NBER WORKING PAPER SERIES

THE WEALTH OF NATIONS: FUNDAMENTAL FORCES VERSUS POVERTY TRAPS

David E. Bloom
David Canning
Jaypee Sevilla

Working Paper 8714
http://www.nber.org/papers/w8714

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
January 2002

The Wealth of Nations: Fundamental Forces Versus Poverty Traps
David E. Bloom, David Canning and Jaypee Sevilla
NBER Working Paper No. 8714
January 2002
JEL No. O1

ABSTRACT

We test the view the large differences in income levels we see across the world are due to differences in underlying characteristics, i.e. fundamental forces, against the alternative that there are poverty traps. Taking geographical variables as fundamental characteristics, we find that we can reject fundamental forces in favor of a poverty trap model with high and low level equilibria. The high level equilibrium state is found to be the same for all countries while income in the low level equilibrium, and the probability of being in the high level equilibrium, are greater in cool, coastal countries with high, year-round, rainfall.

David E. Bloom
Harvard University
Department of Population
and International Health
Building I, Room 1110B
677 Huntington Avenue
Boston, MA 02115
and NBER
dbloom@hsph.harvard.edu

David Canning
School of Economics
and Management
Queen's University of Belfast
d.canning@qub.ac.uk

Jaypee Sevilla
School of Public Health
Harvard University
jsevilla@fas.harvard.edu

1. Introduction

Why are some countries so rich while others are so poor? There are two fundamentally different approaches to answering this question. One approach is to trace economic outcomes back to fundamental forces, such as geography, climate, or culture. The alternative approach is to argue for the existence of multiple equilibria. If there are multiple equilibria, one country may be stuck in a poverty trap while another may be very wealthy, even though both countries have exactly the same exogenous characteristics.

Most approaches to estimating and testing poverty trap models have focused on short run and medium run dynamics as countries gravitate towards different "convergence clubs," depending on their initial positions. Some recent examples of this approach are Durlauf and Johnson (1995), Berthelemy and Varoudakis (1996), Desdoigts (1999), and Feve and Le Pen (2000). Instead we concentrate on the implications of a poverty trap for long run behavior. In the long run, a theory of "fundamental forces" corresponds to a unique relationship between exogenous factors and income levels, while the existence of a poverty trap generates at least two relationships.

Our test consists of asking whether moving from a single relationship between exogenous factors and economic performance, to multiple relationships (in practice we allow only two), significantly improves the model's fit. The advantage of this approach is that it depends only on the "reduced form" mapping from exogenous factors to the level of income, while the approaches based on short and medium run behavior require the inclusion, and modeling, of all the endogenous variables in the development process. There is, however, a cost in that we test only the existence of a poverty trap; we do not specify the mechanism by which the poverty trap emerges.

If we assume that the "fundamental factors" are the same in every country, the income levels of countries should be distributed around a single expected value, while multiple equilibria would generate income levels that are clustered around several distinct values. This model has been extensively investigated by Quah (1996, 1997) who argues that there are "twin peaks" in the empirical distribution of national per capita incomes. Bianchi (1997) and Paap and van Dijk (1998) test the existence of "twin peaks" against a single peak and reject a single peak; "twin peaks" capture the structure of the data much better than a single peak.

However, this result may be due to the omission of some "fundamental forces." Intuitively, if countries differ in some fundamental ways, these differences may explain the actual distribution of income per capita. The "twin peaks" may be capturing the distribution of the exogenous variables that underlie economic performance, rather than indicating the existence of multiple equilibria. We therefore extend the model to test if there is a single or multiple relationship between exogenous variables and economic outcomes. In the case of multiple equilibria, our theory suggests that not only may we have multiple equilibria that are functions of some underlying, exogenous, variables but that the probability a country is close to a one steady state rather than the other is also a function of these underlying characteristics.

We begin by adding latitude as a fundamental exogenous variable that is clearly correlated with economic outcomes; countries at higher latitudes tend to be better off. When we test this single relationship against a model with two regimes we find that we can reject the single relationship model. The two regimes we find are a high level steady state which is independent of latitude and a low level steady state in which income rises with latitude. This is suggestive of a low level equilibrium (perhaps primarily agricultural) in which climate matters and a higher level equilibrium based on industry and services that is independent of climate. The probability of being in the high level equilibrium also rises with latitude, being close to zero throughout the tropics but rising sharply as we enter the temperate zones.

When we extend the model to include more meaningful geographical variables we continue to find multiple equilibria following a similar pattern. Cool, coastal countries, with heavy rainfall, evenly spread throughout the year, tend to be better off. In the low level equilibrium, income rises with these favorable characteristics, and they also increase the chance that a country is in a high level equilibrium in which income is independent of geography.

Our test involves comparing the hypothesis of a single equilibrium with that of two steady states. However, we assume that our geographical variables identify the "fundamental forces". It is always possible that a more complex single regime model, with extra exogenous explanatory variables, might adequately represent the data. Taken is this light, our results imply only that we can reject simple geographical determinism as the explanation of the "twin peaks" in the distribution of income levels across countries. It remains to be seen if other fundamental forces could explain the empirical distribution of income levels we observe.

Section 2 provides a brief discussion of how multiple equilibria can arise and the selection of exogenous variables that act as fundamental forces. While the model specification we use is apparently very simple, estimation and testing are not as straightforward as they appear at first sight, due to the fact that the model lacks the regularity conditions that underlie standard techniques. Our econometric methodology and the numerical techniques we employ are described in appendices, and our results are given in section 3.

