



# Cleaning up the kitchen sink: Specification tests and average derivative estimators for growth econometrics



Francisco Rodríguez <sup>a</sup>, Cameron A. Shelton <sup>b,\*</sup>

<sup>a</sup> Emerging Markets Research Group, Bank of America, USA

<sup>b</sup> Robert Day School of Economics and Finance, Claremont McKenna College, USA

## ARTICLE INFO

### Article history:

Received 22 April 2013

Accepted 22 July 2013

Available online 17 August 2013

### JEL classification:

O40

O47

### Keywords:

Economic growth

Cross-country growth regressions

Non-linearities

Non-parametric econometrics

## ABSTRACT

Theory and case-study evidence suggest that non-linearities are pervasive in the growth process. Growth empirics have attempted to characterize these non-linearities with regression trees, additively separable non-parametric estimates, or simple interaction terms. Each method requires specific assumptions about functional form which we demonstrate may not be defensible. We provide two alternate mechanisms for making inference about the growth effects of production–function shifters that do not make *a priori* assumptions about functional form: monotonicity tests and average derivative estimation. Our results suggest that the growth effects of policies are country-specific while the effects of institutions are more robustly monotonic.

© 2013 Elsevier Inc. All rights reserved.

## 1. Introduction

The process of economic growth is riddled with non-linearities and interaction effects. Modern growth empirics are dominated by attempts to map these pervasive non-linearities (see Section 2). Unfortunately, the paucity of data and the large number of plausible growth factors prevent exploration of a fully non-parametric model. As a result, all studies of growth must make an *a priori* choice of how to restrict the specification so as to focus on a particular set of non-linearities. Unfortunately, we show that the data reject common *a priori* restrictions.

Both to learn about the proper functional form and to provide information useful to policy reform, we suggest two complementary approaches. First, we employ an average derivative estimator (ADE) to consistently estimate the sample average marginal effect of growth factors. In a non-linear world, this ADE estimate does not provide the local marginal effect relevant to any particular country. However, given that such a local partial effect is highly sensitive to the presumed specification and thus cannot be estimated with confidence, we believe the sample average marginal effect represents useful information.

Second, we test each growth factor for monotonicity. We do so by imposing the restriction on a general function form and asking whether that restriction can be rejected by the data. Failure to reject monotonicity adds useful information for policy-makers by suggesting that reform in a certain direction will at least have positive effects on growth, albeit of unknown magnitude. Rejection of monotonicity suggests greater caution. Combining monotonicity tests with the ADE gives us a better picture of the marginal effect of each growth determinant.

\* Corresponding author. Address: Robert Day School, Claremont McKenna College, 500 E. 9th St., Claremont, CA 91711, USA. Tel.: +1 909 607 1692.  
E-mail address: [cshelton@cmc.edu](mailto:cshelton@cmc.edu) (C.A. Shelton).

More generally, we believe that “testing down” represents the proper approach in the face of model uncertainty. By starting from a general functional form and imposing restrictions such as separability, non-linearity, and monotonicity, the econometrician can learn about the true functional form with less fear of erroneous conclusions driven by omitted variables.

The rest of the paper proceeds as follows. The next section reviews the rapidly growing econometric literature on non-linearities and growth and gives an overview of our methods. Section 3 explains why coefficients from a linear model do not represent the sample average marginal effect when the true model is non-linear. Section 4 presents our data. Section 5 presents the weighted average derivative estimator and monotonicity tests. Section 6 concludes.

## 2. The relationship between our approach and prior work

There is a large and sophisticated empirical literature investigating non-linearities in the growth process with cross-country regressions. Initially focused on demonstrating heterogeneity in the parameters—the existence of multiple growth regimes—the literature has progressed to attempts to map the interactions between growth determinants. It is characterized by four main approaches<sup>1</sup>

- i. Adding quadratic terms of a variable ( $z_1$ ) thought to have a non-linear effect on growth ( $\gamma$ ) and/or interaction terms between two variables thought to be complements or substitutes ( $z_1$  and  $z_2$ ).<sup>2</sup>

$$\gamma_i = \beta'X_i + \phi_1 z_{1,i} + \phi_{11}(z_{1,i})^2 + \phi_{12} z_{1,i} z_{2,i},$$

