the 2.5^{th} and 97.5^{th} percentiles of the MCMC draws.

Figure 1*: Simple Tree Schematic



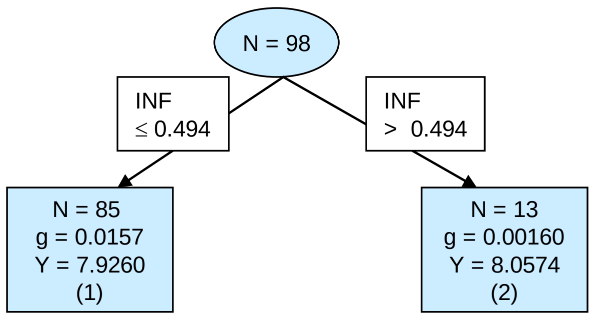
*Schematic due to Hastie, Tibshirani, and Friedman (2001).

Figure 2(a)*: Tree Diagram For BTREED Baseline Model



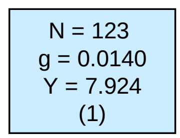
*Note that N gives the number of observations in each node, g stands for the average growth rate in real per capita GDP from 1965-94 for countries in the terminal node, and Y is the corresponding average level of log initial real per capita GDP in 1965.

Figure 2(b)*: Tree Diagram For BTREED Solow Model



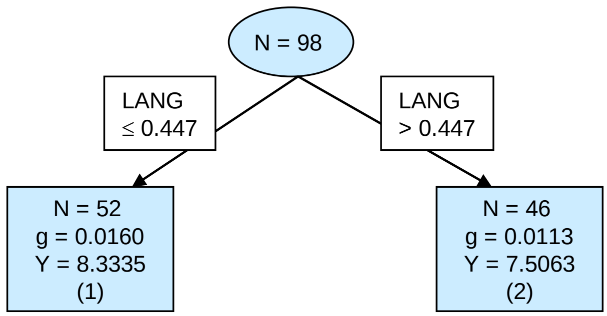
*Note that N gives the number of observations in each node, g stands for the average growth rate in real per capita GDP from 1965-94 for countries in the terminal node, and Y is the corresponding average level of log initial real per capita GDP in 1965.

Figure 2(c)*: Tree Diagram For BTREED Parsimonious Model



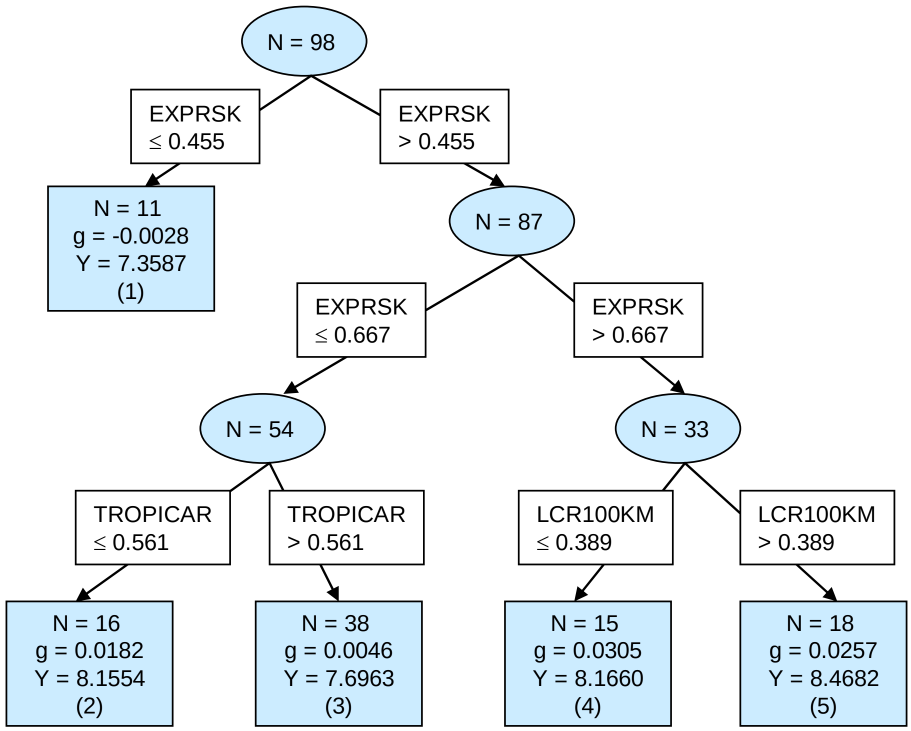
*Note that N gives the number of observations in each node, g stands for the average growth rate in real per capita GDP from 1965-94 for countries in the terminal node, and Y is the corresponding average level of log initial real per capita GDP in 1965.

Figure 3(a)*: Tree Diagram For TR Baseline Model



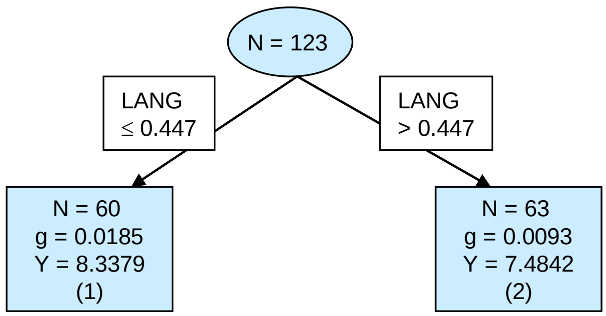
*Note that N gives the number of observations in each node, g stands for the average growth rate in real per capita GDP from 1965-94 for countries in the terminal node, and Y is the corresponding average level of log initial real per capita GDP in 1965.

Figure 3(b)*: Tree Diagram For TR Solow Model



*Note that N gives the number of observations in each node, g stands for the average growth rate in real per capita GDP from 1965-94 for countries in the terminal node, and Y is the corresponding average level of log initial real per capita GDP in 1965.

Figure 3(c)*: Tree Diagram For TR Parsimonious Model



*Note that N gives the number of observations in each node, g stands for the average growth rate in real per capita GDP from 1965-94 for countries in the terminal node, and Y is the corresponding average level of log initial real per capita GDP in 1965.

Figure 4(a): Hansen (2000) Confidence Intervals for Figure 3(a)

First Split:

Threshold Variable	LANG
Threshold Estimate	0.44720000
95% Confidence Interval:	[0.411000, 0.458600]