2. The Model

There is a huge theoretical literature on models that have the potential to generate multiple equilibria and poverty traps (e.g. see Azariadis (1996)). We do not focus on a particular mechanism. Instead, if there is a poverty trap we argue that the GDP per capita, y, of a country can be written as

$$
\begin{aligned}
y &= y_1^*(x) + u_1 \text{ with probability } p(x) \\
y &= y_2^*(x) + u_2 \text{ with probability } 1 - p(x)
\end{aligned}
\qquad (1)
$$

where $y_1*(x)$ and $y_2*(x)$ represent two possible steady state levels of income for the country, given its exogenous characteristics, x. The disturbance terms, u_1 and u_2, represent short run deviations from steady state. Note that we allow the probability of being in a particular steady state to depend on x.

One way of justifying equation (1) is to think of economic development as a multidimensional stochastic dynamical system. If the deterministic part of the system is non-linear it can give rise to multiple steady states. If most stochastic shocks are small, the countries will spend most of their time in the neighborhood of one of these equilibria. However, rare large shocks can move a country between equilibria, and which equilibrium a country is in at a particular point in time is random and depends on the history of these large shocks. This interpretation of equation (1) is set out more formally in appendix 1.

It is natural to thinks of one of the equilibria in equation (1) as a high income equilibrium while the other is a low income poverty trap. The model set out in appendix 1 implies that countries can "jump" from the poverty trap to a high income level and vice versa, so that the probability p(x) represents the proportion of time a country is in equilibrium 1. An alternative way of thinking about equation (1) would be to assume that all countries start in the low level equilibrium and, at different times, jump irreversibly to the high level equilibrium. In this case, equation (1) represents a snapshot of a disequilibrium state and p(x) is the probability of a country with characteristics x still being in the low level equilibrium at the time of the snapshot. We do not try to distinguish between these two interpretations of equation (1) here since this would require an analysis of the dynamics of the system.

We can estimate the model of multiple equilibria set out in (1) and then test it against the model of a single steady state given by

$$y = y^*(x) + u \qquad (2)$$

A key issue is what variables should be included as exogenous in the vector x. As is made clear in appendix 1, we must be careful not to include any variables that may be endogenous. Many of the variables often used in growth models, such as the savings rate, the school enrollment rate, the population growth rates, and even policy variables such as tax rates and openness to trade, may themselves be endogenously determined. Including such variables may give the impression of a unique equilibrium relationship when in reality they are a function of the equilibrium being observed. “Fundamental forces” must be characteristics that determine a country’s economic performance, but are not determined by it.

Landes (1998) and Diamond (1997) emphasize the role of geography and culture as underlying factors that historically have determined the pace of economic development. Sachs and Warner (1997) and Gallup and Sachs and Mellinger (1999) have argued that geography has a significant impact on modern economic growth, while Hall and Jones (1999) use geography as an exogenous instrument in their estimation. On the other hand, Easterly and Levine (1997) argue that ethnic fragmentation (the opposite of our homogeneity measure) is a significant factor in Africa's growth tragedy.

In what follows we concentrate on physical geography as exogenous characteristics of a country that affects its equilibrium income level. We do not use cultural variables (with the exception of homogeneity of the population, which we find not to be significant). This is not a rejection of the role of culture in development. The problem with using culture as a determinant of income levels is that there are good arguments that cultural variables, such as religious affiliation, are endogenous in the long run and that their inclusion would bias our results. Note that cultural variables may be appropriate in a short or medium run analysis when they can be regarded as fixed in the time period under consideration.

Even in the case of geography, exogeneity may be called into question in the long run. While the physical geography of a given land mass may be fairly exogenous, exogeneity may fail because a country's borders may change. For example, both the United States and Russia underwent large expansions in the 18th and 19th centuries, while in this century the Austro-Hungarian and Ottoman empires disintegrated. Alesina and Spolagre (1997) and Bolton and Roland (1997) have suggested socio–economic mechanisms that lead to the endogenous formation of nations. If we take a long enough time frame, the composition of countries cannot be regarded as exogenous. We think of the time period under analysis here as being long run but with country borders exogenously given. That is, we have an implicit assumption that border changes happen relatively slowly with respect to the mechanisms that determine long run income levels.

Table 1 gives definitions, sources, and some summary statistics on the data we use. Table 2 shows a correlation matrix for our exogenous variables and income per capita; among our variables, latitude has the highest correlation with income per capita. We can think of latitude as a catchall for a variety of geographical indicators; it has a high degree of correlation with temperature and rainfall levels, as well as with their variability.[1] While we begin by using latitude as our exogenous variable, we eventually proceed to a richer model allowing the different geographical variables we have identified to have independent roles.

3. Estimation and Testing

Given our choices of exogenous factors, we wish to estimate the unique relationship given by equation (2) and the multiple equilibrium relationship in equation (1) and test between them. To make matters simple, we assume that the functions being estimated are linear and that the disturbance terms around the steady states in equations (1) and (2) are independent and normally distributed[2], though we allow each relationship to have its own variance.

While equation (2) is easy to estimate (we simply use ordinary least squares), the system of equations (1) poses some problems. The likelihood function corresponding to equation (1) is

$$L = p(x)\frac{1}{2\pi s_1}\exp[-\frac{1}{2}(\frac{y-\beta_1 x}{s_1})^2] + (1-p(x))\frac{1}{2\pi s_2}\exp[-\frac{1}{2}(\frac{y-\beta_2 x}{s_2})^2] \quad (3)$$

which says that with probability p(x), GDP per capita, y, is distributed normally with standard deviation s_1 around $\beta_1 x$, while with probability 1-p(x) it is distributed normally with standard deviation s_2 around $\beta_2 x$.