- ii. A regression tree (Durlauf and Johnson, 1995) or threshold estimation (Hansen, 2000) procedure to partition the set of countries into several growth regimes based on whether a country is above or below endogenously determined threshold values for a particular set of variables,  $z$ . Coefficients from a linear model are then common within the regime but heterogeneous across regimes.<sup>3</sup>

$$\gamma_i = \beta'X_i$$

$$\beta = \beta_j \text{ iff } z_i \in Z_j.$$

- iii. Latent class models. As with regression trees, countries are assumed to belong to one of a finite number of growth regimes. A growth regime is defined as a distribution of growth rates conditional on observable growth determinants. Using multinomial logit, the parameters of each regime are estimated along with the probability that a given country is a member of each regime.<sup>4</sup>

- iv. Partially linear models in which a single variable is allowed to enter in an unrestricted non-linear form,  $f(z)$ , which is estimated non-parametrically.<sup>5</sup>

$$\gamma_i = \beta'X_i + f(z_i).$$

Unfortunately, the paucity of data demands parsimony, ensuring that each study restricts attention to only a very few of the potential non-linear terms. Only a few variables enter nonlinearly or provide thresholds for sample splitting. Those few variables that enter nonlinearly are often additively separable, precluding interaction with other growth determinants. When interactions between variables are considered, attention is usually restricted to one or two.

While limiting the scope of non-linearities to a few chosen variables is necessary given the limited data, it is likely a consequential misspecification of the true model. Erroneous omission of non-linear terms introduces omitted variables bias of unknown sign. Cuaresma and Doppelhofer (2007) provide an example of the dangers of such misspecification. They argue that the East Asian and African dummies, which have been found to be robust determinants of growth, are actually picking up a non-linear effect of trade openness. East Asian countries exhibit high openness while African countries exhibit low openness and so the additional marginal effect of openness beyond linearity is incorrectly attributed to the regional

<sup>1</sup> See Durlauf et al., 2005 for an early survey.

<sup>2</sup> A partial list of influential papers adopting this approach includes: Barro (1996) for democracy; Banerjee and Duflo (2003) for inequality; DeJong and Ripoll (2006) for tariffs; Chang et al. (2009) for trade openness; Minier (2007a) for a variety of policy variables; and Minier (2007b) for various political institutions.

<sup>3</sup> A partial list of influential papers adopting this approach includes: Durlauf and Johnson (1995), Hansen (2000), Canova (2004) and Minier (2007a) for threshold effects based on initial income and literacy; Papageorgiou (2002) adds trade openness; Minier (2007b) looks at splits based on political institutions; Cuaresma and Doppelhofer (2007) combine threshold effects with Bayesian model averaging; and Kourtellis et al., 2007 use a variety of threshold variables and an order-independent method to show that ethno-linguistic fractionalization affects the relationship between foreign aid and growth; Sirimaneetham and Temple (2009) argue that macroeconomic policy stability is the best threshold variable; Tan (2010) finds that high quality institutions can mitigate the detrimental effects of ethnic fragmentation.

<sup>4</sup> A partial list of influential papers adopting this approach includes: Paap et al., 2005 argue that growth regimes are not well characterized by geographic clusters; Owen et al., 2009 argue that institutions are more important than geography in determining regime membership; Bos et al. (2010) estimate a stochastic production frontier to allow for inefficient use of inputs; Di Vaio and Enflo (2011) argue that the link between globalization and convergence holds only after 1950.

<sup>5</sup> A partial list of influential papers adopting this approach includes: the seminal paper by Liu and Stengos (1999) which focuses on initial income and schooling; Kalatzidakis et al., 2000 who treat the nonparametric variables as controls to focus on the robustness of the (linear) Solovian growth determinants; Durlauf et al., 2001 who estimate the nonlinear effects of the Solovian growth determinants; and Banerjee and Duflo (2003) who estimate the nonlinear effects of inequality; and Henderson et al. (2008) who confirm that several measures of policy, geography, and institutions have robustly nonlinear effects on growth.

dummies in a growth regression where trade openness is erroneously restricted to enter linearly. Moreover, they suggest that the non-linear relationship between initial income and growth, perhaps the strongest result in the recent literature on growth regimes, is not robust to model uncertainty.

Furthermore, it is often assumed that when the true specification is non-linear a linear specification will nonetheless provide the average marginal effect in a sample of countries and thus remains a useful approximation. Unfortunately, as we explain in Section 3.1, this is only the case when the explanatory variable is normally distributed, which is not the case for many growth determinants. Nonparametric approaches to point estimation offer an alternative, but typically deliver very large standard errors making inference difficult and require assumptions about the true (non-linear) functional form.

