Nonlinearities in Growth: From Evidence to Policy

Ethan B. Cohen-Cole

Steven N. Durlauf

Giacomo Rondina

University of Wisconsin at Madison[1]

April 28, 2005

Abstract

This paper considers the question of how one can translate evidence of nonlinearities and threshold effects in growth into policy recommendations. We argue that the current evidence of these effects, while important in terms of scholarly debates, does not readily lend itself to policy evaluation. The reasons for this are two-fold. First, the existing evidence on nonlinearities is relatively difficult to integrate into a common coherent view. Different models of nonlinearity appear in different papers; these models are often nonnested and do not present a clear alternative to linear growth models. Second, we argue that the econometric evidence of nonlinearities is often developed in ways that do not allow one to explicitly examine the effects of alternative policies on growth. We describe some recent econometric methods that can address these problems.

---

[1] Department of Economics, 1180 Observatory Drive, Madison WI, 53706-1393. Corresponding Author: Durlauf. The John D. and Catherine T. MacArthur Foundation and University of Wisconsin Graduate School have provided financial support.

> Economics, therefore, cannot give us any valuable information concerning social reform. Only a pseudo-economics can seek to offer a background for rational economic planning. Truly scientific economics can merely help to reveal the driving forces of economic development though different historical periods. It may help us to foresee the outlines of future periods, but it cannot help us to develop and put into operation any detailed plan for the new period.
>
> Karl Popper, *The Poverty of Historicism* (p. 48-49)

1. Introduction

The purpose of the paper is to explore some issues involved in translating statistical evidence of nonlinearities into policy recommendations. Of course, actual policy decisions cannot be reduced to mappings from statistical statements to particular policy choices. But for the purposes of understanding how statistical exercises should influence policy evaluation, it is appropriate to assume that there is such a mapping and ask how it should be constructed.

From the perspective of new growth theories, there are good reasons to expect that nonlinearities exist in cross-country data. One reason for this is that models ranging from the now-classic Azariadis and Drazen (1990) to the recent Howitt and Mayer-Foulkes (2002) provide microeconomic mechanisms by which multiple steady states and convergence clubs can emerge in a cross-section of countries. The presence of multiple steady states is of particular interest to policymakers since these raise the possibility of development traps which require government interventions if a country is to escape from one. Nonlinearities can also arise for more mundane reasons, such as deviations of the aggregate production function in the neoclassical growth model from the Cobb-Douglas specification that is conventionally assumed; Duffy and Papageorgiou (2000) provide evidence in favor of a constant elasticity of substitution specification and show how this implies nonlinearities should be present in cross-country growth regressions.

Within the modern empirical growth economics literature, there is a growing body of evidence that nonlinearities in the growth process are in fact present. This evidence falls into two main categories.[2] First, there is evidence that different countries obey different growth models. This type of analysis, of which Bloom, Canning, and Sevilla (2003) Canova (2004), Desdoigts (1999), Durlauf and Johnson (1995), Johnson and Takeyama (2001), Kourtellos (2003a,b), Masanjala and Papageorgiou (2004a), Papageorgiou (2002), and Tan (2004) are examples, attempts to identify groups of countries that obey similar growth processes. As such, this approach primarily draws from classification methods which attempt to sort observations into common statistical models. Second, a number of authors have moved beyond linear growth models to consider environments in which regression parameters are functions of various economic characteristics. Examples of work of this type include Banerjee and Duflo (2003), Durlauf, Kourtellos, and Minkin (2001), Fiaschi and Lavezzi (2003), Kalaitzidakis, Mamuneas, and Stengos (2000) and Liu and Stengos (1998). These different papers use a wide range of statistical methods; see Durlauf, Johnson and Temple (2004) for a survey of empirical growth research as a whole as well as the location of nonlinear growth studies in this large literature. Beyond studies that have attempted to identify nonlinearities explicitly, it is now common practice in growth empirics to add the products of variables as additional regressors in linear models; this practice is an ad hoc way of introducing nonlinearities. While our arguments will apply at least as much to these ad hoc formulations as well as to more formal analyses, we will not focus on the latter in referring to the empirical literature.

While the body of evidence generated by this empirical work seems persuasive that some nonlinearities matter for understanding growth, there has not been any systematic direct investigation of the implications of nonlinearities for policy evaluation.

One can interpret some of various policy debates in the growth literature as depending on whether nonlinearities are present, but these debates focus more on the validity of the claim of nonlinearity per se than on how policy choices should be altered if the nonlinearity is present. For example, the disagreement between Burnside and Dollar

---

[2] In this discussion, we elide the difference between papers that search for nonlinearities in growth rates versus those that explore nonlinearities in per capita output levels. While these differences matter of course for policy evaluation, they have not been carefully

(2000) and Easterly, Levine and Roodman (2004) on the efficacy of foreign aid in improving growth performance may be interpreted as a disagreement about whether the regression coefficient on foreign aid depends on the level of policy or not. The first of these papers, which argues that in the presence of good policies, foreign aid can affect growth, has been widely used in policy circles as evidence for the need for policy reform in developing countries. While we are sympathetic to the need for policy reform, our concern is with the extent to which the findings in such papers justify claims of this type.

In this paper, we discuss the relationship between policy evaluation and growth nonlinearities. The main claim in this paper is that the available evidence on nonlinearities in growth does not provide a basis for policy recommendations, even when, as in the debate of the effects of foreign aid, the focus of the analysis is explicitly on policy variables. To be clear, we do not question the presence of nonlinearities per se. Rather, we argue that this evidence, as currently constituted cannot be used to say much about how growth policies should be constructed. To some extent, the reasons why policy inferences cannot be made in the context of nonlinearities will apply generally to the ways in which empirical evidence is adduced in economics; in this respect the literature on nonlinearities and growth is simply an example of some of the limitations of much current econometric practice.

Section 2 of the paper discusses the relationship between conventional econometric practice and policy evaluation for growth regressions. The discussion will focus on the relationship between evidence of growth nonlinearities, as conventionally presented, and policy analysis. Section 3 discusses a second problem with the existing body of evidence on nonlinearities: the presence of model uncertainty in growth regression specifications. The failure of standard analyses of growth nonlinearities to account for model uncertainty not only has implications for policy evaluation, but calls into question some of the evidence that nonlinearities are present. Section 4 discusses some issues of observational equivalence that have yet to be addressed in the nonlinear growth literature. Section 5 makes some suggestions on how to provide better links between evidence of growth nonlinearities and policy. Section 6 concludes.

---

distinguished in most empirical work; see Durlauf, Johnson and Temple (2004) for

2. Statistical decision theory and econometrics

The first reason why one cannot integrate evidence of growth nonlinearities into policy advice stems from the fact that econometric evidence, as conventionally presented in growth contexts, does not translate into the analysis of decision problems under uncertainty. This is a general problem with conventional econometric analyses, as argued by Chamberlain (2001), Sims (2002), Brock, Durlauf, and West (2003) among others; the latter paper focuses specifically on growth issues. In this discussion, we follow the framework in Brock, Durlauf, and West (2003) which in turn follows standard statistical decision theoretic arguments, e.g. Berger (1980).