We can estimate the coefficients in (3) using maximum likelihood methods, but there are several difficulties with this approach. In addition to problems of estimation, there are difficulties in constructing valid hypothesis tests. These problems are discussed in detail in appendix 2. Our approach to overcome these problems of estimation and testing is to use the techniques suggested Feng and McCulloch (1994, 1996). The most important point to note is that we must place a lower bound on the variance of the disturbances in each equilibrium relationship in (1). This rules out the case where one equilibrium relationship simply gives exact outcomes, without any error, for a small number of countries while the other fits the rest of the data. This restriction is necessary for estimation purposes, but it also seems a plausible restriction on what we mean by multiple equilibria. We wish to find multiple stable steady states, around which countries congregate, not just argue that adding a small number of dummy variables, that exactly predict a number of outliers, improves the fit of the model.

In addition to the theoretical and conceptual problems of estimation and testing, the nonlinear maximum-likelihood grid search method we use for estimation, and Monte Carlo methods we use for testing, are sometimes sensitive to the numerical procedures employed. In appendix 3 we report the exact numerical methods that were used, to allow replication of our results, and we discuss the robustness of our results to alternative numerical procedures.

Our dependent variable is GDP per capita in 1985 from the Penn World Tables 5.6 (see Summers and Heston (1991)). For our purposes the year used is of little significance; using 1985 gives us the most complete dataset, with all 152 countries having data available for this year.

Table 3 reports estimates and tests the simple "twin peaks" model. The first column of table 3 is a regression of the log of 1985 GDP per capita on a constant. Log income per capita in 1985 had an average value of around 7.9, with a standard deviation of about 1. Column (2) of table 3 reports the results of fitting a model with two steady states using the maximum likelihood methods described in the appendices. The results suggest that about 85% of the countries in our sample are clustered around a low level poverty trap (regime I), though the standard deviation of these countries around this low level of income is quite large. On the other hand, 15% of countries are clustered much more tightly around a high level equilibrium (regime II). This clearly matches the “twin peaks” in described by Quah (1996).

The last row of table 3 gives likelihood ratio test of the multiple equilibria model against a single relationship. Due to the failure of the standard regularity conditions (as discussed in appendix 2) this likelihood ratio test does not have the usual chi-squared asymptotic distribution. Instead, we give the 5% critical value for the test, calculated by the Monte Carlo method explained in appendix 3 and reported in table 7. As already found by Bianchi (1997) and Paap and van Dijk (1998) using similar techniques, the test decisively rejects the single steady state model in favor of the alternative that countries are clustered around two distinct steady states.

The key question is whether or not the addition of extra exogenous explanatory variables makes these “multiple equilibria” disappear. We begin by introducing latitude as a proxy for climate. In table 4 we report the result of adding absolute latitude to the model as an exogenous variable. In the model with a single regime (column (1)) we find that income per capita is significantly greater at higher latitudes. On its own, latitude can explain about 40% of the variation in income levels across countries. We then estimate the two regime model (column (2)), assuming for now that the steady state level of income varies with latitude but that the probability of each regime is the same for every country. We again find a low level steady state and a high level steady state. In both these equilibria income level rises with latitude, though in the low level steady state the sensitivity to latitude is twice as great. However, for countries at high latitudes the two steady states are very close together, while in the tropics the two steady states are very far apart.

In column (2) we are assuming that the probability of each steady state is fixed; every country in this model has a 70% chance of being in the low level equilibrium and a 30% chance of being in the high level equilibrium. While this two regime model fits the data better than a single regime, the improvement in the log likelihood is not significant at the 5% significance level.

Therefore we cannot reject that income depends on a “fundamental force” in the shape of latitude. However, the results in column (2) of table 4 depend on the unrealistic assumption that the probability of each regime is the same across countries. Equatorial countries may not only have lower income when in the poverty trap, they may also find it harder to escape from the trap than countries with temperate zone. In column (3) of table 4 we report results where we allow the probability of each regime to vary with latitude. The functional form we use is

$$\text{Probability}(\textit{regime I}) = C(p_0 + \lambda \; \textit{latitude}) \quad (4)$$

where C is the cumulative normal distribution (ensuring the probability lies between zero and one) and p_0 and λ are parameters to be estimated.

The results in column (3) of table 4 are shown graphically in figure 1, where we plot income per capita against latitude. Again, we have a low level regime (regime I) where the level of income is very sensitive to latitude. We also have a high level steady state (regime II), but in this regime the level of income seems independent of latitude, since the latitude variable is not statistically significant. The probability of being in the low level steady state is high for countries at the equator but falls with latitude. In fact, the probability of being in the low level regime is close to one at the equator, but declines to around 25% at latitude 60^0, which corresponds to Northern Europe. The improvement in fit (as measured by the log-likelihood) when we allow the regime probability to vary with latitude is quite large and we can reject the model of a single regime against the alternative of two regimes with the probability of each regime varying with latitude.

Latitude is really only a proxy for some underlying geographical variables. We wish to introduce our geographical variables directly. Unfortunately, our statistical methodology is not very good at handling a large number of independent variables – a large number of degrees of freedom tends to push the estimates towards the boundary where some explanatory variables become "dummy variables" and are used to explain perfectly a small number of countries. In table 5 we report simple OLS estimates in which we explain income per capita with a range of exogenous variables. We begin with all our variables in column (1); parsing down by sequential elimination of variables that are not statistically significant at the 5% levels leaves the four geographical variables whose coefficients are reported in column (2): percentage of land within 100km of the coast, log of maximum average monthly temperature, log of average monthly rainfall, and the log of the standard deviation of rainfall month to month over the year. These four variables account for over 60% of the variation in income per capita. The richer countries tend to be cool and coastal, with high but steady rainfall over the year. Note that favorable geography tends to increase a country's the income level, but the simple O.L.S. estimates may be a combination of a rising equilibrium level of income as geography improves and a higher chance of jumping to a high level equilibrium with better geography.

