Do Institutions Rule? The Role of Heterogeneity in the Institutions vs. Geography Debate*

Andros Kourtellos
Department of Economics
University of Cyprus†

Thanasis Stengos
Department of Economics
University of Guelph‡

Chih Ming Tan
Department of Economics
Tufts University§

June 16, 2009

Abstract

We uncover evidence of substantial heterogeneity in the growth experience of countries using a structural threshold regression methodology. Our findings suggest that studies that seek to promote mono-causal explanations in the institutions versus geography debate in growth are potentially misleading.

Keywords: Threshold Regression, Endogenous Threshold Variables, Growth, Institutions, Geography.

JEL Classifications: C21, C51, O47, O43.

*We thank Steven Durlauf and Yannis Ioannides for their insightful comments and suggestions. We are also very grateful to Francesco Trebbi for sharing the Rodrik et. al. dataset with us.

†P.O. Box 537, CY 1678 Nicosia, Cyprus, e-mail: andros@ucy.ac.cy.

‡Guelph, Ontario N1G 2W1, Canada, email: tstengos@uoguelph.ca

§8 Upper Campus Road, Medford, MA 02155, email: chihming.tan@tufts.edu.

1 Introduction

In this paper, we consider one of the important ongoing debates in the growth empirics literature: the "institutions vs. geography" debate. The key question in this debate is whether geography has direct effects on long-run economic performance or if its influence is limited only to its effects on other growth determinants, such as institutions. Attempts to resolve this debate have centered on the use of linear cross-country regressions where the dependent variable is purchasing-power parity adjusted GDP per capita in 1995 while proxies for institutional quality, macroeconomic policies, and geographic endowments form the set of regressors. Acemoglu et. al. (2001), Easterly and Levine (2003), and Rodrik et. al. (2004) conclude that geography's influence on long-run income levels is solely indirect through its effects on institutions, while Sachs (2003) argues that their conclusions are overturned once a measure of malaria transmission is included. Sachs goes further by suggesting that the search for mono-causal effects of fundamental growth determinants on growth may be misdirected. He concludes that, "[t]here is good theoretical and empirical reason to believe that the development process reflects a complex interaction of institutions, policies, and geography [Sachs (2003), p. 9]".

We re-evaluate, in this paper, the conclusions in the institutions versus geography debate by taking Sachs' observation above seriously. Linear cross-country regression exercises potentially ignore possible misspecification of the long-run development process. However, there is both substantial theoretical and empirical support for heterogeneity in the cross-country development process. Canova (1999), Desdoigts (1999), Durlauf and Johnson (1995), Durlauf, Kourtellos, and Minkin (2000), and Mamuneas, Savvides, and Stengos (2006) all find substantial heterogeneity in growth processes across countries using a range of empirical approaches. Further, theoretical work such as the seminal work of Azariadis and Drazen (1990) on threshold externalities as well as the recent work on endogenous economic take-offs (see the survey by Galor (2005)) suggest that heterogeneity across countries may be characterized by threshold nonlinearities.

It is unclear whether previous findings based on the assumption of linearity will be robust once we account for specification issues such as nonlinearity and parameter heterogeneity suggested by the broader growth literature. In this paper, we model nonlinearity and heterogeneity using sample splitting and threshold regression methods (Hansen (2000), Caner and Hansen (2004)). These methods internally sort the data, on the basis of some threshold variable, into groups of observations each of which obeys the same model. This allows researchers to consider differences in the marginal effects of growth determinants (e.g., disease ecology) depending on whether the value of some other growth variable (e.g., institutions) exceeds a threshold level (split value). The attractiveness of this model stems from the fact that it does not treat the split value as predetermined by the researcher, but as a parameter to be estimated. The threshold regression model is therefore a particularly appropriate alternative to the linear model in empirical growth research as it nests the latter, but allows the growth researcher to investigate the possibility of threshold nonlinearities in the growth process and also to uncover the interactions between various growth determinants and their effects on long-run development.