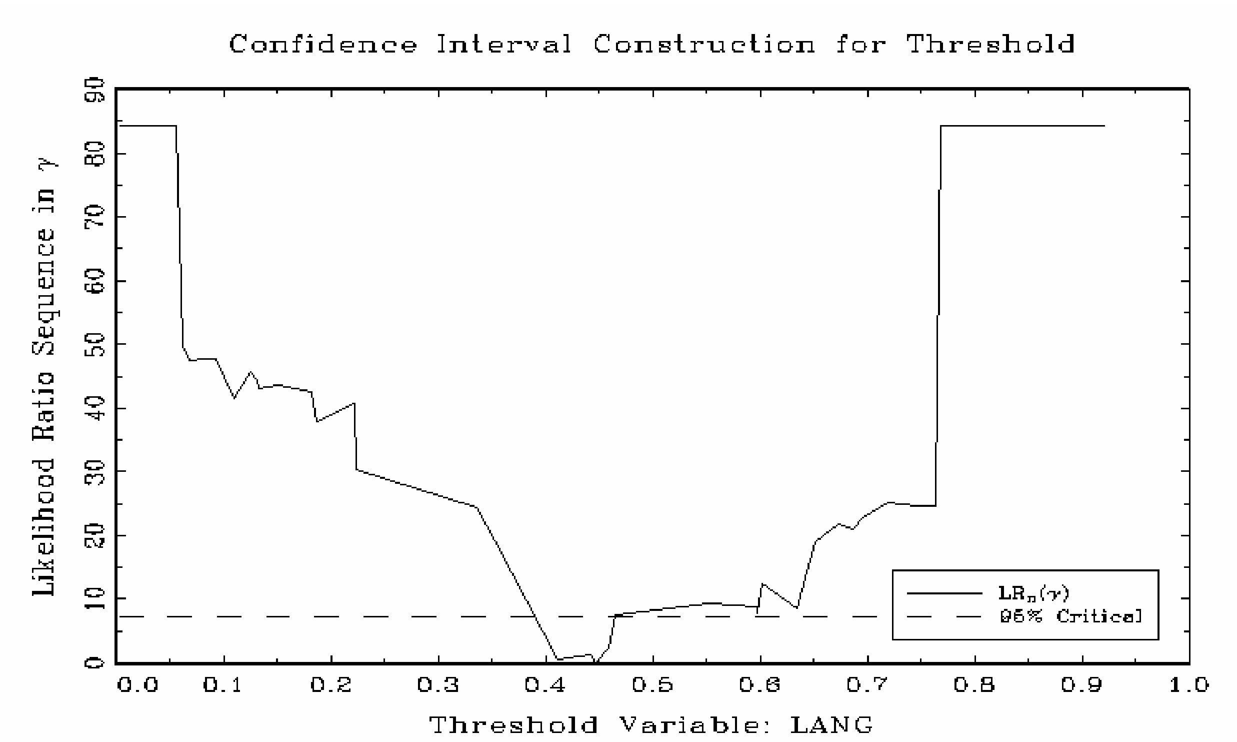
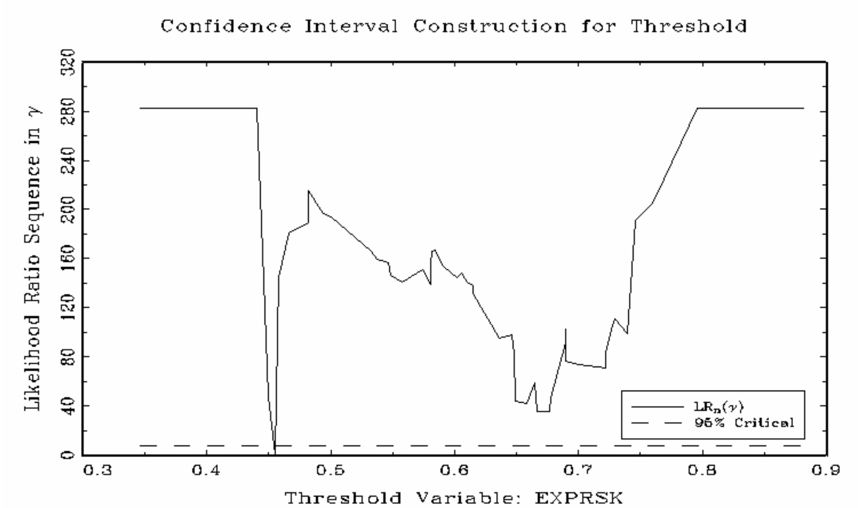


Figure 4(b): Hansen (2000) Confidence Intervals for Figure 3(b)

First Split

Threshold Variable	EXPRSK
Threshold Estimate	0.45538462
95% Confidence Interval:	[0.455384, 0.455384]



Second Split

Threshold Variable	EXPRSK
Threshold Estimate	0.66692308
95% Confidence Interval:	[0.649230, 0.677692]

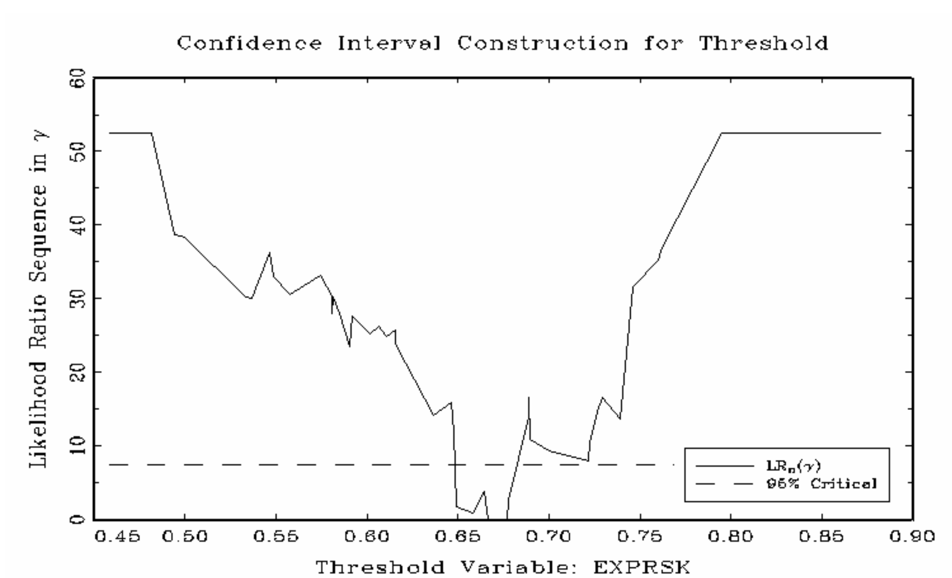
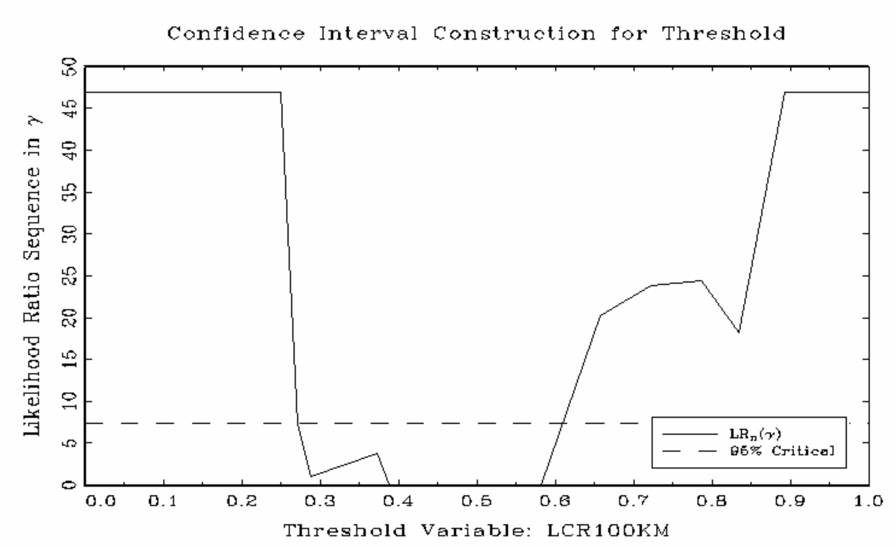


Figure 4(b) (cont.): Hansen (2000) Confidence Intervals for Figure 3(b)