Having selected our exogenous variables, we could follow our theoretical framework and estimate a two regime model in which all the parameters of the model vary across regimes, and the probability of being in a regime depends on all the geographical variables. However, when we try this the maximum likelihood estimate again converges to the boundary where one regime has a minimal error term - the model has too many parameters relative to the number of countries being explained; it always goes to a "dummy variable" solution where one regime is used to model a group of countries exactly.

To avoid this problem, we impose the restriction that in the high level equilibrium the level of income is fixed and is independent of geography. In addition, we assume that the relative coefficients on our geography variables are the same in low level regime and when determining the regime probability.[3] The model we actually estimate is:

$$y_i = a_1 + (\sum_{j=1}^{4} \beta_j x_{ji}) + \varepsilon_{1i} \quad \textit{with proability } p(x_i)$$
$$y_i = a_2 + \varepsilon_{2i} \quad \textit{with probability } 1 - p(x_i) \qquad (5)$$
$$\text{where } p(x_i) = C(p_0 + \lambda(\sum_{j=1}^{4} \beta_j x_{ji}))$$

where i represents a country, j a geographical variable and C is again the cumulative normal distribution. These restrictions clearly make it more difficult for the multiple equilibrium model to perform well as compared with completely unrestricted estimation. The results of this estimation are shown in column (2) of table 6. The estimates of the effect of geography in the low level equilibrium are very similar to those in the simple OLS model reported in column (1).

However, the likelihood ratio test again rejects the model in column (1), with a single relationship, in favor of our multiple equilibrium model. This multiple equilibrium model has a high level equilibrium which is independent of geography (given by the regime II constant), and a low level equilibrium which is very sensitive to geography. Good geography not only improves a country's steady state income in the low level equilibrium, it also raises the probability that a country is in the high level equilibrium. The probability of being in the high level equilibrium improves slowly with improvements in geography. For a wide range of countries there is a positive probability of being in each equilibrium and this model has real multiple equilibria.

5. Conclusion

We have found empirical evidence of the existence of a poverty trap. There appears to be a high level equilibrium that is the same for all countries. In the low level equilibrium, cool coastal countries, with high year-round rainfall, have high incomes. Countries with favorable geography have quite high income levels in the low level equilibrium and they find it relatively easy to jump to the high level equilibrium. On the other hand, hot landlocked countries, with low or very seasonal rainfall, have very low levels of income in the low level steady state. In addition, it is very hard for them to break out of this poverty trap and reach the high level equilibrium. Geography matters. But we reject simple geographical determinism in favor of a multiple equilibrium model.

While we find evidence of multiple equilibria, we have nothing to say about why these multiple equilibria come about, or how countries can make the transition from one equilibrium to the other. To do this requires a more structural, dynamic, model that allows for the interaction between the endogenous social and economic forces that determine human development and economic growth.

References

Alesina A. and Spolagre E. (1997), "On the Number and Size of Nations", *Quarterly Journal of Economics*, Vol. 112, pp 1027-1056.

Azariadis C. (1996), "The Economics of Poverty Traps Part One: Complete Markets*," Journal of Economic Growth*, Vol. 1, pp 449-486.

Bair F.E. (1990), *Countries of the World and Their Leaders: Yearbook 1990*, Detroit, Gale Research Inc.

Berthelemy, J. and Varoudakis, A. (1996), "Economic Growth, Convergence Clubs, and the Role of Financial Development," *Oxford Economic Papers*, Vol. 48, pp 300-328.

Bianchi M. (1997), "Testing for Convergence: Evidence from Non-Parametric Multimodality Tests," *Journal of Applied Econometrics*, Vol. 12, pp 393-409.

Bolton P. and Roland G. (1997), "The Breakup of Nations: A Political Economy Analysis," *Quarterly Journal of Economics*, Vol. 112, pp 1057-1090.

Central Intelligence Agency (1985), *World Factbook 1985*, Washington D.C. CIA.

Central Intelligence Agency (1997), *World Factbook 1997*, Washington D.C. CIA

Desdoigts A. (1999), "Patterns of Economic Development and the Formation of Clubs," *Journal of Economic Growth*, Vol. 4, pp 305-330.

Diamond J. (1997), *Guns, Germs and Steel*, New York, W.W. Norton.

Durlauf, S. and Johnson, P. (1995), "Multiple Regimes and Cross-Country Growth Behavior," *Journal of Applied Econometrics*, Vol. 10, 365-384.

Easterly W. and Levine R. (1997), "Africa's Growth Tragedy: Policies and Ethnic Divisions," *Quarterly Journal of Economics*, Vol. 112, pp 1251-1288.

Feng, Z. and McCulloch, C. (1994), "On the Likelihood Ratio Test Statistic for the Number of Components in a Normal Mixture with Unequal Variances," *Biometrics*, Vol. 50, pp 1158-62.

Feng, Z. and McCulloch, C. (1996), "Using Bootstrap Likelihood Ratios in Finite Mixture Models," *Journal of the Royal Statistical Society,* Series B, Vol. 58, pp 609-17.

Feve P and Le Pen Y. (2000), "On Modeling Convergence Clubs," *Applied Economics Letters*, Vol. 7, pp 311-314.

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

Hall R.E. and Jones C.I. (1999), "Why do Some Countries Produce so Much more Output per Worker than Others?" *Quarterly Journal of Economics*, Vol. 114, pp 83-116.

Kifer Y. (1988), *Random Perturbations of Dynamical Systems*, Boston, Birkhauser.

Kurian R. (1991), *The New Book of World Rankings,* New York, Facts on File.

Landes, D. (1998). *The Wealth and Poverty of Nations: Why Some Are So Rich and Some So Poor?*, New York, W.W. Norton.

McLachlan, G. (1987), "On Bootstrapping the Likelihood Ratio Test Statistic for the Number of Components in a Normal Mixture," *Applied Statistics*, Vol. 36, pp 318-24.