This work is not the first to employ sample splitting and threshold regression methods to a problem in empirical growth. However, previous work using threshold models to account for parameter heterogeneity in growth (e.g., Papageorgiou (2002), Tan (2009)) have assumed that the threshold variable is exogenous. This assumption may be plausible if geography[1] variables were responsible for the threshold effect, but certainly not if institutional quality was the threshold variable since the literature has argued strongly that institutions are endogenous. In this paper, we revisit the institutions versus geography debate within the framework of Kourtellos, Stengos, and Tan (2009); henceforth KST, where we consider a threshold regression model with an endogenous threshold variable. KST can be viewed as an extension of work by Bai and Perron (1998), Hansen (2000), and others who consider threshold models with both exogenous threshold and slope variables, and Caner and Hansen (2004) who then allow for endogeneity in the slope variables.

The main strategy in KST is to exploit the intuition obtained from the limited dependent variable literature, and to relate the problem of having an endogenous threshold variable with the analogous problem of having an endogenous dummy variable or sample selection in the limited dependent variable framework. However, as we pointed out in that paper, there is one important difference. While in limited dependent variable models, we observe the assignment of observations into groups but the (threshold) variable that drives this assignment is taken to be latent, here, we have the opposite: we do not know which observations belong to which group (we do not know the split value), but we can observe the threshold variable. KST show that, just as in the limited dependent variable framework, consistent estimation of parameters requires the inclusion of a set of inverse Mills ratio bias correction terms. Therefore, threshold regression methods that ignore the endogeneity of the threshold variable will generate parameter estimates that are inconsistent. Analogous to Caner and Hansen (2004), KST propose a concentrated least squares estimator for the threshold parameter and estimate the slope parameters by GMM based on the sample split implied by the threshold estimate, which is shown to be consistent and they provide guidance for inference.

When we apply KST to growth data, we find results that offer a markedly more nuanced view from those in the existing institutions versus geography debate where the presence of possible heterogeneity is ignored. Our results also differ substantially from those obtained using methods that ignore the possible endogeneity of the threshold variable. Our results certainly confirm that the

[1] Though both Brock and Durlauf (2001) and Deaton (2009) have pointed out that just because variables are predetermined or external does not automatically make them exogenous.

quality of institutions is an important growth determinant. But, what they really highlight is the role of institutional quality in classifying countries into two long-run development regimes. If the quality of institutions is sufficiently high, then both institutions and geography proponents would agree that higher levels of institutional quality have a positive and significant impact on long-run per capita income. Geography proponents could also legitimately argue that disease prevalence has a significant negative impact on long-run performance. However, for low-quality institutions countries, institutions and geography proponents are likely to hold to their positions and bitterly disagree over the true deep determinant of under-development. Our findings therefore affirm Sachs' conjecture; the development process certainly appears to be an outcome of complex interactions between fundamental causes.

The paper is organized as follows. Section 2 describes the model and the setup. Finally, section 3 provides the results from our empirical application.

2 The Structural Threshold Regression (STR) model

Consider the following structural threshold regression (STR) model of log income per capita

$$y_i = \mathbf{x}_i'\beta_1 + u_i, \quad q_i \leq \gamma \tag{2.1}$$

$$y_i = \mathbf{x}_i'\beta_2 + u_i, \quad q_i > \gamma \tag{2.2}$$

where y_i is the log income per capita in country i, q_i is an endogenous threshold variable (such as the quality of institutions) with γ being the sample split value, $\mathbf{x}_i$ is a vector of growth determinants, β_1 and β_2 are regime-specific slope coefficients, and u_i is an error. We assume that $E(u_i|\mathbf{z}_i) = 0$ where $\mathbf{z}_i$ is a $l \times 1$ vector of instruments with $l \geq p$ where p is the number of endogenous variables. We assume a random sample $\{y_i, x_i, q_i, z_i, u_i\}_{i=1}^n$. A reduced form equation for q_i is given by

$$q_i = \mathbf{z}_i'\pi_q + v_{q,i} \tag{2.3}$$

where $E(v_{q,i}|\mathbf{z}_i) = 0$.