<u>Third Split</u>

```
Threshold Variable            LCR100KM
Threshold Estimate            0.38887620
95% Confidence Interval:      [0.270992, 0.582102]
```



<u>Fourth Split</u>

```
Threshold Variable            TROPICAR
Threshold Estimate            0.56140000
95% Confidence Interval:      [0.561400, 0.561400]
```

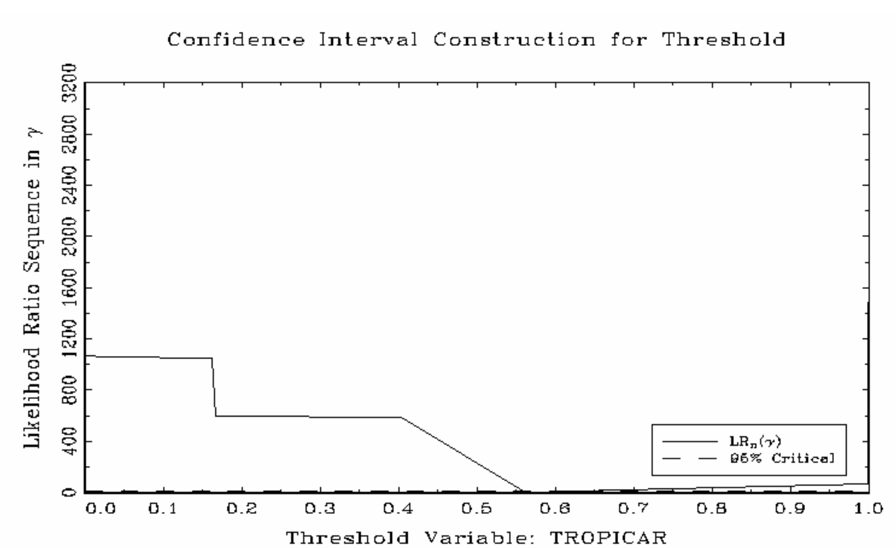


Figure 4(c): Hansen (2000) Confidence Intervals for Figure 3(c)

First Split:

Threshold Variable	LANG
Threshold Estimate	0.44720000
95% Confidence Interval:	[0.335800, 0.597500]