Paap R. and R.K.van Dijk (1998), "Distribution and Mobility of Wealth of Nations," E*uropean Economic Review*, Vol. 42, pp 1269-1293.

Quah, D. (1996), "Twin Peaks: Growth and Convergence in Models of Distribution Dynamics," *Economic Journal*, Vol. 106, pp 1045-55.

Quah, D. (1997), "Empirics for Growth and Distribution: Stratification, Polarization, and Convergence Clubs," *Journal of Economic Growth*, Vol. 2, pp 27-59.

Sachs, J. and Warner, A. (1997), "Fundamental Sources of Economic Growth," *American Economic Review*, Pap. and Proc.,Vol87, pp 184-8.

Summers R and Heston A. (1991), "The Penn World Table (Mark 5): An Expanded Set of International Comparisons, 1950-1988," *Quarterly Journal of Economics*, Vol. 106, pp 327-368.

Appendix 1

We assume that the economic system comprises a vector of exogenous variables, x, and endogenous variables, z. The discussion of which variables should be considered exogenous and which endogenous is postponed until section 3. The evolution of the endogenous variables at time t is assumed to be given by

$$z_t = f(z_{t-1}, x) + \varepsilon_t \qquad (1.1)$$

where longer lags can be incorporated by redefining lagged variables as current variables (at the expense of increasing the size of the parameter space). The steady states, or equilibria, of the deterministic part of this system are given by

$$g(x_t) = \{z^* : f(z^*, x) = z^*\} \qquad (1.2)$$

In general, equation (2) may give rise to many different steady states. However, if the function f is linear we can write the system as

$$z_t = Az_{t-1} + Bx + \varepsilon_t \qquad (1.3)$$

and, if A is invertible, the steady state is unique and is given by

$$z^* = (I - A)^{-1} Bx \qquad (1.4)$$

and provided the matrix A has a maximal eigenvalue less than one, the steady state is stable and we can write

$$z_t = z^*(x) + u_t \text{ where } u_t = \sum_{s=0}^{\infty} A^s \varepsilon_{t-s} \qquad (1.5)$$

In a linear system we either have one steady state, or if A is not invertible, no steady states or an infinite continuum of steady states. In order to generate distinct, isolated, steady states we require the function f to be nonlinear.

One approach to modeling multiple steady states would be to estimate a relationship like (1.1) directly, allowing for non-linearities, and solve equation (1.2) for its steady states. This could be done using a system of cross-country regressions (for example, as a vector autoregression), explaining the evolution of all the endogenous variables in the development process. The disadvantage of this approach is that it requires the construction of a complete model of economic and social development, a daunting undertaking. Instead of estimating the full model, it is possible to take a short cut and estimate only the growth equation for income per capita. However this may be misleading. It is easy to construct examples in which the single equation in (1.1) that generates economic growth is linear, but non-linearities elsewhere in the system generate multiple equilibria, including multiple steady state values of the income level. Similarly, if non-linearities are found in the income growth equation, they must be combined with other parts of the system to determine if there are multiple equilibria.

This gives us an incentive to find alternative methods of testing for multiple equilibria that do not require estimation of the entire system. One method is to note that the random dynamical system (1.1) represents the short term evolution of the system and we can consider its behavior on longer time scales. Consider a non-linear model for which the deterministic part of equation (1.1) generates multiple steady states. Further, let us simplify matters by assuming that, for each level of the exogenous variables, the set of locally stable attractors of the deterministic part of equation (1.1) consists of at most two steady states. This means that for almost all initial conditions the deterministic part of the system converges to one of these steady states.

If the noise in equation (1.1) in small, we can image that the system will converge to a steady state and then oscillate in the neighborhood of that steady state. If the noise is sufficiently small, the oscillations near the steady state can be confined to a neighborhood in which the dynamical system is approximately linear, and we can apply equations (1.3), (1.4) and (1.5) to model the local behavior of the system. Hence we can write

$$\begin{aligned} z_t &= z_1^*(x) + u_{1t} \qquad \text{if } z_0 \in Z_1(x) \\ z_t &= z_2^*(x) + u_{2t} \qquad \text{if } z_0 \in Z_2(x) \end{aligned} \qquad (1.6)$$

If the system starts at time zero in the basin of attraction $Z_1(x)$ it converges to the steady state $z_2^*(x)$ and at time t will large be in the neighborhood of this steady state. Similarly, if it starts in the basin of attraction $Z_2(x)$ it converges to $z_2^*(x)$. The boundary between these two basins of attraction is a unstable separatrix; eventually the random noise in the system will force the model into one of the basins and lead to convergence to a steady state.

The system (1.6) provides a feasible method of estimation, however, while the two steady states are functions of the exogenous variables, which basin of attraction the initial condition lies in depends on all the endogenous variables in the vector of initial conditions. Equation (6) is essentially a model of convergence clubs (subtracting z_0 from both sides of the equations in (6) gives a growth model) an economy converges to its local "club," which "club" it is a member of depending on both its exogenous factors and its initial position. Again, however, in this approach we have to model all the endogenous variables in the development process as part of its "initial position" in order to decide to which club a country belongs.

We can view equation (1.6) as the medium run behavior of the model. Most of the time the shocks to the system are small, so that economies converge to the steady state in whose basin of attraction they lie. Now suppose at longer time scales large shocks may occur. We can regard the system (1.6) as a two state process. The system is usually close to one steady state or the other but when a large shock occurs it is possible that the endogenous variables in the model jump out of one basin of attraction and into the other, after which they converge to the neighborhood of the new steady state. After a sufficiently long period of time the system becomes ergodic and losses its dependence on initial conditions; we can describe long run behavior as a stationary equilibrium probability distribution over neighborhoods of the two states.