STR is similar in nature to the case of the error interdependence that exists in limited dependent variable models between the equation of interest and the sample selection equation; see Heckman (1979). For example, in the endogenous dummy model, the variable q_i that determines the assignment of observations to regimes is latent, but the assignment is known (given by the dummy variable). However, in the STR case, we observe q_i, but the sample split value γ is unknown, and we estimate it.

Hence, as in the limited dependent case, under joint normality of $(u_i, v_{q,i})$, we have the following conditional expectations

$$E(y_i|\mathbf{x}_i, \mathbf{z}_i, q_i \leq \gamma) = \boldsymbol{\beta}_1'\mathbf{x}_i + \kappa\lambda_1\left(\gamma - \mathbf{z}_i'\boldsymbol{\pi}_q\right) \tag{2.4}$$

$$E(y_i|\mathbf{x}_i, \mathbf{z}_i, q_i > \gamma) = \boldsymbol{\beta}_2'\mathbf{x}_i + \kappa\lambda_2\left(\gamma - \mathbf{z}_i'\boldsymbol{\pi}_q\right) \tag{2.5}$$

where κ is the covariance between u_i and $v_{q,i}$, $\lambda_1(\gamma - z_i'\pi_q) = -\frac{\phi(\gamma-\mathbf{z}_i'\pi_q)}{\Phi(\gamma-\mathbf{z}_i'\pi_q)}$ and $\lambda_2(\gamma - z_i'\pi_q) = \frac{\phi(\gamma-\mathbf{z}_i'\pi_q)}{1-\Phi(\gamma-\mathbf{z}_i'\pi_q)}$ are the inverse Mills ratio bias correction terms, and $\phi(\cdot)$ and $\Phi(\cdot)$ are the Normal pdf and cdf, respectively.

Let $I(\cdot)$ be an indicator function that defines two regimes depending on the value of the threshold variable q_i, where $I(q_i \leq \gamma) = 1$. Further define $\lambda_i(\gamma) = \lambda_1\left(\gamma - \mathbf{z}_i'\boldsymbol{\pi}_q\right) I(q_i \leq \gamma) + \lambda_2\left(\gamma - \mathbf{z}_i'\boldsymbol{\pi}_q\right)(1 - I(q_i \leq \gamma))$ and $\boldsymbol{\delta} = \boldsymbol{\beta}_1 - \boldsymbol{\beta}_2$. Then we can rewrite the STR model as a single equation

$$y_i = \mathbf{x}_i'\beta + \mathbf{x}_{i,\gamma}'\delta + \kappa\lambda_i(\gamma) + \varepsilon_i \tag{2.6}$$

where ε_i is a regression error.

Notice that when the threshold variable q_i is exogenous, i.e. $\kappa = 0$, (2.6) becomes the threshold regression model of Hansen (2000). Additionally, when $\mathbf{x}_i$ is also endogenous then we get the threshold regression model of Caner and Hansen (2004). In both cases, the inverse Mills ratio bias correction terms are omitted so that naively estimating the STR model using Hansen (2000) or Caner and Hansen (2004) would generally result in inconsistent estimation. In a series of Monte Carlo exercises, KST confirm that this is indeed the case.