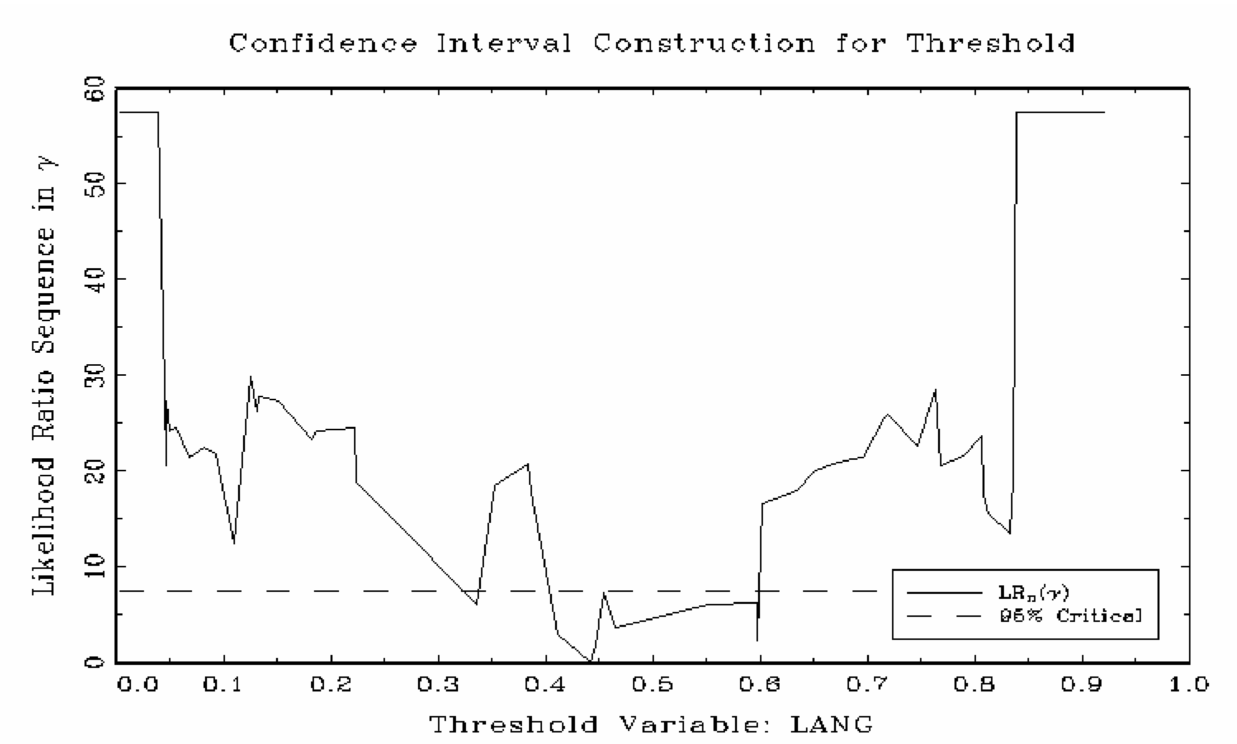
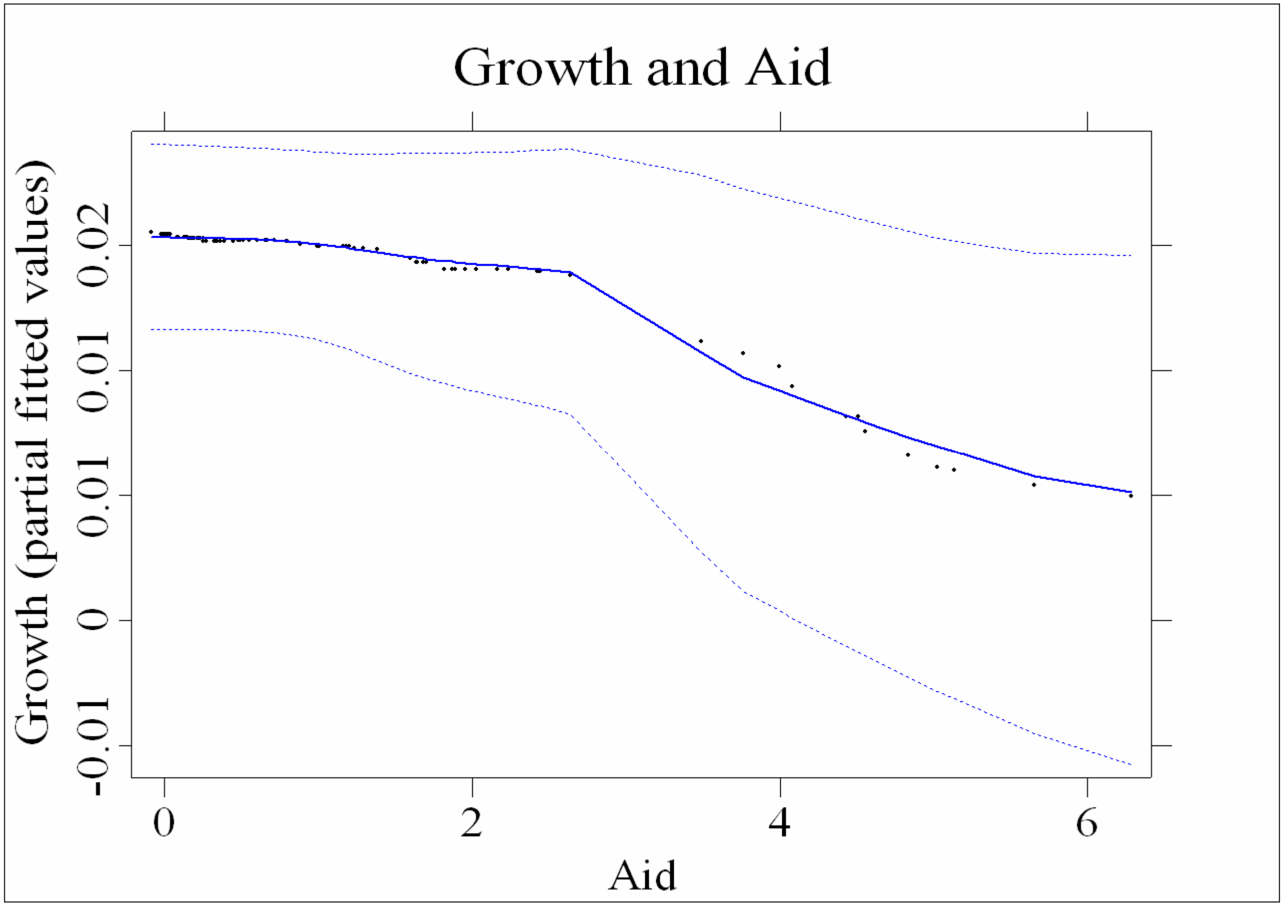
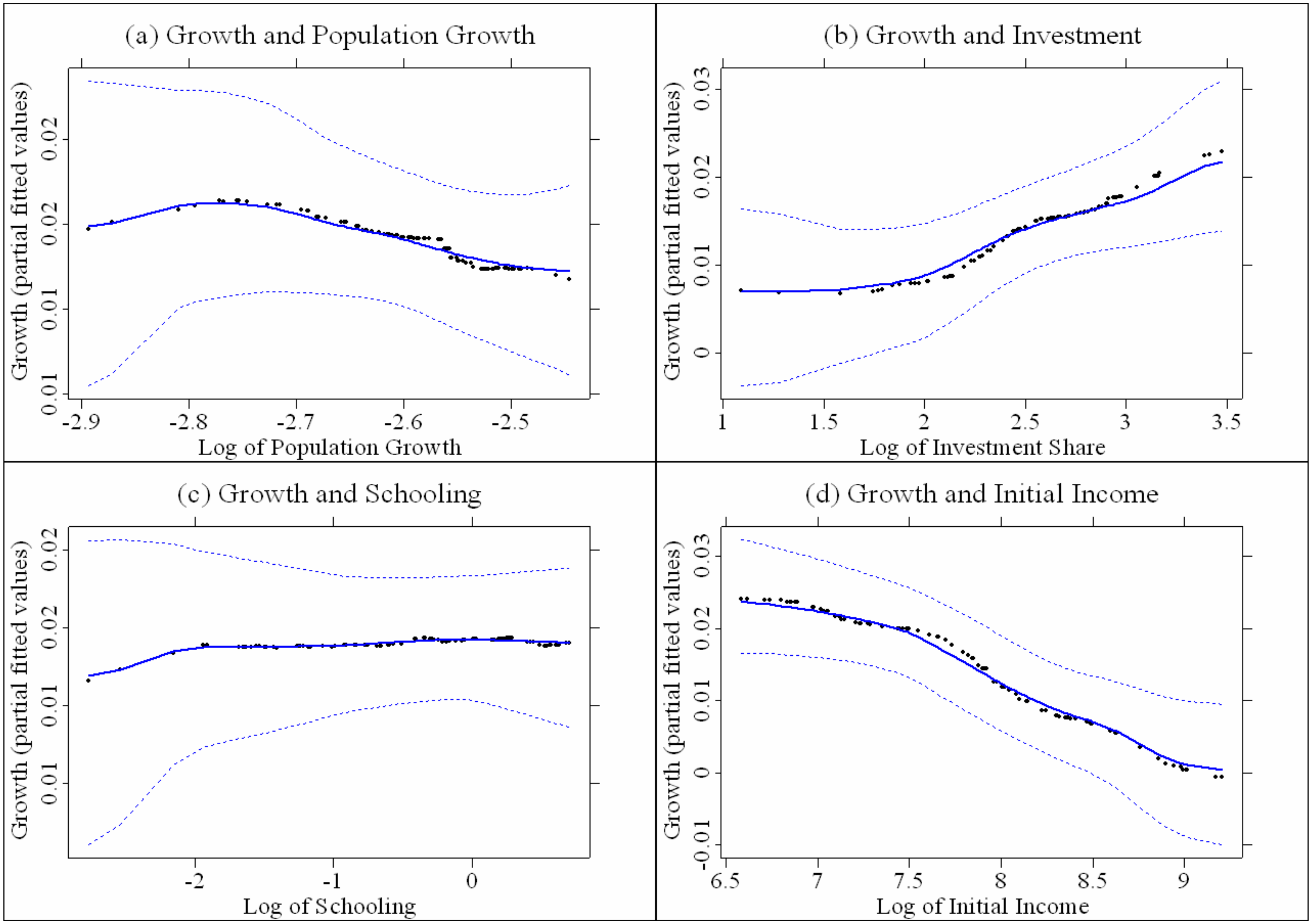


Figure 5(a): Partial Dependence Plot for Aid (Baseline/Solow[ϒ] Model)