These intuitive ideas can easily be formalized. Equation (1.1) gives rise to a random dynamical system and can be regarded as a Markov process, mapping points in the state space of the endogenous variables into probability measures over that state space. Under some technical regularity conditions[4], after sufficient time has passed, the empirical cross section distribution of independent observations drawn from such models approximates the invariant probability measure on the state space. Kifer (1988) shows that, provided large shocks are sufficiently rare in the error term in (1.1), the invariant probability measure puts almost all its weight in the neighborhoods of the attractors (i.e. the two stable steady states) of the deterministic part of the system.

The invariant measure which describes the long run behavior of the system is an equilibrium probability distribution which depends on the transition probabilities of moving from one steady state to the other; if a steady state has a relatively large basin of attraction it will be more difficult to exit, and it will have a higher weight in the stationary distribution. The size of the basin of attraction may depend on the exogenous variables as shown in (1.6); this implies the long run probability of the system being in a particular steady state also depends on the exogenous variables. Formally, this gives rise to the "reduced form" model

$$\begin{aligned} z_t &= z_1^*(x) + u_{1t} \qquad \text{with probability } p(x) \\ z_t &= z_2^*(x) + u_{2t} \qquad \text{with probability } 1 - p(x) \end{aligned} \tag{1.7}$$

where both the steady states, and the probability of being in each steady state, depend on the exogenous variables. The advantage of this approach is that in the very long run, the endogenous variables can be written as depending on the exogenous variables alone.

Clearly, all three approaches to the problem are valid under the assumptions set out above. However, estimating the full multidimensional development process in the short run dynamic relationship does not appear to be a feasible proposition at present. Estimating a model with "convergence clubs" is well established in the literature. However, our approach here will be to use the framework set out in (1.7). Note that compared to (1.6) equation (1.7) ignores the information contained in "initial conditions," in practice lagged values of the endogenous variables. The advantage we achieve by doing this is that we do not have to specify the (potentially very large) set of endogenous variables in the model and estimate the boundary between the two steady states.

While equation (1.7) specifies a system of equations, determining all the endogenous variables, we can simply pick out the income per capita (y) to give the model

$$\begin{aligned} y_t &= y_1^*(x) + u_{1t}^y \text{ with probability if } p(x) \\ y_t &= y_2^*(x) + u_{2t}^y \text{ with probability } 1 - p(x) \end{aligned} \qquad (1.8)$$

Equation (1.8) corresponds with equation (1) in the main text.

Appendix 2

We wish to estimate equation (1) using maximum likelihood techniques. However, if we allow one of the relationships to collapse, to completely explain a small set of data points while the variance of its error term goes to zero, the model's usual continuous probability density function collapses to a discrete atom of probability, causing the likelihood function to become unbounded.

Essentially, this means that the m parameters in one of the relationships in (8) are used to completely explain m data points, without noise, while the other relationship is fitted to the rest of the data. If we estimate (1) with no constraint on the error variances, the maximum likelihood estimator will always converge to this boundary.

However, this is really a model of a single relationship, with m special cases being explained in ad hoc way. It is similar to including a number of dummy variables to explain outliers from a particular relationship. Conceptually, our notion of multiple equilibria is that there are more than one basin of attraction, with a substantial number of countries lying in each basin. If we wish to estimate such a model we must place a lower bound on standard deviation of the regime disturbances in (1) to rule out this "dummy variables" estimate.

Imposing such a constraint allows us to estimate the relationship (1). We really want to find a greatest local maximum that corresponds to non-zero variance of the error terms in (1). In practice, the likelihood surface is quite irregular even away from the boundary, with multiple local maxima. To try to ensure convergence to the greatest local maximum we therefore undertake a grid search over initial starting values. This is discussed in more detail in appendix 3.

Having found the maximum likelihood estimate of equation (1), we can construct the likelihood ratio statistic $-2(\log(\lambda_r / \lambda_u))$, where λ_r is the likelihood of the restricted (one regime) model and λ_u is the likelihood of the unrestricted (two regime) model. The two models are nested, but the parameter restrictions imposed by the hypothesis of a single relationship are not straightforward. Under the null of a single relationship, the true parameters in equation (1) lie on the boundary of the parameter space and some the variables in the model become irrelevant, making their associated parameters unidentified. It follows that the model fails the usual regularity conditions and the likelihood ratio statistic will not have the usual chi-squared asymptotic distribution.

We can overcome this problem by generating appropriate critical values for the likelihood ratio test via Monte Carlo methods (as in, for example, McLachan (1987)). This involves a Monte Carlo study, using random data generated from the estimated relationship given by equation (2). Using this data we can calculate critical values of the likelihood ratio statistic under the null hypothesis that the estimated model with one relationship is the true model. In doing this we must be careful to use the same approach in estimating the two regime model in the Monte Carlo study as used in our actual estimator. In particular, as pointed out by Feng and McCulloch (1994), the critical values of the likelihood ratio test depend on the cut-off used to constrain the minimum allowable regime variance in the two regime model.

For each value of the cut-off variance we can construct a valid test at the appropriate significance level, but the simulations in Feng and McCulloch (1996) suggest that imposing a higher cut-off can substantially increase the power of the test. Intuitively, a very low cutoff allows the estimated model to approach the boundary where we have the "dummy variable" model, and produces a very high log likelihood for the two-regime model, whatever the true underlying structure of the data. This points towards using a relatively high cut-off for the minimal variance allowed in the two regime model, though we do not wish to make it so large as to exclude the "true" relationship. In our estimation we limit the standard deviation of each regime to be at least 0.05 (with income measured in natural logs) which corresponds to countries lying around their steady state GDP per capita with a standard error of at least 5%. We also experimented with limits of 0.1 and 0.01 (corresponding to standard deviations in income level of 10% and 1% respectively) but found that tightening the constraint sometimes led us to exclude the previously found maximum, while weakening the constraint never produced greater local maxima but instead led the maximum likelihood estimate to converge to the boundary.