We estimate the parameters of (2.6) in three steps. First, we estimate the reduced form parameter π_q in (2.3) by LS and obtain $\widehat{\lambda}_i(\gamma) = \widehat{\lambda}_{1,i}(\gamma) + \widehat{\lambda}_{2,i}(\gamma)$,with $\widehat{\lambda}_{1,i}(\gamma) = \lambda_1\left(\gamma - \mathbf{z}_i'\widehat{\boldsymbol{\pi}}_q\right)$ and $\widehat{\lambda}_{2,i}(\gamma) = \lambda_2\left(\gamma - \mathbf{z}_i'\widehat{\boldsymbol{\pi}}_q\right)$. Second, we estimate the threshold parameter γ by minimizing a Concentrated Least Squares (CLS) criterion

$$\widehat{\gamma} = \arg\min_{\gamma} \sum_{i=1}^{n}(y_i - \mathbf{x}_i'\boldsymbol{\beta} - \mathbf{x}_{i,\gamma}'\boldsymbol{\delta} - \kappa\widehat{\lambda}_i(\gamma))^2 \tag{2.7}$$

Finally, once we obtain the split samples implied by $\widehat{\gamma}$, we estimate the slope parameters using GMM.

Using a similar set of assumptions as in Hansen (2000) and Caner and Hansen (2004), KST show that the STR estimator is consistent. They further obtain asymptotic confidence intervals for γ using a suggestion by Seo and Linton (2007).

3 The Institutions versus Geography Debate

In this section, we revisit the institutions versus geography debate. Our work follows most closely Rodrik et. al. The data we use also comes primarily from that paper. The dependent variable is the log of GDP per capita in 1995. As in Rodrik et. al., the set of regressors consists of a measure of institutional quality, the rule of law index (RULE); a measure of trade openness, the logarithm of the ratio of nominal imports plus exports relative to GDP in purchasing power parity-adjusted US dollars (LNOPEN); and two alternative geography measures, distance from the equator of the capital city (DISTEQ) and the malaria index in 1994 (MALFAL94). We consider the sample of countries that corresponds to Rodrik et. al.'s large cross-country set since their findings were shown to be robust to sample variations. Here, RULE is instrumented using the proportion of the population that speaks either English (ENGFRAC) or a major European language (EURFRAC), as suggested by Hall and Jones (1999). We instrument the trade openness variable with Frankel and Romer's (1999) logarithm of predicted trade shares variable (LOGFRANKROM). Following Sachs, we also instrument MALFAL94 using an index of malaria ecology (ME).

Table 1 presents our main findings. We contrast results where the model is assumed to be linear (columns (1)-(2)) against those where the model is a threshold regression model (columns (3)-(10)) that sorts the countries into two regimes. We found evidence for RULE as an endogenous threshold variable. Each threshold model presents the sample split value and the corresponding 90% confidence interval. We also present the GMM slope estimates for each regime.

The linear GMM results replicate those in the literature. When DISTEQ is the geography variable, we find, as Rodrik et. al. do, that RULE is the only variable to have a significant impact on long-run performance (column 1). However, when we replace DISTEQ with MALFAL94, as recommended by Sachs, we find, as he does, that both RULE and MALFAL94 have significant effects on long-run performance. As expected, in both these cases, higher institutional quality was found to be good for long-run performance, while, in the latter case, more severe disease prevalence was shown to have a negative impact (column 2).

When we account for heterogeneity, however, we find that STR delivers more nuanced results compared to the established findings based on the linear model. Compared to Rodrik et. al.'s findings, the STR GMM results (columns (5) and (6)) suggest that there exists substantial heterogeneity in the effect of institutional quality on long-run performance for countries above and below a threshold level. For countries with RULE below -0.736, which corresponds to Pakistan, the marginal impact of improving institutional quality is about 5.5 times larger than that for countries above the threshold value. A one standard deviation improvement of institutional quality would raise long-run performance by 3.3 standard deviations for the low-quality institutions set of countries compared to only less than 0.6 for the higher-quality institutions group. Hence, for this exercise, while the STR GMM results do affirm that "institutions rule" overall, we find that institutions are particularly important for the worst-off countries.