For expositional purposes, we will work with a simple form of nonlinearity; this is done in order to draw clear comparisons between standard econometric practice and the evaluation of decision problems. To do this, define the canonical cross-country growth regression developed by Barro (1991) and Mankiw, Romer and Weil (1992) as

$$g_i = p_i\delta + Z_i\gamma + \varepsilon_i \tag{1}$$

Where g_i is real per capita growth of country i across some fixed time interval, p_i is the policy instrument of interest, Z_i is a set of additional regressors that reflect the growth determinants that a modeler chooses to control, and ε_i is an error. One form of nonlinearity in the growth process that matters for policy is described by

$$g_i = p_i\delta + f\left(p_i, X_i\right)\delta_n + Z_i\gamma + \varepsilon_i \tag{2}$$

where $f\left(p_i, X_i\right)$ is some nonlinear function; X_i allows the policy nonlinearity to interact with various country-specific characteristics. For expositional purposes, suppose

---

discussion of an analogous problem that arises in linear growth regression models.

that the functional form of $f\left(\cdot,\cdot\right)$ is known. Let $\hat{\delta}_n$ and $\hat{\sigma}^2_{\delta_n}$ denote the ordinary least squares estimates of the regression parameter and its associated variance. Conventional econometric practice involves evaluating whether $\hat{\delta}_n$ is statistically significantly different from 0, with a researcher concluding that nonlinearity in the growth process is present if the associated *t*-test statistic is statistically significant; under the standard 5% confidence level, this means that the magnitude of the estimated parameter and associated standard deviation must approximately obey $\frac{\left|\hat{\delta}_n\right|}{\hat{\sigma}_{\delta_n}} \geq 2$

While this is a defensible way in which to engage in the development of positive statements about nonlinearities and growth,[3] its relationship to policy advice is not obvious. To do this, it is helpful to convert the discussion from uncertainty about $\hat{\delta}_n$ to uncertainty about δ_n. This shift is useful in thinking directly about policy as it allows one to formulate a well-posed statistical decision theory problem. Let d denote the data that are available to a researcher and m denote the growth model she is employing. The use of a given growth model requires that a researcher has made an assumption on the set of growth determinants she wishes to control, i.e. the choice of elements of Z as well as an assumption on the sorts of nonlinearity that are incorporated in the model. (We will consider relaxation of these assumptions in the next section). From this perspective, a statistical analysis of nonlinearity in (2) produces a conditional probability measure

$$\mu\left(\delta_n \middle| d, m\right) \tag{3}$$

Assume that the regressors in (2) are nonstochastic and that the model errors are independent and identically normally distributed with known variance (deviations from

---

[3] Of course, there have long existed deep controversies as to the usefulness and even the interpretation of statistical significance calculations; many of the disagreements between Bayesian and frequentist approaches to statistics apply to this context. The frequentist/Bayesian distinction matters for our discussion to the extent that the Bayesian framework more naturally lends itself to decision-theoretic formulations of policy evaluation.

these are not of importance to the substance of our argument). In this case, $\mu\left(\delta_n \left| d,m\right.\right) \sim N\left(\hat{\delta}_n,\hat{\sigma}^2_{\delta_n}\right)$ so that

$$E\left(\delta_n \left| d,m\right.\right) = \hat{\delta}_n \tag{4}$$

and

$$\operatorname{var}\left(\delta_n \left| d,m\right.\right) = \hat{\sigma}^2_{\delta_n} \tag{5}$$

We make these assumptions in order to be able to translate uncertainty about the effects of policy into the standard statistical significance requirements employed in empirical work. Specifically, under these assumptions

$$\frac{\hat{\delta}_n}{\hat{\sigma}_{\delta_n}} = \frac{E\left(\delta_n \left| d,m\right.\right)}{\operatorname{var}\left(\delta_n \left| d,m\right.\right)^{1/2}} \tag{6}$$

so that *t*-statistics may be interpreted as moments of (3) and hypothesis tests using a *t*-statistic as restrictions on these moments.

How can information about the posterior density of δ_n be translated into inferences about policy? Such a translation necessarily requires that one specify the policymaker's preferences. Suppose that a policymaker possesses a loss function of the form $l\left(g_i, p_i, \vartheta_i\right)$; ϑ_i denotes those characteristics of country *i* that affect loss calculations such as initial income; there is no reason to believe that growth may be evaluated in isolation when assessing losses. This function further embodies any benefits or costs induced by the policy that are not related to its effect on growth by incorporating p_i as an argument.[4] Suppose that the policymaker has decided to employ (2) as his

---

[4] For example, the level of educational spending may produce benefits with respect to political outcomes.

model of the growth process. The evaluation of the expected loss associated with a policy is, using standard arguments,

$$E\left(l\left(g_i, p_i, \vartheta_i\right) \middle| p_i, d, m\right) = \int_G l\left(g_i, p_i, \vartheta_i\right) \mu\left(g_i \middle| p_i, d, m\right) \tag{7}$$

where G is the support of g_i. The associated optimal policy problem is based on

$$\max_{p_i \in P_i} \int_G l\left(g_i, p_i, \vartheta_i\right) \mu\left(g_i \middle| p_i, d, m\right) \tag{8}$$

where P_i denotes the support of the policy variable and m reflects the dependence of these calculations on the choice of a particular model.

The relevant point for us is that there is little connection between conditions on probability measure of the form (6) and policy evaluation calculations of the form (7) or (8). To understand the ways in which statistical significance and policy assessment differ, we consider a policy comparison of the type studied in Brock, Durlauf, and West (2003), in which a policymaker is considering whether or not to implement a one unit change in p_i. Suppose that $\delta = 0$ and that $f\left(p_i, X_i\right) = p_i \cdot 1_{X_i \in \bar{X}}$; in this expression $1_{X_i \in \bar{X}}$ is an indicator function. We are therefore considering the case where the policy is only efficacious for countries whose characteristics fall into a particular category. Threshold models, whose theoretical foundations are associated with Azariadis and Drazen (1990), often can be written in this form. Notice that the debate between Burnside and Dollar (2000) and Easterly, Levine, and Roodman (2004) may be approximated by this model, where $\bar{X}$ denotes the set of values of country-specific variables that constitute good policies and p_i denotes foreign aid.

When does the question of whether the policy variable should be changed equate to the statistical significance of $\hat{\delta}_n$? In terms of the expected loss function, an increase in the policy instrument is advisable if and only if it reduces the expected loss.

$$E\left(l\left(g_i, p_i+1, \vartheta_i\right) \middle| p_i+1, d, m\right) - E\left(l\left(g_i, p_i, \vartheta_i\right) \middle| p_i, d, m\right) < 0 \tag{9}$$

For the canonical statistical significance test, one would conclude that higher policy levels increase growth if

$$\left(-\hat{\delta}_n + \kappa(\alpha)\hat{\sigma}_{\delta_n}\right)1_{X_i \in \bar{X}} = \left(-E\left(\delta_n \middle| d, m\right) + \kappa(\alpha)\operatorname{var}\left(\delta_n \middle| d, m\right)^{1/2}\right)1_{X_i \in \bar{X}} < 0 \tag{10}$$

where $\kappa(\alpha)$ denotes the value necessary to obtain a given significance level α. It follows that statistical significance equates to the desirability of a policy change if

$$\begin{gathered} E\left(l\left(g_i, p_i+1, \vartheta_i\right) \middle| p_i+1, d, m\right) - E\left(l\left(g_i, p_i, \vartheta_i\right) \middle| p_i, d, m\right) = \\ \left(-E\left(\delta_n \middle| d, m\right) + \kappa(\alpha)\operatorname{var}\left(\delta_n \middle| d, m\right)^{1/2}\right)1_{X_i \in \bar{X}} = \\ \left(-\hat{\delta}_n + \kappa(\alpha)\hat{\sigma}_{\delta_n}\right)1_{X_i \in \bar{X}} \\ < 0 \end{gathered} \tag{11}$$

The equivalence of the desirability of a policy change with statistical significance implies that a policymaker must possess a loss function that only depends on the posterior mean and variance of δ_n, i.e. $\hat{\delta}_n$ and $\hat{\sigma}^2_{\delta_n}$. This places very strong restrictions on preferences. First, the analysis requires that the policymaker have mean variance preferences, which may not be appropriate for growth contexts where issues of the effects of negative growth on health and mortality naturally arise. Second, (11) requires the policymaker to assess losses only with respect to the component of growth associated with the policy component $f(p_i, X_i)\delta_n$; other components of the growth regression, $Z_i\gamma$ and ε_i (the latter matters for the conditional variance of the growth rate) are irrelevant. The loss function formulation implicit in (11) therefore requires that the policymaker only cares about the increment to the growth rate induced by the policy, not the growth rate per se.

We are not aware of any good argument as to why this loss function is appealing from either a normative or a positive perspective. It in fact seems easy to think of reasons why such preferences are unappealing. Suppose that a policymaker has to decide how to allocate a fix amount of investment across two countries i and j with identical initial incomes and each which obeys the same model (2), in particular the marginal effect of policy δ_n is the same across the two countries. Suppose that one country is growing at an expected rate of 1% and the other at 10% at the pre-allocation capital levels. The loss function in (11) would make the policymaker indifferent with respect to the allocation of investment across the countries.