In addition to our constraint on the standard deviation of each regime, the probability of each regime must lie between zero and one. To ensure that these constraints hold, we actually replace the probabilities and standard errors in (3) with

$$p(x) = C(p_0 + \lambda x)$$

$$s_i = 0.05 + 0.95\, C(\sigma_i)$$

where C is the cumulative normal distribution. This constrains the probability of regime I to lie between zero and one while the standard deviation of each regime is between 0.05 and one. The upper bound on the standard deviations of the regimes is not actually required, but we find that, in practice, imposing such a bound does not affect the estimates. In reporting the results we report estimates of derived probabilities and standard deviations with standard errors calculated by the delta method.

Appendix 3

In order to find the maximum likelihood estimate we undertake a grid search over initial conditions. Experimentation showed that the crucial parameters to search over are the probability of each regime, the standard deviation of the two regimes, and the initial gap between the intercepts of the two regimes. For the other parameters we use as initial conditions the estimates from the single regime model. The starting points for the two intercepts were positioned to be equally spaced around the intercept of the single regime model. The starting values (with number of options in brackets) used were:

intercept difference (3)	d: 0.5, 1, 2.
standard error (5)	σ_i : –1.282, –0.524, 0, 0.524, 1.282.
probability intercept (7)	p_0 : -7, –1.282, –0.524, 0, 0.524, 1.282, 7.
probability slope (3)	λ : -1, 0, 1.

Note that the initial starting conditions for the two regimes are symmetrical, so there is no gain from allowing a negative intercept differential in the starting conditions.

In the Monte Carlo study we first generated random data points based on the single regime model (the null hypothesis) assuming a normal distribution of errors, with the estimated standard deviation, around the relationship. Then, using the same grid search program as above, we find the maximum likelihood estimate of the two regime model and calculate the likelihood ratio statistic, testing the two regime model against a single regime. Repeating this process 500 times using randomly generated data gives us the empirical distribution of the likelihood ratio statistic under the null.

Table 1: Variable Sources, and Summary Statistics

DEFINITION	SOURCE	MEAN	MEDIAN	S.D.	MIN	MAX	OBS
GDP per Capita:GDP per capita in 1985 in international prices	Penn World Tables 5.6, Summers and Heston (1991)	4423	2564	4424	299	19648	152
Latitude: Absolute latitude of approximate center of country	Central Intelligence Agency (1997)	23.24	19.50	15.93	0.00	65.00	152
Landlocked: dummy variable equaling 1 for landlocked countries and zero for all others	Our construction	0.17	0.00	0.38	0.00	1.00	152
% Land Coastal: % land within 100km of coast	Gallup and Sachs and Mellinger, (1999)	0.46	0.35	0.40	0.00	1.00	148
Max Temp: Average maximum temperature of capital city in degrees centigrade	Bair (1990)	24.19	25.97	7.44	4.17	37.22	120
Rainfall: average rainfall in capital city in centimeters	Bair (1990)	10.01	7.90	7.86	0.34	40.26	120
SD Max Temp: standard deviation of maximum temperature in capital at 3-month intervals from the yearly mean in degrees centigrade	Derived from Bair (1990)	4.89	3.62	3.89	0.28	16.87	120
SD Rainfall: standard deviation of rainfall at monthly intervals from yearly mean in cm	Derived from Bair (1990)	7.30	4.78	8.86	0.31	56.91	120
Homogeneity: percentage of largest single group with of ethnic, religious and linguistic homogeneity	Kurian (1991)	59.69	62.50	29.67	7	100	124

Table 2
Correlations

	GDP per capita	latitude	log rainfall	log s.d. rainfall	log max temp	log s.d. max temp	% land coastal	land locked	homogeneity
GDP per capita	1.000								
latitude	0.636	1.000							
log rainfall	-0.241	-0.397	1.000						
log s.d. rainfall	-0.621	-0.628	0.743	1.000					
log max temp	-0.537	-0.812	0.273	0.538	1.000				
log s.d. max temp	0.414	0.812	-0.557	-0.582	-0.597	1.000			
% land coastal	0.347	0.012	0.207	0.055	0.087	-0.219	1.000		
land locked	-0.232	0.066	-0.009	0.048	-0.110	0.127	-0.479	1.000	
homogeneity	0.546	0.501	-0.248	-0.425	-0.386	0.306	0.415	-0.193	1.000

Table 3
"Twin Peaks"

Dependent Variable: Log GDP per capita 1985, 152 observations.

	Single Regime	Two Regimes	
		Fixed Regime Probabilities	
		Regime I	Regime II
Constant	7.891 (0.084)	7.612 (0.097)	9.411 (0.052)
Standard deviation of regime disturbance	1.041	0.876	0.180
Probability of regime I		0.845 (0.043)	
R^2	0		
Log likelihood	-221.4	-208.7	
Likelihood ratio test of two regimes vs. one regime [5% critical value]		26.0 [11.6]	

Standard deviation of the regime disturbance constrained to exceed 0.05.

Table 4
Latitude and Income Level

Dependent Variable: Log GDP per capita 1985, 152 observations.

	Single Regime	Two Regimes		Two Regimes	
		Fixed Regime Probabilities		Regime Probabilities Vary with Latitude	
		Regime I	Regime II	Regime I	Regime II
Constant	6.922 (0.118)	6.392 (0.178)	8.136 (0.331)	6.979 (0.137)	9.261 (0.145)
Latitude	0.041 (0.003)	0.049 (0.005)	0.025 (0.010)	0.035 (0.007)	0.003 (0.003)
Standard deviation of regime disturbance	0.806	0.574	0.538	0.809	0.102
Probability I Intercept		0.532 (0.328)		3.588 (0.943)	
Effect of latitude on probability of regime I				-0.071 (0.022)	
Probability of regime I at equator.		0.703 (0.114)		1.000 (0.001)	
Probability of regime I at latitude 60.		0.703 (0.114)		0.255 (0.158)	
R^2	0.401				
Log likelihood	-181.9	-175.1		-166.9	
Likelihood ratio test of two regimes vs. single regime [5% critical value]		13.8 [17.9]		30.2 [21.8]	

Standard deviation of the regime disturbance constrained to exceed 0.05.