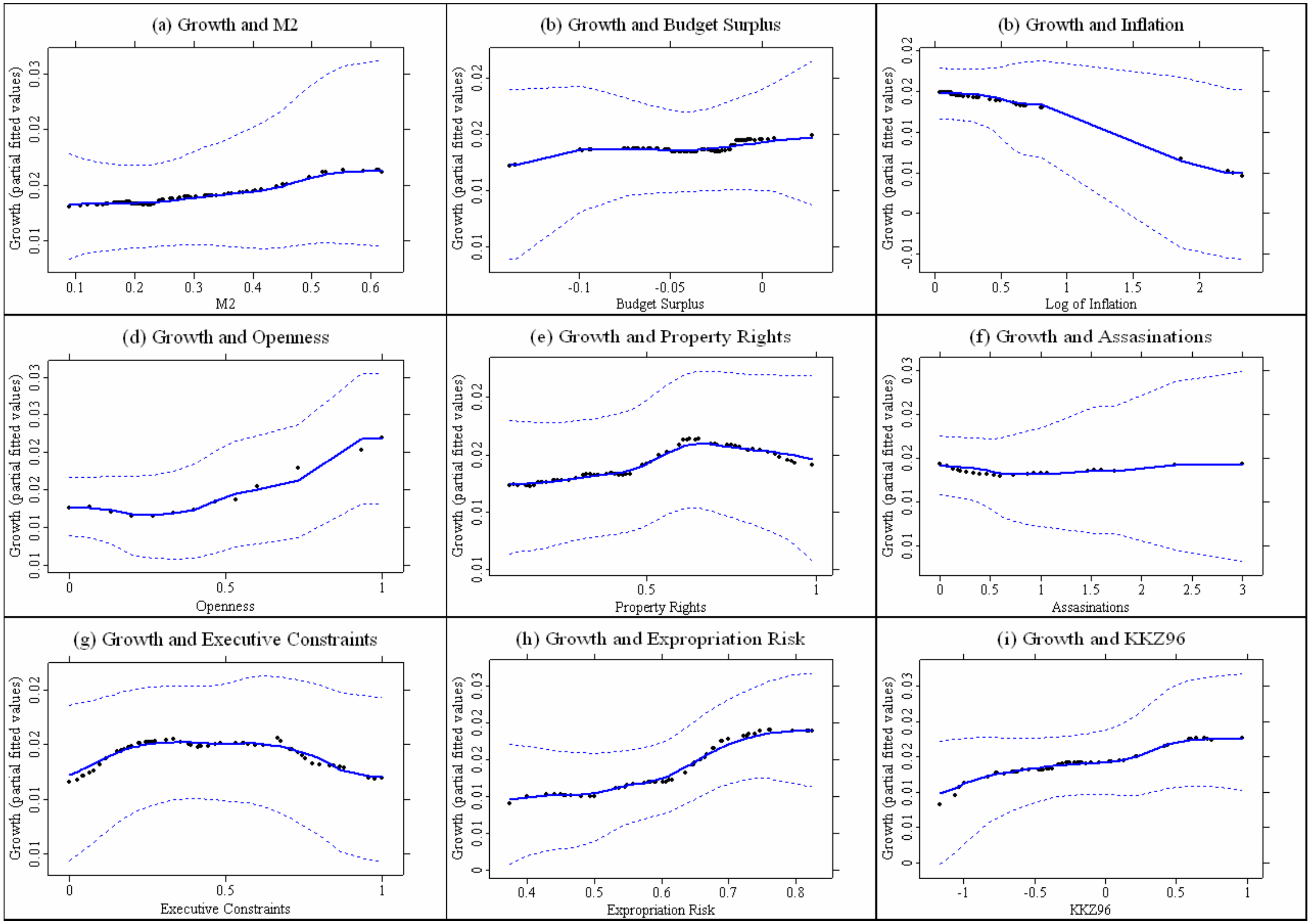
[ϒ] The BART partial dependence diagrams are the same for the Baseline and Solow Models (see Table 3 for model specifications) since both model specifications have the same set of variables.

Figure 5(b): Partial Dependence Plots for Solow Variables (Baseline/Solow[ϒ] Model)



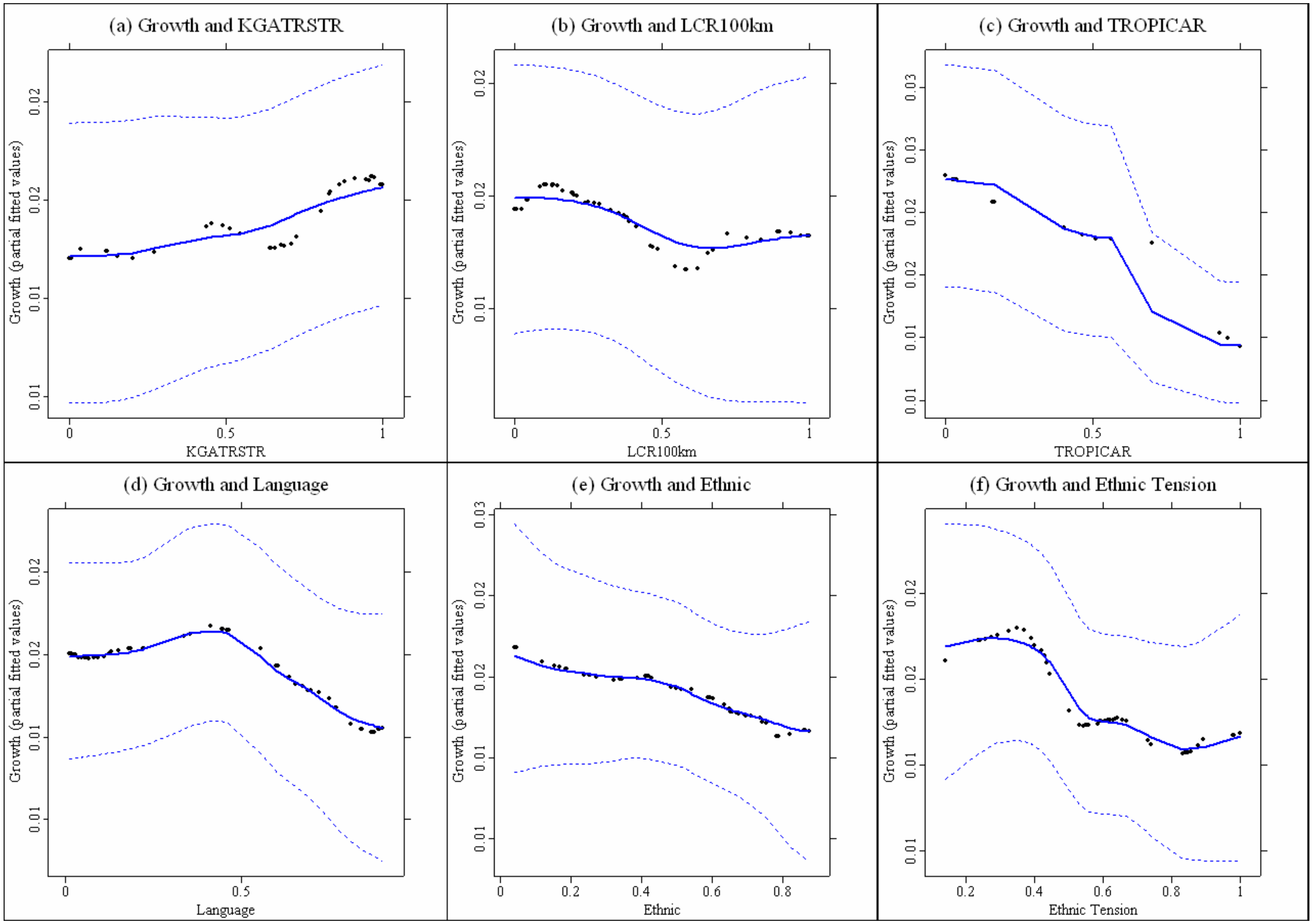
[ϒ] The BART partial dependence diagrams are the same for the Baseline and Solow Models (see Table 3 for model specifications) since both model specifications have the same set of variables.