Our analysis of the lack of policy relevance of growth empirics is an example of a very general criticism with frequentist statistical practice that is made by Bayesians, namely the absence of any firm decision-theoretic basis for statistical significance tests. The force of our argument, however, does not depend on deep issues that distinguish frequentist and Bayesian approaches. For our purposes, what matters is that the standard evaluative criterion for whether or not nonlinearity is present in the growth process does not equate to a natural statement about its implications for policy evaluation.

Does the absence of any natural equivalence between statistical significance of nonlinearities and an associated policy evaluation metric matter in practice? The answer is potentially yes at two levels. First, loss functions must matter if one is engaged in policy analysis. For asymmetric loss functions, it is clear that efforts to exploit nonlinearities may not be justified even in the presence of statistically significant evidence. Suppose the policy question is the transfer of foreign aid from a poorer country to a wealthier one. Suppose that the policymaker's loss function is the sum of the country specific loss functions. It is easy to imagine that one would decline to move the resources if the country-specific loss function is concave in the growth rate. Hence, if the policy question of interest is the allocation of finite investment resources, then nonlinearities may not affect the optimal allocation.

Second, even if one assumes mean/variance loss functions, then there potentially exist major differences between comparing coefficient distributions to growth rate distributions. Brock, Durlauf, and West (2003), for example, consider the question of whether one should recommend a reduction in tariffs to sub-Saharan African countries. That paper finds that when one accounts for the overall uncertainty of growth rates under higher versus lower tariffs, the effects on the variance of growth of the tariff change are an order of magnitude smaller than the variance associated with the tariff parameter when analyzed in isolation.

The upshot of this discussion is that one cannot translate evidence of nonlinearities, as developed via hypothesis testing, into policy recommendations, except under very special conditions. As a result, there is no natural way to draw policy inferences from the existing empirical work on growth nonlinearities. To do this, it is necessary to follow a statistical decision theoretic formulation such as the one we have described, which includes explicit attention to the question of loss functions.[5]

At a minimum, the analysis of this section may be summarized as saying that the appropriate object of interest for policymakers is $\mu\left(g_i \middle| p_i, d, m\right)$, the conditional probability measure for growth under a policy, which in turn means that objects of this type should be reported to policymakers looking for guidance. Nonlinearities are only policy-relevant to the extent their presence affects this probability. To give an example of what sorts of calculations this approach suggests, in the context of the Burnside and Dollar (2000) versus Easterly, Levine and Roodman (2004) debate on the efficacy of foreign aid, our approach would shift the focus of the argument away from evaluation of the statistical significance of $\hat{\delta}_n$ to an evaluation of questions such as how different allocations of some amount of aid alters the distribution of g_i for each of a given set of countries. In principle, the policymaker can then use his loss function to evaluate which, if any allocation, he regards as optimal.

3. Model uncertainty

A second problem with translating this wide-ranging evidence into policy prescriptions is that there does not exist a well defined nonlinear alternative to the canonical cross-country growth regression. This is true in two senses.

---

[5] We do not claim that this is the only appropriate way to formulate statistical decision problems; the obvious competitor is frequentist-based decision problems.

First, there is still no consensus as to which growth determinants need to be included when specifying a growth regression, be it linear or nonlinear. This problem is not unique to the analysis of nonlinear growth models and is in fact ubiquitous in growth studies. Durlauf, Johnson and Temple (2004) identify 145 different regressors that have appeared in published growth studies; these regressors correspond to 43 distinct growth theories. This contrasts with the 87 different growth regressors found by Durlauf and Quah (1999). It is easy to identify reasons why the failure to account appropriately for the full gamut of growth determinants can lead to spurious evidence of nonlinearity. For example, if there is a nonlinear relationship between initial income and the degree of democracy, and the degree of democracy has a causal effect on growth, then this would induce spurious evidence of a nonlinear relationship between initial conditions and growth if democracy were omitted from the regression.

While there is a growing body of work that accounts for model uncertainty in order to identify robust growth inferences,[6] this literature has generally not addressed the question of the robustness of evidence on growth nonlinearities. Exceptions include Brock and Durlauf (2001), Brock, Durlauf and West (2003) and Masanjala and Papageorgiou (2004b) which do assess the robustness of evidence on parameter heterogeneity between sub-Saharan Africa and the rest of the world; the first and third papers provide relatively strong evidence that parameter heterogeneity is present.

Second, and specific to nonlinear growth models, there has been essentially no effort to date that integrates the various findings on nonlinearity into a coherent description of a data generating process. The empirical literature on nonlinear growth models contains many different specifications of the nonlinearity. Some models are based on identifying subsets of countries which obey a common growth model, e.g Bloom, Canning, and Sevilla (2004), Canova (2004), Durlauf and Johnson (1995), and Tan (2004). These models have threshold-like structures in that economies obey one of a set of discrete models, such that the model that applies to a given country is determined by some set of initial conditions. These threshold models may be well approximated by the generic form

$$g_i = p_i\delta_i + Z_i\gamma_i + \varepsilon_i \tag{12}$$

with

$$\delta_i = \delta^k,\ \gamma_{\mathrm{i}}=\gamma^k \text{ if } X_i \in X^k; X^k \cap X^{k'} = \phi \text{ if } k \neq k' \tag{13}$$

Intuitively, countries are partitioned according to the values of some vector X into subsets, each of which obeys a common linear model. Within this literature there are a range of formulations. Durlauf and Johnson (1995) and Tan (2004) employ classification algorithms that allow for the number of distinct growth models to be endogenously determined by the data; one cost of this is that the procedure treats the mapping from initial conditions to a model as deterministic. Bloom, Canning, and Sevilla (2004), in contrast, employ a mixture distribution approach; each country is described by one of two models, and the probability that a given country is described by a particular model is determined by its initial conditions.

Other models, such as Banerjee and Duflo (2003), Durlauf, Kourtellos and Minkin (2001), Kalaitzidakis, Mamuneas, and Stengos (2000) and Liu and Stengos (1999) are interpretable as varying coefficient models[7]; models of this type take the form

$$g_i = p_i\delta(X_i) + Z_i\gamma(X_i) + \varepsilon_i \tag{14}$$

where $\delta(\cdot)$ and $\gamma(\cdot)$ are smooth functions. Within this class, there are differences as to which coefficients vary and what determines elements of the variation. For example one finds some studies that allow for a mix of linear and nonlinear coefficients (e.g. Liu and Stengos (1999)) whereas others (e.g. Durlauf, Kourtellos and Minkin (2001)) do not.

---

[6] Early contributions include Levine and Renelt (1992) and Sala-i-Martin (1997); recent contributions which use model averaging methods of the type we describe below.

[7] See Hastie and Tibshirani (1993) for a discussion of such models from a statistical perspective.

This type of framework encompasses the introduction of nonlinearities via ad hoc interaction terms. To see this, suppose that one accounts for the interaction of policy with some country-specific characteristic x_i by adding $x_i p_i$ as a regressor to (1). This is equivalent to the specification $\delta\left(X_i\right)=\delta_0+x_i\delta_1$.

Finally, some work on nonlinearity has focused on issues of dimension reduction, i.e. finding low dimensional nonlinear models to capture growth dynamics. Desdoigts (1999) and Kourtellos (2003a,b) pursue this approach. These types of models amount to estimating growth models of the form

$$g_i=\sum_k f_k\left(p_i\delta^k+Z_i\gamma^k\right)+\varepsilon_i \tag{15}$$

where each $f_k\left(\cdot\right)$ denotes a distinct nonlinear function. This type of approach is designed to look for low dimensional approximations for high dimensional nonlinear functions.[8]

One can develop a more elaborate typology of nonlinear statistical growth models, but the three types of models so far described are sufficient to illustrate the general problem. While there are now studies that exist that justify the statistical models, there does not exist a consensus on how to model nonlinearity. This lack of consensus applies both across and within the nonlinear model classes we have described.

How should one account for model uncertainty in policy evaluation? From the perspective of the earlier decision problem, eq. (7), the natural way to proceed is to explicitly calculate an analog to (7) that accounts for the fact that the researcher does not know the true data generating process for growth.[9] In other words, the evaluation of the effects of a policy should be based on

---

[8] In the projection pursuit formulation (15), each of the one-dimensional $p\delta^k+Z\gamma^k$ terms is known as a projection. Hence projection pursuit approximates a possibly complex nonlinear process with the sum of a set of simple one dimensional nonlinear functions $f_k\left(\cdot\right)$ whose arguments are the projections.