Table 5
Geography and Income Level: OLS Estimates

Dependent Variable: Log GDP per capita 1985.

Constant	8.199 (1.106)	10.237 (0.867)
Log Rainfall	0.423 (0.133)	0.387 (0.139)
Log Standard Deviation of Rainfall	-0.652 (0.114)	-0.724 (0.125)
Log Maximum Temperature	-0.277 (0.312)	-0.734 (0.285)
Percent of Land within 100km of the coast	0.696 (0.229)	0.805 (0.172)
Log of Standard Deviation of Maximum Temperature	0.083 (0.140)	
Landlocked	-0.249 (0.194)	
Homogeneity	0.321 (0.325)	
Latitude	0.010 (0.011)	
R^2	0.663	0.594
N	102	115
Log likelihood	-92.14	-112.65

Table 6
Geography and Income Level

Dependent Variable: Log GDP per capita 1985, 115 observations.

	Single Regime	Two Regimes: Regime II independent of geography
Regime I Constant	10.237 (0.867)	10.213 (0.479)
Regime II Constant		9.448 (0.034)
Log Rainfall	0.387 (0.139)	0.334 (0.120)
Log Standard Deviation of Rainfall	-0.724 (0.125)	-0.639 (0.136)
Log Maximum Temperature	-0.734 (0.285)	-0.749 (0.152)
Percent of Land within 100km of the coast	0.805 (0.172)	0.801 (0.163)
Standard deviation of regime I disturbance	0.659	0.652
Standard deviation of regime II disturbance		0.092
Probability I Intercept (p_0) Effect of geography on probability of regime I (λ)		-2.477 (1.186) -2.034 (0.711)
R^2	0.594	
Log likelihood	-112.7	-100.9
Likelihood ratio test of two regimes vs. single regime [5% critical value]		23.6 [10.1]

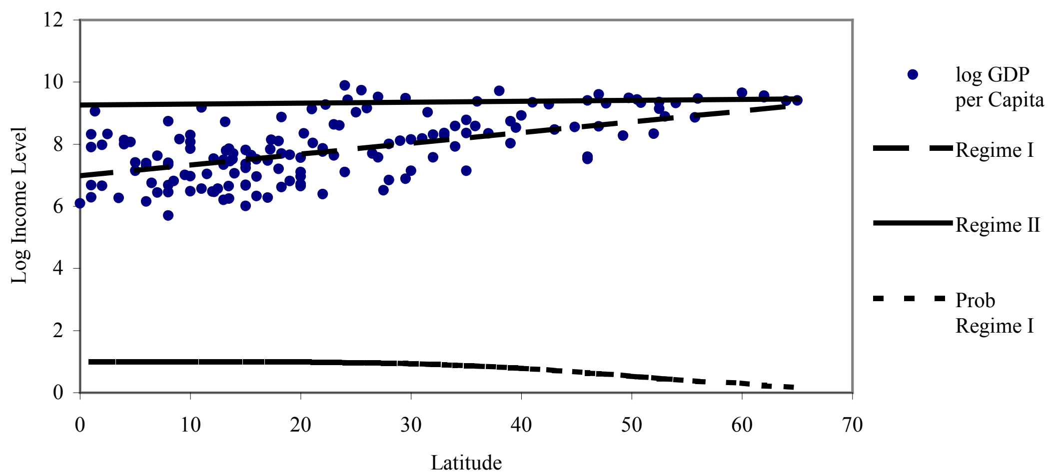
Standard deviation of the regime disturbance constrained to exceed 0.05.

Table 7
Critical Values for the Likelihood Ratio Test of Two Regimes Versus One Regime

Income Level Depends on:	Intercept	Latitude	Latitude	Geography: Regime II Independent of Geography
Regime Probabilities	Fixed Regime Probabilities	Fixed Regime Probabilities	Regime Probabilities Vary with Latitude	Regime Probabilities Vary with Geography
Constraint on Regime Disturbance	Standard Deviation ≥ 0.05	Standard Deviation ≥ 0.05	Standard Deviation ≥ 0.05	Standard Deviation ≥ 0.05
N	152	152	152	115
Significance Level				
10%	10.14	15.92	19.54	7.76
5 %	11.66	17.94	21.76	10.05
1 %	15.13	21.14	26.51	15.53

Critical values based on the empirical distribution of the likelihood ratio statistic under the null of one regime (500 repetitions).

Figure 1
Latitude and Income Per capita



Endnotes

[1] Using a principal components analysis of the geographic variables produces a first principal component that captures about 46% of the geographical variation across countries, but does markedly less well than latitude in explaining income levels.

[2] It is tempting to take a more general approach, and not specify the distribution of the error terms. However, a model with distributions around two steady states is observationally equivalent to one with a bimodal distribution around a single steady state. In order to identify the two models we have to impose some distributional assumptions on the distribution of the error terms.

[3] Extending the model to allow geography to affect a country's income level in the high–level equilibrium, but keeping the relative weights on geography the same across the two equilibria, produced an insignificant coefficient on geography in the high level equilibrium.

[4] These require that the randomness in the system, the mapping from the state space into probability measures, be continuous in the total variation norm, and tight, so that the sequence of outcomes generated by the system remains in a bounded area with probability one.