Similarly, our STR GMM results (columns (9) and (10)) for the case where MALFAL94 replaces DISTEQ do support Sachs' finding that "malaria transmission, which is strongly affected by ecological conditions, directly affects the level of per capita income after controlling for the quality of institutions [Sachs (2003), Abstract]". We find that MALFAL94 has a significant negative impact on long-run performance for both low- and high-institutions countries. However, institutional quality (RULE) only has a significant positive impact on long-run performance after countries exceed a threshold level (RULE $>$ -0.195; which corresponds to China). This finding actually strengthens Sach's position since it suggests that the only thing that could deliver marginal improvements for the worst-off countries is the alleviation of the negative effects of a disadvantageous disease ecology. For this group of countries, small changes to institutional quality are unlikely to do much good (unless it gets the country above the threshold point).

Our STR results also contrast with those obtained using Caner and Hansen's (IVTR; 2004) approach that allows for slope covariates to be endogenous, but maintains the assumption of an exogenous threshold variable. We showed in the discussion in the previous section, and also using Monte Carlo experiments that are reported in KST, that the omission of the inverse Mills ratio bias correction terms results in the estimators for the slope parameters for IVTR to be inconsistent. However, IVTR has seen recent popularity in its application to growth empirics (e.g., Papageorgiou (2002)), so we also compare our STR findings to those of IVTR. In comparison to STR, for the Rodrik et. al. specification, the IVTR results (columns (3) and (4)) provide weaker support for Rodrik et. al.'s findings. The IVTR results suggest that institutional quality only matters strongly (at the 1% significance level) after a country has attained a minimum level (RULE $>$ 0.231; which corresponds to India). Below that level, variations in none of the growth determinants has any influence on long-run performance at the 5% level. Similarly, the IVTR findings for the Sachs specification (columns (7) and (8)) also dilute the importance of MALFAL94. In contrast to the STR findings, the IVTR results suggest that the negative impact of disease prevalence only applies to the worst-off countries. For high-quality institutions countries, only continued improvements in institutions would deliver significant (positive) marginal payoffs in terms of better long-run performance. The sharp differences between the STR and IVTR findings suggest that not accounting for the endogeneity of the threshold variable in threshold regression exercises could deliver conclusions that are highly misleading in practice.

In sum, we conclude that our findings differ substantially from those obtained from methods that either ignore the presence of thresholds altogether or ignore the possible endogeneity of the threshold variable. There is much evidence to suggest that there exists substantial heterogeneity in the growth experiences of countries, and that studies that seek to promote mono-causal explanations for the variation in long-run economic performance across countries are potentially misleading.

References

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

[2] Bai, J. and P. Perron (1998), "Estimating and Testing Linear Models with Multiple Structural Changes," *Econometrica*, 66(1), 47-78.

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

[4] Caner, M. and B. Hansen (2004), "Instrumental Variable Estimation of a Threshold Model," *Econometric Theory*, 20, 813-843.

[5] Canova F. (2004), "Testing for Convergence Clubs in Income Per Capita: A Predictive Density Approach," *International Economic Review*, 45(1), 49-77.

[6] Deaton, A. (2008), "Instruments of Development: Randomization in the Tropics, and the Search for the Elusive Keys to Economic Development," Keynes Lecture, British Academy, October 9, 2008.

[7] Desdoigts A. (1999), "Patterns of Economic Development and the Formation of Clubs," *Journal of Economic Growth*, 4(3), 305-30.

[8] Durlauf S. and P A. Johnson (1995), "Multiple Regimes and Cross Country Behavior," *Journal of Applied Econometrics,* 10(4), 365-384.

[9] Durlauf, S., A. Kourtellos, and A. Minkin (2001), "The Local Solow Growth Model," *European Economic Review*, 45, 4-6, p. 928-40.

[10] Easterly, W. and R. Levine (2003), "Tropics, Germs, and Crops: How Endowments Influence Economic Development," *Journal of Monetary Economics*, 50(1), p. 3-39.

[11] Galor O. (2005), "From Stagnation to Growth: Unified Growth Theory," in *Handbook of Economic Growth*, P. Aghion and S. Durlauf, eds., Amsterdam: North Holland.

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

[13] Heckman, J. (1979), "Sample Selection Bias as a Specification Error," *Econometrica*, 47(1), p. 153-161.