[9] Other approaches, such as Levine and Renelt (1992) may be interpreted as imposing very special preferences, i.e. minimax preferences with respect to model uncertainty (see

$$E\left(l\left(g_i, p_i, \vartheta_i\right) \middle| p_i, d\right) = \int_G l\left(g_i, p_i, \vartheta_i\right) \mu\left(g_i \middle| p_i, d\right) \tag{16}$$

rather than eq. (7). While (7) computes the expected loss conditional on a policy, the data and a given model, $\mu\left(g_i \middle| p_i, d, m\right)$, (16) computes the expected loss conditional only on a policy and the data, $\mu\left(g_i \middle| p_i, d\right)$. The "model-free" analog to (8) is constructed in the same way:

$$\max_{p_i \in P_i} \int_G l\left(g_i, p_i, \vartheta_i\right) \mu\left(g_i \middle| p_i, d\right) \tag{17}$$

The construction of $\mu\left(g_i \middle| p_i, d\right)$ requires the use of a set of techniques that are known in the statistical literature as model averaging methods. The ideas behind model averaging originate in Leamer (1978) but have recently reemerged via Draper (1995) and Raftery, Madigan and Hoeting (1997); useful introductions are Wasserman (1996) and especially Hoeting, Clyde, Madigan and Raftery (1999). In the growth context, model averaging has been advocated and employed in Brock and Durlauf (2001a), Brock, Durlauf and West (2003), Fernandez, Ley and Steel (2001), Doppelhofer, Miller and Sala-i-Martin (2004) and Masanjala and Papageorgiou (2004b).

The basic idea of model averaging is that uncertainty about the true model for an outcome of interest should be treated symmetrically to other forms of uncertainty. In order to account for this uncertainty, one constructs probability statements that do not condition on a given model but rather condition on a model space M. The elements of M represent different candidates for the "true" growth process. In this discussion, we assume that the true model is an element of this space; the interpretation of model averaging when none of the models is true is an area of current research. Once the model space has been formulated, one can describe the conditional (given data and policy) probability measure for a particular growth rate as

---

Brock, Durlauf, and West (2003); others such as Sala-i-Martin (1997) do not appear to

$$\mu\left(g_i \middle| p_i, d\right) = \sum_{m \in M} \mu\left(g_i \middle| p_i, d, m\right) \mu\left(m \middle| d\right) \tag{18}$$

In this expression, $\mu\left(m \middle| d\right)$ is the posterior model probability associated with model *m*. By Bayes' rule,

$$\mu\left(m \middle| d\right) \propto \mu\left(d \middle| m\right) \mu\left(m\right) \tag{19}$$

Eq. (19) illustrates how these posterior model probabilities are determined by $\mu\left(d \middle| m\right)$, the likelihood of the data given the model and $\mu\left(m\right)$, the prior probability assigned to the model. The assignment of such priors is often problematic since prior beliefs may not lend themselves to ready quantification. As a result, it is common to use simple priors which assign all models equal prior probability (Brock and Durlauf (2001a), Fernandez, Ley, and Steel (2001)) or assign higher prior probability to simpler models (Doppelhofer, Miller, and Sala-i-Martin (2004)). Brock, Durlauf, and West (2003) propose assigning priors based on accounting for the different levels of uncertainty that exist in constructing growth models; in doing this they distinguish between uncertainty about underlying growth theories, uncertainty about specification, uncertainty about measurement and uncertainty about parameter heterogeneity. We are not aware of any evidence that the choice of prior has mattered in the applications of model averaging to economic growth.

How will model averaging affect policy evaluation? This may be seen most easily when one considers the first two moments of the growth process; recall that these moments are sufficient for quadratic preferences. Following Leamer (1978), once model uncertainty has been incorporated, the expected value and variance of the growth rate for a country equal

$$E\left(g_i \middle| p_i, d\right) = \sum_m \mu\left(m \middle| d\right) E\left(g_i \middle| p_i, d, m\right) \tag{20}$$

---

have any decision-theoretic justification.

and

$$\begin{gathered} var\left(g_i \middle| p_i, d\right) = E\left(g_i^2 \middle| p_i, d\right) - \left(E\left(g_i \middle| p_i, d\right)\right)^2 = \\ \sum_{m \in M} \mu\left(m \middle| d\right) var\left(g_i \middle| p_i, d, m\right) + \sum_{m \in M} \mu\left(m \middle| d\right)\left(E\left(g_i \middle| p_i, d, m\right) - E\left(g_i \middle| p_i, d\right)\right)^2 \end{gathered} \tag{21}$$

From the perspective of policy evaluation, the term

$$\sum_{m \in M} \mu\left(m \middle| d\right)\left(E\left(g_i \middle| p_i, d, m\right) - E\left(g_i \middle| p_i, d\right)\right)^2 \tag{22}$$

is of particular interest. This term captures how different model-specific estimates of the expected mean effect of the policy $E\left(g_i \middle| p_i, d, m\right)$ vary around the expected mean effect when model uncertainty has been accounted for, $E\left(g_i \middle| p_i, d\right)$. This term thus captures a component of the uncertainty of outcomes associated with a policy that is ignored whenever a single model is assumed to be the correct one. Draper (1995) argues that one reason why forecasts of the effects of changes in policies or other variables are much less accurate than would be expected given the stochastic structure of the models under use is that this term is conventionally ignored.

When one compares the evaluation of policy via eq. (16) to a standard statistical significance approach, it is clear how far the body of evidence on growth nonlinearities is from allowing one to compare policy counterfactuals. That being said, objects such as (16) can in principle be constructed as is being done in macroeconomic contexts such as monetary policy rule comparison (see Levin and Williams (2003) and Brock, Durlauf and West (2003,2005)). Similar exercises are possible for the analysis of growth policies when nonlinearities may be present. Again, referring to the debate on foreign aid and growth, we would advocate reporting results that avoid conditioning on a particular type of nonlinearity, i.e. a nonlinearity in which growth is affected by the product of aid and policy quality. Clearly there are many nonlinear specifications that capture the idea that the effectiveness of aid depends on policy regime; commitment to one type of nonlinearity is not justified by theoretical reasons.

Model averaging methods are still in their infancy in economics and statistics. There still exist many outstanding questions concerning the formulation of priors on model spaces, the use of Bayesian versus frequentist estimates, the appropriateness of various approximations of posterior probabilities in order to ensure computational feasibility, etc. Further, the selection of a model space *M* raises judgment issues that are difficult to resolve. Nevertheless, this approach seems very promising and should be applied to nonlinear growth analysis.

While model averaging methods seem the most appropriate way of integrating model uncertainty into growth analyses, we conclude this section by noting that for some purposes, it may be useful for a researcher to report, in addition to "model-free" results, specifically, $\mu\left(g_i \middle| p_i, d\right)$, some of model-specific probabilities $\mu\left(g_i \middle| p_i, d, m\right)$ from which is it is constructed as well as the associated posterior model probabilities $\mu\left(m \middle| d\right)$. It may be the case that a policymaker wishes to know how a policy performs for the least favorable economic model among those under consideration.

One reason for this is that a policymaker may possess preferences in which uncertainty about the true model is treated differently than uncertainty within a model; Epstein and Wang (1994) analyze preferences of this type, which lead to a policymaker placing evaluating policies according to

$$(1-e)\int_G l\left(g_i, p_i, \vartheta_i\right)\mu\left(g_i \middle| p_i, d\right) + e\left(\sup_{m\in M}\int_G l\left(g_i, p_i, \vartheta_i\right)\mu\left(g_i \middle| p_i, d, m\right)\right) \quad (23)$$

By this criterion, a policymaker evaluates each policy via a weighted average of the expected loss and the loss under the least favorable model. The parameter *e* is interpreted as a measure of ambiguity aversion in that $e = 0$ reproduces the standard expected loss calculation whereas $e = 1$ means that a policymaker assumes the least favorable model of the economy when evaluating a policy, so that the posterior model probabilities are no long relevant to policy analysis. Hansen and Sargent (2004) provide a deep examination of the implications of ambiguity aversion. We will again refer to ambiguity. Here we wish simply to note that model averaging calculations may supplement rather than replace model-specific calculations. Brock, Durlauf, and West (2005) discuss ways to report model-specific calculations in environments with model uncertainty.