Figure 5(c): Partial Dependence Plots for Macroeconomic Policy and Institutions (Baseline/Solow[ϒ] Model)



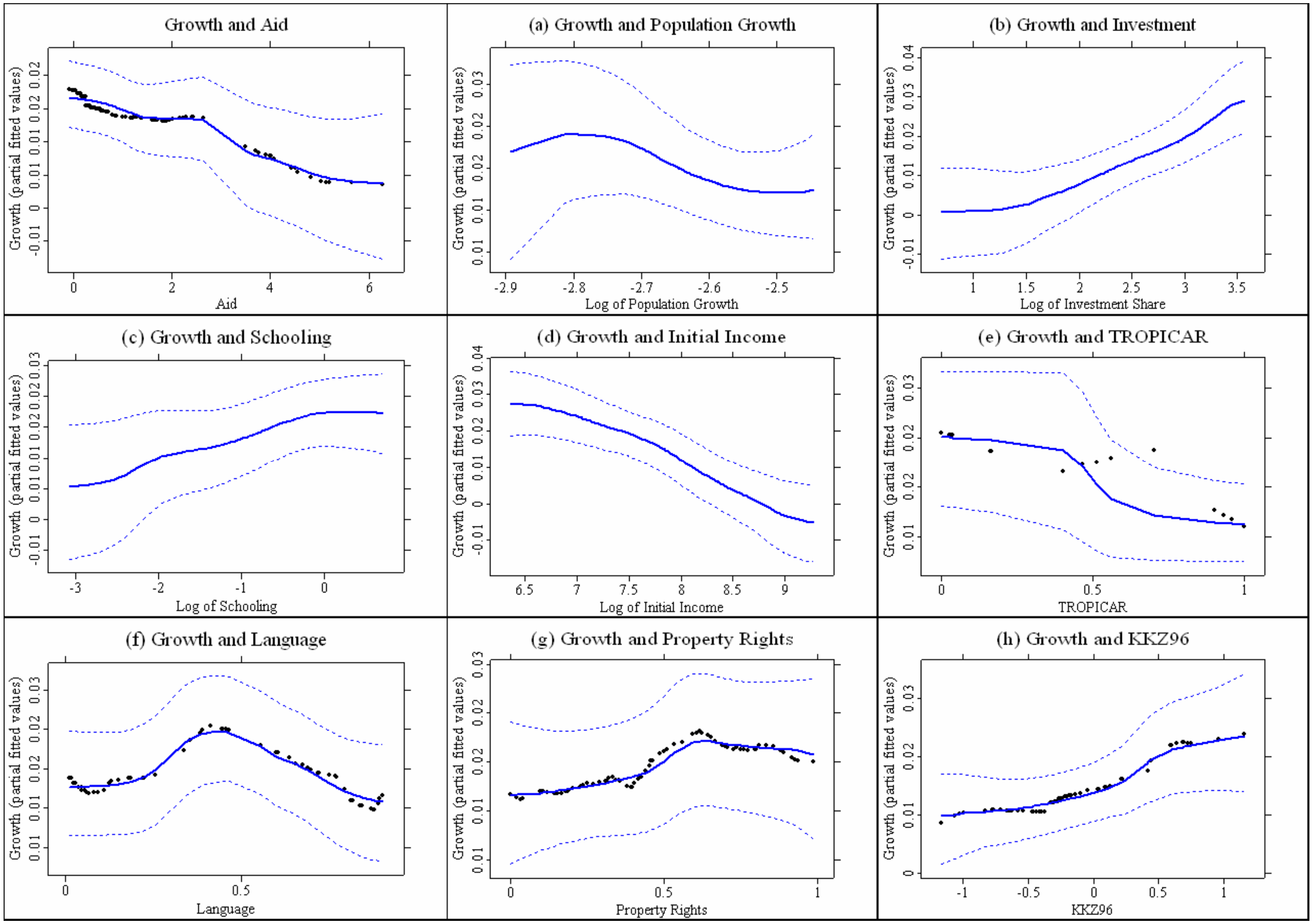
[ϒ] The BART partial dependence diagrams are the same for the Baseline and Solow Models (see Table 3 for model specifications) since both model specifications have the same set of variables.

Figure 5(d): Partial Dependence Plots for Other Fundamental Determinants (Baseline/Solow[ϒ] Model)



ϒ The BART partial dependence diagrams are the same for the Baseline and Solow Models (see Table 3 for model specifications) since both model specifications have the same set of variables.

Figure 6: Partial Dependence Plots (Parsimonious[ϒ] Model)



ϒ Please see Table 3 for model specification.

Table 1: Data Description

VARIABLE	DESCRIPTION	PANEL	SOURCE
g	GDP growth rates (using rgdpch)	1965-79, 1980-94	PWT61
gpop	logarithm of population growth + 0.05	1965-79, 1980-94	PWT61
inv	logarithm of average investments/gdp	1965-79, 1980-94	PWT61
uyrm25	logarithm of average years of male secondary and higher school attainment	1965, 1980	Barro-Lee(2000)
y0	log of initial per capita income	1965, 1980	PWT61
kgatrstr	Percentage of land area classified as tropical and subtropical via the Koeppen-Geiger system.		CID, Harvard University
lcr100km	Percentage of a country's land area within 100km of an ice- free coast.		CID, Harvard University
tropicar	Fraction of land area in geographic tropics.		Gallup and Sachs, 1999
lang	Measure of linguistic fractionalization based on data describing shares of languages spoken as "mother tongues".		Alesina, A., A. Devleeschauwer, W. Easterly, S. Kurlat, and R. Wacziarg (2003)
ethnic	Measure of racial and linguistic characteristics.		Alesina, A., A. Devleeschauwer, W. Easterly, S. Kurlat, and R. Wacziarg (2003)
prights	Political Rights. The variable was tranformed using (7-x)/6 so that lower ratings (closer to zero) are given to countries with poor political rights and higher ratings (closer to one) are given to countries with better political rights.	1972-79, 1980-94	Freedom House 2005

Table 1 (cont.): Data Description

VARIABLE	DESCRIPTION	PANEL	SOURCE
assas	Assassinations per capita	1965-79, 1980-94	Banks (2002)
excon	Rescaled, from 0 to 1, with a higher score indicating more constraint: 0 indicates unlimited authority; score of 1 indicates executive parity or subordination. We calculated the average for each period.	1965-79, 1980-94	Polity IV dataset
exprsk	Risk of “outright confiscation and forced nationalization" of property. Rescaled, from 0 to 1, with a higher score indicating higher less risk of expropriation.	1982-94, 1982-94	IRIS
kkz96	Composite Governance index. It is calculated as the average of six variables: voice and accountability, political stability and absence of violence, government effectiveness, regulatory quality, rule of law, and control of corruption.	1996, 1996	Kaufmann, Kraay and Mastruzzi (2005)
aid	Effective Development Assistance/ real GDP	1970-79, 1980-94	Roodman (2004)
bs	Budget surplus	1965-79, 1980-94	Roodman (2004)
inf	ln(1+ inflation rate)	1965-79, 1980-94	Global Development Network Growth Database.
m2	Average Ratio of M2 to GDP	1965-79, 1980-94	Global Development Network Growth Database.
sacw	Average openness measure proposed by Sachs and Warner	1965-79, 1980-94	Sachs and Warner, 1995; Easterly et al., 2004; Wacziarg and Welch, 2002
easia	A dummy variable for East Asia		
latincar	A dummy variable for Latin America		
ssafr	A dummy variable for sub-Saharan		