[14] Kourtellos, A., T. Stengos, and C. M. Tan (2009), "Structural Threshold Regression," Rimini Centre of Economic Analysis, working paper series.

[15] Mamuneas, T., A. Savvides, and T. Stengos (2006), "Economic Development and Return to Human Capital: A Smooth Coefficient Semiparametric Approach," *Journal of Applied Econometrics*, 21(1), p. 111-132.

[16] Papageorgiou, C. (2002), "Trade as a Threshold Variable for Multiple Regimes," *Economics Letters*, 77(1), 85-91.

[17] 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), p.131-165.

[18] Sachs, J. (2003), "Institutions Don't Rule: Direct Effects of Geography on Per Capita Income," National Bureau of Economic Research Working Paper No. 9490.

[19] Seo, M. H. and O. Linton (2007), "A Smoothed Least Squares Estimator For Threshold Regression Models," *Journal of Econometrics*, 141(2), p. 704-735.

[20] Tan, C. M. (2009), "No One True Path: Uncovering the Interplay between Geography, Institutions, and Fractionalization in Economic Development", *Journal of Applied Econometrics*, forthcoming.

Table 1: Regressions of log per capita GDP

	Linear-GMM		IVTR-GMM		THRET-GMM		IVTR-GMM		THRET-GMM	
Threshold **Sample Split** **90 % CI**			*q=Rule* 0.231 [0.158, 0.231]		*q=Rule* -0.736 [-0.867, -0.736]		*q=Rule* 0.231 [0.158, 0.551]		*q=Rule* -0.195 [-0.442, 0.722]	
			$q \leq$ 0.231	q> 0.231	$q \leq$ -0.736	q> -0.736	$q \leq$ 0.231	q> 0.231	$q \leq$ -0.195	q> -0.195
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
RULE	1.334*	0.700*	5.319**	0.924*	3.336*	0.589*	0.407	0.938*	1.208	1.102*
	(0.287)	(0.148)	(3.087)	(0.284)	(1.405)	(0.254)	(0.878)	(0.379)	(0.802)	(0.405)
LNOPEN	-0.286	-0.034	-0.995	0.038	0.487	0.074	0.058	-0.015	0.371	0.050
	(0.255)	(0.178)	(1.090)	(0.110)	(0.667)	(0.153)	(0.212)	(0.169)	(0.467)	(0.136)
DISTEQ	0.001	-	-0.031	0.002	-0.012	0.003	-	-	-	-
	(0.009)	-	(0.037)	(0.006)	(0.024)	(0.005)	-	-	-	-
MALFAL94	-	-1.375*	-	-	-	-	-1.436*	-0.763	-1.324*	-1.243*
	-	(0.213)	-	-	-	-	(0.215)	(0.987)	(0.391)	(0.252)
No. of observations	120	120	76	44	28	92	70	37	55	52
J-stat: $\chi^2(1)$	6.555	1.350	1.647	0.066	0.628	3.323	2.590	0.274	1.653	0.0183

All the regressions include an unreported constant. Standard errors are in parentheses. "*" denotes significance at 1%, "**" at 5%. The quality of institutions variable, RULE, is the Rule of Law Index for 2001. The dependent variable is the natural logarithm of per capita GDP in PPP US dollars in 1995. LNOPEN is the natural logarithm of real openness defined by the ratio of nominal imports plus exports to GDP in PPP US dollars. DISTEQ is the distance from Equator of capital city measured as abs(Latitude)/90. MALFAL94 is the Malaria index in 1994. ENGFRAC is the fraction of the population speaking English. EURFRAC is the fraction of the population speaking on of the major languages of Western Europe. LOGFRANKROM is the natural logarithm of predicted trade shares computed from a bilateral trade equation with "pure geography" variables. ME is a population weighted Malaria Ecology index that includes temperature, species abundance, and vector type (the type of mosquito). We refer the reader to Rodrik et. al. (2004) for detailed data references.