4. Observational equivalence

A final problem in mapping evidence of nonlinearity into policy recommendations is the absence of a set of results on observational equivalence for nonlinear growth models. We employ the phrase observational equivalence rather than identification as the analysis of identification typically revolves around determining conditions on exogenous variables and errors for a null model such that no competing model will provide the same probability statements about observables as the null model. Such exercises are important, as they tell us what can "ultimately" be learned from observables. However, for the analysis of policy, what matters is determining the space of models that are compatible with the available data. Put differently, identification asks under what conditions one model can generate different probability statements from all competitors. For policy purposes, the question is, given the data available, what range of models cannot be distinguished.

The idea of observational equivalence is illustrated in Figure 1, which generalizes a figure in Durlauf and Johnson (1995). In this figure, four different capital transition functions $\psi(\cdot)$ with threshold externalities à la Azariadis and Drazen (1990) are graphed with associated data points marked. Observations, represented by ovals, are the same in each of the cases; each of the candidate transition functions, by construction matches the data perfectly. Hence, all four models are observationally equivalent. One can in principle identify one model versus the other in a cross-section if the cross section density of k fulfills certain conditions. However, for the evaluation of policy given the limited data available in this example, this is not what is relevant.

These observationally equivalent structures presumably have very different policy implications. We initially focus on cases 1-*A* and 1-*B*. The difference between the two transition functions is that in case 1-*A*, there exist 3 locally stable steady states, whereas in case 1-*B*, there exist two locally stable steady states. Suppose we are allocating capital across countries and are considering the benefits of allocating capital to country *i* versus *j*. In case 1-*A* , country *j* is trapped at an intermediate steady state whereas in case 1-*B* country *j* will convergence to the high steady state. The benefits of the allocation of resources shift will depend critically on whether the capital transfer will change the steady state for *j*. (One obvious reason is that we are assuming that country *i* has lower initial income than *j*, so egalitarian considerations would argue against the transfer; such considerations being in principle possibly trumped by the magnitude of the benefit to *j*). But this is precisely what cannot be determined in the cross-section given the available data. More generally, the benefits of a policy to raise capital in *j* will sensitively depend on where the country is located relative to any thresholds that produce multiple steady states, which may not be distinguishable given available cross-country data.

The issue of observational equivalence cannot be disassociated from the questions that arise in specifying a model space. We now consider cases 1-*C* and 1-*D* which are also observationally equivalent to 1-*A* and 1-*B*. In case 1-*C* there is a capital interval that is totally unproductive. In case 1-*D*, there is a capital interval with extremely high marginal product. A policymaker would presumably want to move capital across countries in order to avoid the trap or in order to exploit the region of high marginal product. However, neither model makes sense in terms of any underlying economic reasoning, so one would not wish to place weight on them when designing policies.

From the perspective of future work on nonlinear growth, we believe the development of a typology of observationally equivalent nonlinear models is arguably the most important outstanding research question.

5. What is to be done?

In evaluating the implications of our three criticisms, it is important to note that they fall into two distinct categories. The criticisms we have made of the existing nonlinear growth literature based on the absence in current research of a statistical decision theoretic approach and the failure to appropriately account presence of model uncertainty may each, in principle, be addressed using existing statistical methods. This may require the use of Bayesian methods or Bayesian/frequentist hybrid methods[10] which are currently available. On the other hand, problems of observational equivalence are by definition not amenable to "solution" by employing different statistical methods.[11] Rather, problems of observational equivalence require subsequent research in two directions. The first is the construction and evaluation of new sources of evidence that may be used to break observational equivalence. The second is the evaluation of how a policymaker should engage in decisionmaking when observational equivalence is present. Based on these directions, we suggest some specific possibilities that seem important in developing policies that account for nonlinearities.

i. historical studies

The observational equivalence problems we have discussed make clear why historical approaches to understanding growth are so important: historical approaches, by careful attention to transitions, can provide exactly the sort of evidence to restrict the class of potential nonlinearities that statistical studies cannot. The work of Acemoglu, Johnson, and Robinson (2001,2002) and Engerman and Sokoloff (2003,2005) are exemplary in terms of their ability to supplement the standard growth regression framework. One need not claim that a question such as "do institutions affect growth?" has been resolved by such papers to realize that they have made much deeper progress on this question than has come from most regression approaches Such studies, of course,

---

[10] Brock, Durlauf, and West (2003,2005) and Doppelhofer, Miller and Sala-i-Martin (2004) employ frameworks in which frequentist estimates of parameters are averaged over models using posterior model probabilities. While, following Raftery (1995), the approach may be interpreted as an approximation of a fully Bayesian procedure, the philosophy of the approach is to employ a Bayesian solution to model uncertainty in a frequentist estimation framework.

[11] A strict Bayesian might object that observational equivalence simply means the posteriors of certain models are equal to their prior probabilities and so may be handled using Bayesian methods; our point is that statistical methods do not address the observational equivalence problem in a nontrivial way.

correspond to the views of Karl Popper with which this paper began. This does not mean that there is no role for regression work for those questions that may be addressed with more historical approaches. Tan (2004), for example, using state-of-the-art nonlinear statistical methods, shows how debates between geography and institutions are to some extent artificial, as their respective importance depends on initial conditions.

ii. taking dynamics seriously

A second promising course is the development of explicit efforts to understand transitions. While panel data methods have become increasingly popular in empirical growth studies, their main uses have been to eliminate unobserved fixed effects or country specific differences in coefficients (eg. Islam (1998) and Lee, Pesaran and Smith (1997)) or to allow for the construction of instrumental variables using lagged variables (eg. Caselli, Esquivel and Lefort (1996)). As such, the typical panel growth analyses do not account for how transition dynamics will imply coefficient evolution (for nonlinear growth models) nor how these dynamics can be informative about phenomena such as multiple steady states.

In this respect, recent work by Phillips and Sul (2003) is quite important. This analysis uses time series methods to study growth dynamics but attempts to allow for parameter evaluation of the type that is naturally suggested by transitional growth dynamics of the type found in the standard neoclassical growth model. By this we mean that the Phillips and Sul approach can capture the differences in growth model coefficients associated with different stages of the transition towards steady state behavior. An important extension of their analysis would be the investigation of models whose transition dynamics can accommodate interesting nonlinear growth frameworks. Bernard and Durlauf (1996) indicate some of the ways in which endogenous growth alternatives affect the appropriate ways of constructing cross-country growth regressions whose purpose is to allow discrimination between neoclassical and endogenous growth perspectives.

iii. microeconomic studies

Finally, it seems likely that policy relevant work on nonlinearities will require a shift away from aggregated to microeconomic level data. Many of the generative mechanisms that produce aggregate nonlinearities seem to capture various types of social interactions such as information flows across social networks, social norms as substitutes for incomplete contracts, etc. If so, then the evidence of these phenomena at a micro-level will clearly matter if one wishes to exploit these nonlinearities via policy.

There exist deep identification problems in environments with social interactions; see Brock and Durlauf (2001b,2004) and Manski (1993,2000) for formal analyses. However, what matters for growth economics is that one can establish empirically plausible conditions under which social interactions may be identified. Further experimental approaches to empirical development economics, of which Banerjee et al (2004) and Kremer, Miguel and Thornton (2004) are nice examples, provide an important complement to the use of observational data to identify socioeconomic mechanisms. One example of how such evidence matters is Banerjee and Duflo (2004) who address the empirical significance of credit market constraints, which represent one of the generative mechanisms for poverty traps, see Azariadis and Stachurski (2004) for an overview.

We suspect that unless aggregate evidence of nonlinearities is supplemented with evidence of specific micro-based mechanisms, the aggregate evidence will have limited policy relevance. The reason for this is that the aggregate evidence of nonlinearities is compatible with a range of microfoundations, whose details matter for how a policymaker should react to nonlinearities. To see why this is so, consider Durlauf and Johnson (1995) who find evidence of convergence clubs based on initial income and initial literacy. The evidence they present may imply a causal role for these variables in determining the aggregate production function or it may be the case that these initial conditions proxy for other variables.