Table 2: Summary Statistics

	Min.	Max.	Median	Mean	Std. Dev.
kgatrstr	0.000	1.000	0.656	0.547	0.400
lcr100km	0.000	1.000	0.363	0.433	0.349
tropicar	0.000	1.000	1.000	0.689	0.432
lang	0.003	0.923	0.427	0.418	0.320
ethnic	0.039	0.930	0.540	0.507	0.235
ethtens	0.131	1.000	0.587	0.570	0.241
prights	0.000	1.000	0.417	0.484	0.278
assas	0.000	4.000	0.067	0.373	0.722
excon	0.000	1.000	0.389	0.471	0.328
exprsk	0.346	0.883	0.613	0.614	0.120
kkz96	-1.869	1.159	-0.270	-0.195	0.589
easia	0.000	1.000	0.000	0.092	0.290
ssafr	0.000	1.000	0.000	0.214	0.412
latincar	0.000	1.000	0.000	0.245	0.432
m2	0.073	1.001	0.225	0.278	0.153
bs	-0.206	0.092	-0.033	-0.039	0.040
inf	0.031	3.127	0.119	0.320	0.587
sacw	0.000	1.000	0.067	0.255	0.339
aid	-0.328	9.482	0.495	1.316	1.823
gpop	-3.059	-2.365	-2.580	-2.602	0.111
inv	0.698	3.563	2.568	2.501	0.520
uyrm	-4.017	1.226	-0.298	-0.489	0.936
y0	6.094	9.344	7.906	7.843	0.742
dum2	0.000	1.000	1.000	0.571	0.497
g	-0.053	0.081	0.014	0.014	0.024

Table 3[◊]: Model Specifications

		Baseline Exercise		Solow Exercise		Parsimonious Exercise	
		Slope	Threshold	Slope	Threshold	Slope	Threshold
1	kgatrstr		X		X	-	-
2	lcr100km		X		X	-	-
3	tropicar	X	X		X	X	X
4	lang	X	X		X	X	X
5	ethnic		X		X	-	-
6	ethtens		X		X	-	-
7	prights		X		X	X	X
8	assas	X	X		X	-	-
9	excon		X		X	-	-
10	exprsk	X	X		X	-	-
11	kkz96		X		X	X	X
12	easia	X			X	X	
13	ssafr	X			X	X	
14	latincar	X			X	X	
15	m2	X	X		X	-	-
16	bs	X	X		X	-	-
17	inf	X	X		X	-	-
18	sacw	X	X		X	-	-
19	aid	X	X	X	X	X	X
20	gpop	X		X		X	
21	inv	X		X		X	
22	uyrm	X		X		X	
23	y0	X	X	X	X	X	X
24	dum2	X		X		X	
	# of obs.	98		98		123	

◊ This Table describes the set of variables in the model space for each of the three specifications – Baseline, Solow, and Parsimonious. An "X" means that a variable was designated either to be a potential threshold variable, or a slope covariate (or, as the case may be, both). An "-" means that that variable was dropped from the model space.

Table 4[∇]: Country Breakdowns by Growth Regimes for BTREED and TR Models

	BTREED			TR		
Country	**Baseline**	**Solow**	**Pars.**	**Baseline**	**Solow**	**Pars.**
Africa						
Benin	-	-	1	-	-	2
Botswana	1	1	1	1^c	4	1^c
Cameroon	2	1	1	2	3	2
Central African Rep.	-	-	1	-	-	2
Congo, Rep.	2	1	1	2	1	2
Gambia	2	1	1	2	5	2
Ghana	2	1	1	2	3	2
Kenya	2	1	1	2	3	2
Lesotho	-	-	1	-	-	1
Malawi	2	1	1	2	3	2
Mali	2	1	1	2	1	2
Mauritius	-	-	1	-	-	2^c
Mozambique	-	-	1	-	-	2
Niger	2	1	1	2	3	2
Senegal	2	1	1	2	3	2
Sierra Leone	2	1	1	2	3	2
South Africa	2	1	1	2	4	2
Togo	2	1	1	2	3	2
Uganda	2	2	1	2	1	2
Congo, Dem. Rep.	2	2	1	2	1	2
Zambia	2	2	1	2	3	2
Zimbabwe	2	1	1	1^c	3	2^c
Asia						
Bangladesh	1	1	1	1	2	1
China	1	1	1	1	4	1
India	2	1	1	2	4	2
Indonesia	2	1/2	1	2	5	2
Korea, Rep. of	1	1	1	1	5	1
Malaysia	2	1	1	2	5	2^c
Nepal	-	-	1	-	-	2
Pakistan	2	1	1	2	2	2
Papua New Guinea	1	1	1	1	5	1^c
Philippines	2	1	1	2	3	2
Singapore	-	-	1	-	-	1^c
Sri Lanka	2	1	1	2	3	2^c
Thailand	2	1	1	2^c	4	2

[∇] A superscript "c" denotes countries within Hansen's 95% CI bound for the first threshold split as given in Figures 4(a)-(c). "1/2" indicates that a country was in one regime in one time period and another in the other.

Table 4[▽] (cont.): Country Breakdowns by Growth Regimes for BTREED and TR Models

	BTREED			TR		
Country	**Baseline**	**Solow**	**Pars.**	**Baseline**	**Solow**	**Pars.**
Latin America & the Caribbean						
Argentina	1	2	1	1	2	1
Bolivia	1	1/2	1	1	3	1
Brazil	1	2	1	1	4	1
Chile	1	2	1	1	5	1
Colombia	1	1	1	1	4	1
Costa Rica	1	1	1	1	3	1
Dominican Republic	1	1	1	1	3	1
Ecuador	1	1	1	1	3	1
Guatemala	2	1	1	2^c	1	2^c
Honduras	1	1	1	1	3	1
Jamaica	1	1	1	1	3	1
Mexico	1	1	1	1	4	1
Nicaragua	1	2	1	1	1^c	1
Panama	-	-	1	-	-	1^c
Paraguay	2	1	1	2	2	2^c
Peru	1	2	1	1	3	1^c
Trinidad & Tobago	1	1	1	1	5	1
Uruguay	1	1	1	1	2	1
Venezuela	1	1	1	1	3	1
Europe, North Africa, & Middle East						
Algeria	2	1	1	1^c	2	1^c
Egypt, Arab Rep.	1	1	1	1	2	1
Hungary	1	1	1	1	5	1
Iran	2	1	1	2	1	2
Israel	2	2	1	2	5	2^c
Jordan	1	1	1	1	2	1
Poland	1	2	1	1	5	1
Syrian Arab Rep.	1	1	1	1	2	1
Tunisia	1	1	1	1	2	1
Turkey	1	1	1	1	4	1

[▽] A superscript "c" denotes countries within Hansen's 95% CI bound for the first threshold split as given in Figures 4(a)-(c). "1/2" indicates that a country was in one regime in one time period and another in the other.