This distinction is of obvious importance for policy evaluation. If the reason for the dependence on initial literacy is a discontinuity in the aggregate production function that represents a real nonconvexity (i.e. once workers achieve a certain level of human capital, their productivity jumps up) that will have implications for how an international aid agency might allocate resources for education across countries or for how a country should make internal human capital investment decisions. In contrast, if the initial conditions in Durlauf and Johnson proxy for a variable such as culture which, although hard to measure, nevertheless represents the main causal mechanism by which different aggregate production functions emerge, then the implications for the allocation of educational resources are very different.

This example may lead policy analysis into the world of decisionmaking under ambiguity rather than decisionmaking under uncertainty. By this, we mean that it may no longer be possible to engage in expected loss calculations since it will not be possible to assign probabilities to different outcomes such as the effect of different levels of literacy on social norms. The recognition that policymakers face forms of uncertainty for which (meaningful) probabilities do not exist has motivated a new and growing literature in macroeconomics on robust policy analysis whose implications we discuss next.

iv. robust policy analysis

We see important potential in the application of recent work in macroeconomics on robustness to the analysis of growth policies. The robustness literature initiated by Hansen and Sargent (2004),[12] focuses on the analysis of environments in which model uncertainty is not resolved by computing posterior probabilities. The idea in this work is to consider model spaces in which the implied uncertainty is local in the sense that all potential models are close to each other, according to some metric. In such contexts, one evaluates policies using a minimax criterion. As such, work on robustness represents an approach to decisionmaking in the presence of ambiguity aversion; under minimax, one assumes the least favorable model when assessing a policy. This approach, because it avoids posterior model probability calculations, is relevant to evaluating policies under

---

[12] See Giannoni (2002), Hansen and Sargent (2001), Brock and Durlauf (2005), Onatski and Stock (2002), Onatski and Williams (200) Tetlow and von zur Muehlen (2002) for a range of analyses.

observational equivalence. When different models are observationally equivalent, then differences in their posterior probabilities will be proportional to differences in their prior probabilities and prior probabilities in growth contexts are likely to be arbitrary. The robustness approach does not depend on assigning priors; rather it shifts the focus towards the constructions of policies that work well regardless of the true model.

For the analysis of nonlinearities and growth, the main idea in the robustness literature that we see as important is the shift from the identification of optimal policies toward the identification of policies that perform relatively well across different possible models of the economy,[13] specifically with respect to the presence of possible nonlinearities. Robustness analysis allows one to ask questions such as how should one adjust policies in the presence of possible nonlinearities, which seem natural for the growth case.

6. Conclusions

This paper is designed to outline some of the complications that exist when one moves from positive analyses of nonlinearities in the growth process to policy recommendations. We do not believe that one can say much about policy given the available aggregate evidence of growth nonlinearities. To rectify this, we advocate three research directions. First, we see a need for a more precise orientation of nonlinear growth research toward policy questions in terms of estimating the relevant objects of study, i.e. conditional distributions of growth under different scenarios. Second, we see a need to address model uncertainty in developing robust evidence of nonlinearities. Third, we believe that nonlinear growth analysis must move to develop a typology of observationally equivalent structures for the limited data that are available as well as

---

[13] This is not an endorsement of specific evaluative criteria such as the use of minimax which underlies the robustness literature. For such cases, we see minimax as an appealing way of moving from the consideration of optimal to “good” policies. For non-local model uncertainty, minimax seems less appealing due to the potential for highly improbable models to control decisions. Work by Manski (2004) on minimax regret approaches to policy evaluation may prove to be important in this regard.

more extensive consideration of what evidence may be developed to evaluate nonlinearities.

The statistical approaches we advocate in no way are panaceas for deep difficulties that exist in developing policy implications from the relatively crude data that are available for assessing cross-country experiences. These methods do hold the promise of making policy recommendations more robust, possibly at the cost of greater modesty in predictions. Work of this type will produce less glamorous claims than those one often sees in single-cause or single-model growth exercises, but will provide a far firmer basis for serious policy debates.

Figure 1: Observational Equivalence in the Azariadis-Drazen Growth Model

Figure 1A: Three Steady States

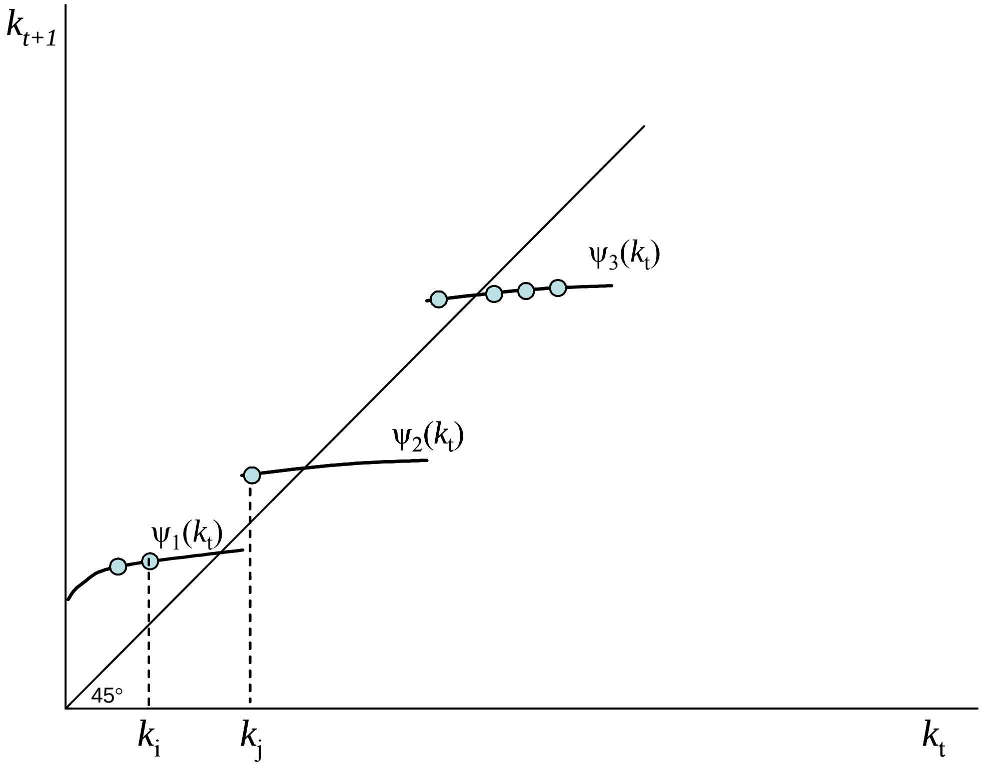


Figure 1B: Two Steady States

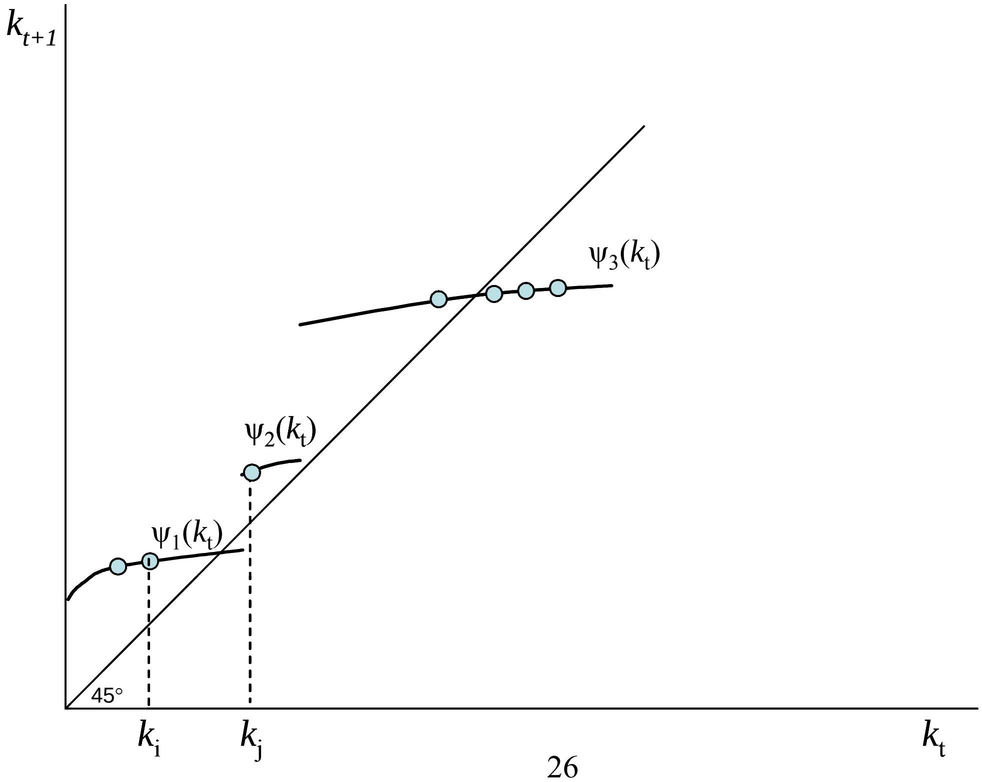


Figure 1C: Three Steady States

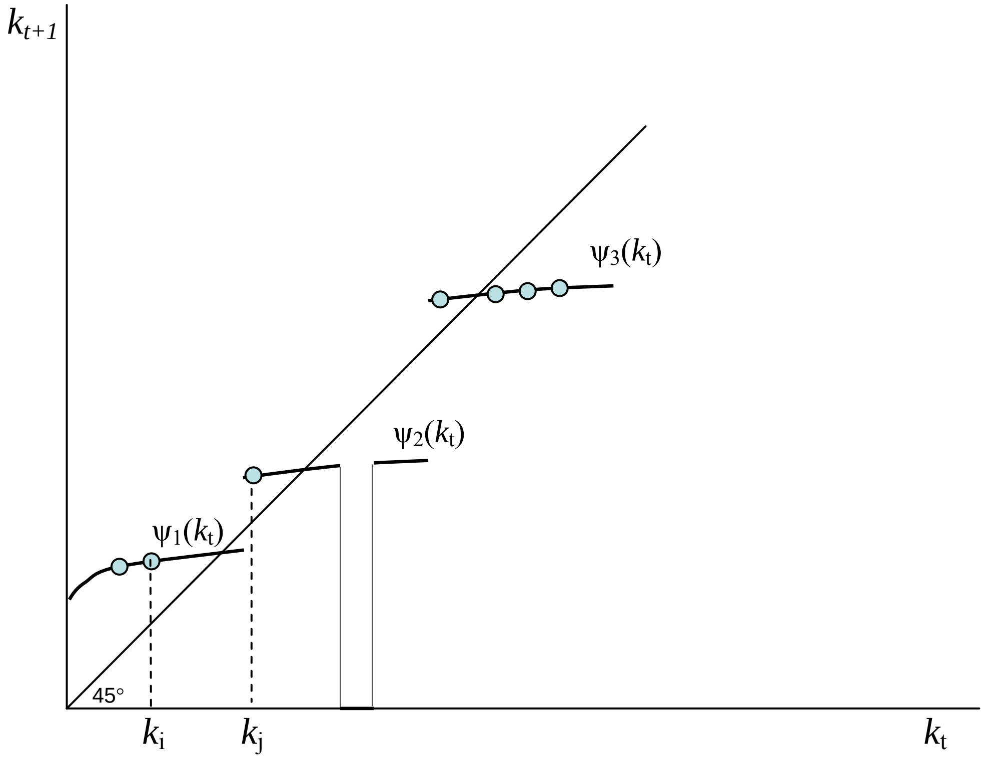
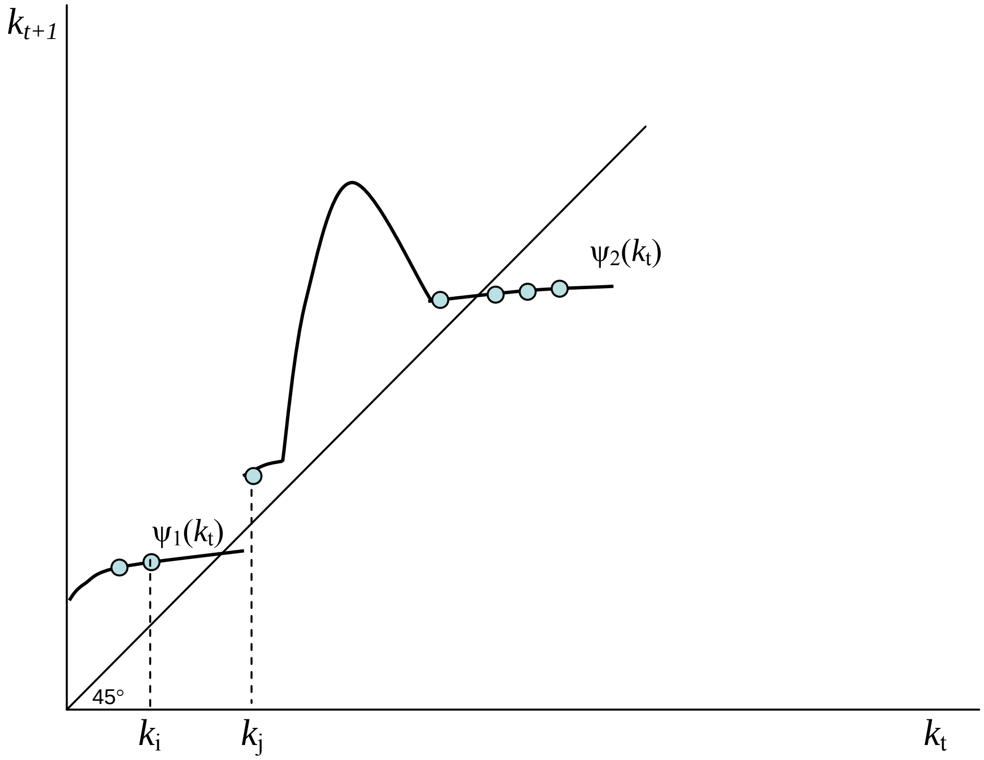


Figure 1D: Two Steady States



References

Acemoglu, D., S. Johnson and J. Robinson, (2001), "The Colonial Origins of Comparative Development: An Empirical Investigation," *American Economic Review*, 91, 5, 1369-1401.

Acemoglu, D., S. Johnson and J. Robinson (2002), "Reversal of Fortune: Geography and Institutions in the Making of the Modern World Income Distribution," *Quarterly Journal of Economics*, 117, 4, 1231-1294.

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

Azariadis, C. and J. Stachurski (2004), "Poverty Traps," in *Handbook of Economic Growth*, P. Aghion and S. Durlauf, eds., Amsterdam: North-Holland.

Barro, R., (1991), "Economic Growth in a Cross-Section of Countries," *Quarterly Journal of Economics*, 106, 407-443.

Banerjee, A. and E. Duflo, (2003), "Inequality and Growth: What Can the Data Say?," *Journal of Economic Growth*, 8, 3, 267-300.

Banerjee, A. and E. Duflo, (2004), "Do Firms Want to Borrow More? Testing Credit Constraints Using a Directed Lending Program," mimeo, MIT.

Banerjee, A., E. Duflo, S. Cole, and L. Linden, (2004), "Remedying Education: Evidence from Two Randomized Experiments in India," mimeo, MIT.

Berger, J., (1980), *Statistical Decision Theory*, New York: Springer-Verlag.

Bernard, A. and S. Durlauf, (1996), "Interpreting Tests of the Convergence Hypothesis", *Journal of Econometrics*, 71, 1-2, 161-73.

Bloom, D., D. Canning and J. Sevilla (2003), "Geography and Poverty Traps", *Journal of Economic Growth*, 8, 355-378.

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

Brock, W. and S. Durlauf (2001b), "Interactions-Based Models", in: *Handbook of Econometrics 5*, J. Heckman and E. Leamer eds., Amsterdam: North-Holland, Amsterdam.

Brock, W. and S. Durlauf, (2004), "Identification of Binary Choice Models with Social Interactions," mimeo, University of Wisconsin.

Brock, W. and S. Durlauf, (2005), “Local Robustness Analysis: Theory and Application,” *Journal of Economic Dynamics and Control*, forthcoming.

Brock, W., S. Durlauf and K. West, (2003), “Policy Evaluation in Uncertain Economic Environments (with discussion),” *Brookings Papers on Economic Activity*, 1, :235-322.

Brock, W., S. Durlauf and K. West, (2005), “Model Uncertainty and Policy Evaluation: Some Theory and Empirics,” mimeo, University of Wisconsin.