Table 5[+]: BTREED Coefficient Estimates for Growth Regimes

	Baseline		Solow		Parsimonious
	(1)	(2)	(1)	(2)	(1)
CONST	0.1798*** (0.0559)	-0.0160 (0.0848)	-0.0027 (0.0499)	-0.0660 (0.4004)	0.1329** (0.0555)
DUM2	-0.0117* (0.0061)	-0.0022 (0.0046)	-0.0134 (0.0039)	-0.0338 (0.0249)	-0.0089** (0.0039)
TROPICAR	-0.0151*** (0.0051)	-0.0155** (0.0064)	- -	- -	-0.0084** (0.0040)
LANG	0.0004 (0.0184)	-0.0003 (0.0123)	- -	- -	-0.0010 (0.0067)
PRIGHTS	- -	- -	- -	- -	0.0084 (0.0065)
ASSAS	0.0019 (0.0026)	-0.0011 (0.0024)	- -	- -	- -
EXPRSK	0.0735*** (0.0238)	0.0271* (0.0155)	- -	- -	- -
KKZ96	- -	- -	- -	- -	0.0128*** (0.0033)
EASIA	-0.0062 (0.0081)	0.0129** (0.0055)	- -	- -	0.0115* (0.0059)
SSAFR	0.0078 (0.0059)	0.0155*** (0.0051)	- -	- -	0.0039 (0.0044)
LATINCAR	-0.0020 (0.0057)	0.0487*** (0.0070)	- -	- -	0.0068 (0.0047)
M2	0.0042 (0.0192)	-0.0072 (0.0205)	- -	- -	- -
BS	0.1509** (0.0621)	-0.0887** (0.0424)	- -	- -	- -
INF	0.0007 (0.0042)	-0.0118*** (0.0035)	- -	- -	- -
SACW	0.0069 (0.0077)	0.0227*** (0.0068)	- -	- -	- -
AID	-0.0022 (0.0017)	-0.0045*** (0.0013)	-0.0009 (0.0011)	-0.0035 (0.0088)	-0.0027** (0.0011)
GPOP	-0.0008 (0.0199)	-0.0486* (0.0280)	-0.0392** (0.0177)	-0.0672 (0.1438)	-0.0019 (0.0184)
INV	0.0072 (0.0083)	0.0150*** (0.0036)	0.0260*** (0.0037)	-0.0172 (0.0405)	0.0173*** (0.0030)
UYRM	-0.0018 (0.0047)	0.0078*** (0.0025)	0.0082*** (0.0024)	0.0117 (0.0253)	0.0061*** (0.0022)
Y0	-0.0260*** (0.0039)	-0.0181*** (0.0032)	-0.0174*** (0.0030)	-0.0045 (0.0344)	-0.0198*** (0.0029)
# of obs.	49	49	85	13	123

[+] Dependent variable is the growth rate of real GDP per capita across, respectively, the periods 1965-79 and 1980-94. Standard errors are in parentheses. Model specifications are described in Table 3. "***" indicates significance at the 1% level while "**" indicates significance at the 5% level and "*" at the 10% level.

Table 6+: TR Coefficient Estimates for Growth Regimes

	Baseline		*Solow*				
	(1)	*(2)*	*(1)*	*(2)*	*(3)*	*(4)*	*(5)*
CONST	0.1858***	-0.0421	-2.2178***	0.1827***	0.1292*	-0.0084	0.1559***
	(0.0451)	(0.0572)	(0.1590)	(0.0247)	(0.0720)	(0.1940)	(0.0500)
DUM2	-0.0109**	-0.0038	-0.0533***	-0.0157***	-0.0079*	-0.0157***	-0.0049
	(0.0052)	(0.0030)	(0.0069)	(0.0033)	(0.0043)	(0.0039)	(0.0047)
TROPICAR	-0.0151***	-0.0169***	-	-	-	-	-
	(0.0041)	(0.0049)	-	-	-	-	-
LANG	-0.0052	0.0060	-	-	-	-	-
	(0.0099)	(0.0122)	-	-	-	-	-
PRIGHTS	-	-	-	-	-	-	-
	-	-	-	-	-	-	-
ASSAS	0.0015	-0.0004	-	-	-	-	-
	(0.0013)	(0.0014)	-	-	-	-	-
EXPRSK	0.0759***	0.0238**	-	-	-	-	-
	(0.0184)	(0.0093)	-	-	-	-	-
KKZ96	-	-	-	-	-	-	-
	-	-	-	-	-	-	-
EASIA	-0.0064	0.0147***	-	-	-	-	-
	(0.0068)	(0.0048)	-	-	-	-	-
SSAFR	0.0079	0.0149***	-	-	-	-	-
	(0.0049)	(0.0041)	-	-	-	-	-
LATINCAR	-0.0012	0.0494***	-	-	-	-	-
	(0.0050)	(0.0062)	-	-	-	-	-
M2	-0.0035	0.0032	-	-	-	-	-
	(0.0154)	(0.0164)	-	-	-	-	-
BS	0.1494***	-0.0877***	-	-	-	-	-
	(0.0445)	(0.0285)	-	-	-	-	-
INF	0.0001	-0.0117***	-	-	-	-	-
	(0.0030)	(0.0021)	-	-	-	-	-
SACW	0.0061	0.0250***	-	-	-	-	-
	(0.0048)	(0.0061)	-	-	-	-	-
AID	-0.0016	-0.0040***	-0.0067***	-0.0015**	-0.0072***	0.0155***	-0.0068***
	(0.0013)	(0.0009)	(0.0016)	(0.0006)	(0.0017)	(0.0036)	(0.0008)
GPOP	-0.0003	-0.0581***	-1.0114***	0.0090	-0.0120	-0.0400	0.0280
	(0.0157)	(0.0187)	(0.0888)	(0.0090)	(0.0262)	(0.0676)	(0.0201)
INV	0.0069	0.0128***	0.0132	-0.0253***	0.0171***	0.0347***	0.0346***
	(0.0071)	(0.0027)	(0.0086)	(0.0058)	(0.0030)	(0.0082)	(0.0040)
UYRM	-0.0006	0.0075***	0.0280***	0.0037	0.0042**	0.0198***	-0.0036
	(0.0030)	(0.0020)	(0.0071)	(0.0045)	(0.0019)	(0.0060)	(0.0039)
Y0	-0.0264***	-0.0179***	-0.0422***	-0.0079***	-0.0233***	-0.0192***	-0.0166***
	(0.0027)	(0.0021)	(0.0096)	(0.0025)	(0.0029)	(0.0044)	(0.0050)
# of obs.	52	46	11	16	38	15	18

+ Dependent variable is the growth rate of real GDP per capita across, respectively, the periods 1965-79 and 1980-94. Standard errors are in parentheses. Model specifications are described in Table 3. "***" indicates significance at the 1% level while "**" indicates significance at the 5% level and "*" at the 10% level.

Table 6[+] (cont.): TR Coefficient Estimates for Growth Regimes

	Parsimonious	
	(1)	*(2)*
CONST	0.1811*** (0.0652)	-0.0307 (0.0662)
DUM2	-0.0099** (0.0047)	-0.0069 (0.0046)
TROPICAR	-0.0101** (0.0042)	-0.0036 (0.0048)
LANG	0.0140 (0.0117)	0.0192 (0.0163)
PRIGHTS	0.0097 (0.0081)	0.0052 (0.0085)
ASSAS	- -	- -
EXPRSK	- -	- -
KKZ96	0.0158*** (0.0032)	0.0098** (0.0045)
EASIA	0.0010 (0.0046)	0.0089* (0.0051)
SSAFR	-0.0014 (0.0068)	0.0028 (0.0049)
LATINCAR	- -	- -
M2	- -	- -
BS	- -	- -
INF	- -	- -
SACW	- -	- -
AID	-0.0023 (0.0014)	-0.0040*** (0.0013)
GPOP	0.0089 (0.0187)	-0.0666*** (0.0212)
INV	0.0263*** (0.0074)	0.0128*** (0.0034)
UYRM	0.0033 (0.0032)	0.0073*** (0.0024)
Y0	-0.0250*** (0.0043)	-0.0171*** (0.0035)
# of obs.	60	63

[+] Dependent variable is the growth rate of real GDP per capita across, respectively, the periods 1965-79 and 1980-94. White standard errors are in parentheses. Model specifications are described in Table 3. "***" indicates significance at the 1% level while "**" indicates significance at the 5% level and "*" at the 10% level.