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

Caselli, F., G. Esquivel and F. Lefort, (1996), “Reopening the Convergence Debate: A New Look at Cross-country Growth Empirics,” J*ournal of Economic Growth* 1, 3, 363-89.

Chamberlain, G., (2001), “Econometrics and Decision Theory,” *Journal of Econometrics*, 95, 255-83.

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

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

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

Draper, D., (1995), “Assessment and Propagation of Model Uncertainty,” *Journal of the Royal Statistical Society, series B*, 57, 45-70.

Duffy, J. and C. Papageorgiou, (2000), “A Cross-Country Empirical Investigation of the Aggregate Production Function Specification,” J*ournal of Economic Growth*, 5, 87-120.

Durlauf, S. and P. Johnson, (1995), “Multiple Regimes and Cross-Country Growth Behavior,” J*ournal of Applied Econometrics,* 10, 4, 363-384.

Durlauf, S., P. Johnson, and J. Temple, (2004), “Growth Econometrics,” in *Handbook of Economic Growth*, P. Aghion and S. Durlauf eds., Amsterdam: North Holland, forthcoming.

Durlauf, S., A. Kourtellos, and A. Minkin, (2001), “The Local Solow Growth Model,” *European Economic Review*, 45, 928-940, Durlauf, S. and D. Quah, (1999), "The New Empirics of Economic Growth," in *Handbook of Macroeconomics*, J. Taylor and M. Woodford, eds., Amsterdam: North Holland.

Easterly, W., R. Levine and D. Roodman, (2004), "Aid, Policies, and Growth: Comment," *American Economic Review*, 94, 3, 774-780.

Epstein, L. and T. Wang, (1994), "Intertemporal Asset Pricing Behavior Under Knightian Uncertainty," *Econometrica*, 62, 283-322.

Engerman, S. and K. Sokoloff, (2003), "Institutional and Noninstitutional Explanations of Economic Differences," *NBER Working Paper no. 9989*.

Engerman, S. and K. Sokoloff, (2005), "Colonialism, Inequality and Long-Term Paths of Development," *NBER Working Paper no. 11057*.

Fernandez, C., E. Ley, and M. Steel, (2001), "Model Uncertainty in Cross-Country Growth Regressions," *Journal of Applied Econometrics*, 16, 563-576.

Fiaschi, D. and M. Lavezzi, (2003), "Distribution Dynamics and Nonlinear Growth," *Journal of Economic Growth*, 8, 4, 355-378.

Giannoni, M., (2002), "Does Model Uncertainty Justify Caution? Robust Optimal Monetary Policy in a Forward-Looking Model," *Macroeconomic Dynamics*, 6, 111-144.

Hansen, L. and T. Sargent, (2001), "Acknowledging Misspecification in Macroeconomic Theory," *Review of Economic Dynamics*, 4, 519-35.

Hansen, L. and T. Sargent, (2004) *Robust Control and Economic Model Uncertainty*, Book manuscript, Hoover Institution, Stanford University.

Hastie, T. and R. Tibshirani, (1993), "Varying Coefficient Models (with discussion), *Journal of the Royal Statistical Society*, *series B*, 55, 757-796.

Hoeting, J., M. Clyde, D. Madigan and A. Raftery, (1999) "Bayesian Model Averaging: A Tutorial," *Statistical Science*, 14, 382-401.

Howitt, P., and D. Mayer-Foulkes, (2002), "R&D, Implementation and Stagnation: A Schumpeterian Theory of Convergence Clubs," *National Bureau of Economic Research Working Paper no. 9104*.

Galor, O., (1997), "Convergence? Inferences from Theoretical Models," *Economic Journal*, 106, 1056-1069.

Islam, N., (1995), "Growth Empirics: A Panel Data Approach," *Quarterly Journal of Economics*, 110, 4,1127-1170.

Johnson, P. and L. Takeyama, (2001), "Initial Conditions and Economic Growth in the US States," *European Economic Review*, 45, 4-6, 919-27.

Kalaitzidakis, P., T. Mamuneas, and T. Stengos, (2000), "A Non-linear Sensitivity Analysis of Cross Country Growth Regressions," *Canadian Journal of Economics*, 33, 3, 604-17.

Kourtellos, A., (2003a), "Modeling Parameter Heterogeneity in Cross-Country Growth Regression Models," mimeo, University of Cyprus.

Kourtellos, A., (2003b), "A Projection Pursuit Approach to Cross-Country Growth Data," mimeo, University of Cyprus.

Kremer, M., E. Miguel, R. Thornton, (2004), "Remedying Education: Evidence from Two Randomized Experiments in India," *National Bureau of Economic Research Working Paper no.*10971.

Leamer, E., (1978), *Specification Searches*. New York: John Wiley and Sons.

Lee, K., M. Pesaran and R. Smith, (1997), "Growth and Convergence in Multi Country Empirical Stochastic Solow Model," *Journal of Applied Econometrics*, 12, 4, 357-92.

Levin, A. and J. Williams, (2003) "Robust Monetary Policy with Competing Reference Models," *Journal of Monetary Economics*, 50, 945-975.

Levine, R. and D. Renelt, (1992), "A Sensitivity Analysis of Cross-Country Growth Regressions," *American Economic Review*, 82, 942-963.

Liu, Z. and T. Stengos, (1999), "Non-Linearities in Cross-country Growth Regressions: A Semiparametric Approach," *Journal of Applied Econometrics*, 14, 5, 527-38.

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

Manski, C., (1993), "Identification of Endogenous Social Effects: The Reflection Problem." *Review of Economic Studies*, 60, 3, 531-542.

Manski, C., (2000), "Economic Analysis of Social Interactions," *Jorunal of Economic Perspectives*, 14, 3, 115- 136.

Manski, C., (2004), "Statistical Treatment Rules for Heterogeneous Populations," *Econometrica*, 72, 4, 1221-1246.

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

Masanjala, W. and C. Papageorgiou, (2004b), "Rough and Lonely Road to Prosperity: A Reexamination of the Sources of Growth in Africa Using Bayesian Model Averaging," mimeo, Louisiana State University.

Onatski, A. and J. Stock, (2002) "Robust Monetary Policy Under Model Uncertainty in a Small Model of the U.S. Economy," *Macroeconomic Dynamics*, 6, 85-110.

Onatski, A. and N. Williams, (2003), "Modeling Model Uncertainty," *Journal of the European Economic Association*, 1, 1078-1122.

Papageorgiou, C., (2002), "Trade as a Threshold Variable for Multiple Regimes," *Economics Letters*, 71, 1, 85-91.

Papageorgiou, C. and W. Masanjala, (2004), "The Solow Model with CES Technology: Nonlinearities with Parameter Heterogeneity", *Journal of Applied Econometrics*, 19, 2, 171-201.

Phillips, P. and D. Sul (2003), "The Elusive Empirical Shadow of Growth Convergence," *Cowles Foundation Discussion Paper no. 1398.*

Popper, K. (1957), *The Poverty of Historicism*, London: Routledge, 1994 reprint.

Raftery, A., (1995) "Bayesian Model Selection in Social Research (with discussion)," in *Sociological Methodology 1995,* P. Marsden ed., Cambridge, MA: Blackwell

Raftery, A., D. Madigan, and J. Hoeting, (1997), "Bayesian Model Averaging for Linear Regression Models," *Journal of the American Statistical Association* 92, 437, 179-91.

Sala-i-Martin, X., (1997), "I Just Ran 2 Million Regressions," *American Economic Review*, 87, 178-83.

Sims, C., (2002), "The Role of Models and Probabilities in the Monetary Policy Process," *Brookings Papers on Economic Activity,* 2, 1-40.

Tan, C. M., (2004), "No One True Path to Development: Uncovering the Interplay Between Geography, Institutions, and Ethnic Fractionalization in Economic Development," mimeo, Tufts University.

Tetlow, R. and von zur Muehlen, P., (2001) Robust Monetary Policy with Misspecified Models: Does Model uncertainty Always Call for Attenuated Policy?," *Journal of Economic Dynamics and Control*, 25, 6-7, 911-949.

Wasserman, L. (2000), “Bayesian Model Selection and Bayesian Model Averaging,” *Journal of Mathematical Psychology*, 44, 97-112.