Given the potential hazards of the current approaches to modeling non-linearities in the growth process, we suggest a complementary approach. First, we note that it is possible to derive a weighted average derivative estimator (WADE) that does provide the sample average marginal effect without making assumptions about the form of the function to be estimated.<sup>6</sup> We estimate the average derivative for each of the 15 growth determinants in our data set and compare them to an OLS estimate of the conditional effect using a linear specification.

Our results must be interpreted with caution. A positive and significant average derivative is not proof of a positive local derivative. In a non-linear setting, the average slope may be positive while the slope faced by a particular country is negative or zero. Thus we pair the WADE estimates with monotonicity tests, explained below, to deliver a fuller picture.

Our second idea is based on a “testing down” approach to model selection. The conventional approach takes the linear model as a baseline and studies potential deviations from it by adding specific non-linear terms. By contrast, a “testing down” approach treats linearity and additive separability as restrictions of a general non-linear model and asks the data to tell us whether these specific restrictions can be rejected.

In the implementation of this approach, we approximate a general model using either a second or third order Taylor expansion or a flexible Fourier expansion.<sup>7</sup> Because these expansions proliferate the terms in the regression, we must limit the number of explanatory variables to preserve sufficient degrees of freedom. We do so by choosing 12 widely used growth determinants which we then sort into three groups: policy variables, economic and political institutions, and geographic and structural variables. Each group thus contains four determinants plus a constructed index summarizing all four. We then pick one variable from each group and estimate the general non-linear form with these three growth determinants plus the Solovian growth factors. We repeat for all 125 possible combinations and report the fraction of the specifications for which restrictions can be rejected. The details are discussed in Appendix B. Unfortunately for the existing literature, we reject most specifications that allow for non-linearity in a particular determinant but restrict growth determinants to be additively separable (as in *iv* above).

The result that interactions between growth determinants are non-excludable is not news. But we believe that testing restrictions of a general non-linear model offers a useful approach for the investigation of non-linearities in the growth process. In the standard approach to growth econometrics, the null hypothesis is that the partial effect of a particular variable is zero. Significance is taken as rejection of the null leading to the conclusion that the variable in question is an important growth determinant. But in the face of model uncertainty, this conclusion is problematic. In a “test down” approach, the alternative is an unrestricted functional form and the null is a specified restriction. Rejection of the null implies that the restricted functional form in question is inconsistent with the growth process.

Monotonicity is a particularly important restriction. Rejecting monotonicity for a particular growth determinant suggests potential non-linearities. Repeating our earlier methods, we test each of the 15 growth determinants (12 basic determinants plus 3 summary indices) for monotonicity in both directions. That is, we test whether the determinant has a monotonically positive effect on growth, and then test whether the variable has a monotonically negative effect on growth. A rejection of both directions is confirmation that the variable in question has non-monotonic effects on growth. Combining monotonicity tests with the WADE gives us a better picture of the marginal effect of each growth determinant.

### 3. Theoretical considerations

The analytical foundations of the linear growth regression were first presented systematically in Mankiw et al. (1992). Starting from a three-factor Cobb–Douglas aggregate production function, log-linearizing about the steady state, and assuming a common value for productivity growth they derived a growth regression that is linear in observable quantities.

$$\gamma_i = \ln A_0 + \beta \ln y_{0,i} + \beta_k \ln s_{k,i} + \beta_h \ln s_{h,i} + \beta_g \ln(n_i + g + \delta) + v_i. \quad (1)$$

where  $\gamma_i$  is the rate of per capita GDP growth in country  $i$ ,  $y_{i,0}$  is initial per capita GDP,  $s_{k,i}$  and  $s_{h,i}$  refer respectively to the rates of investment in physical and human capital,  $n_i$  is the rate of population growth, and  $g$  is the common growth rate of labor-augmenting efficiency.

In order to explain heterogeneity in initial labor-augmenting efficiency ( $A_{i,0}$ ) and steady state labor-augmenting efficiency growth ( $g_i$ ), it has become common to add additional variables  $Z_{i,t}$  as additively separable controls.

$$\gamma_i = \ln A_0 + \beta \ln y_{i,0} + \beta_k \ln s_{k,i} + \beta_h \ln s_{h,i} + \beta_g \ln(n_i + g + \delta) + \beta'_z Z_{i,t} + v_i. \quad (2)$$

<sup>6</sup> Moreover, these estimates are root- $n$  consistent and thus not subject to the curse of dimensionality.

<sup>7</sup> We implicitly assume the existence of the necessary derivatives.

Implicitly, this rests on the idea that labor-augmenting efficiency is a function of various productivity shifters such as political and economic institutions, economic policies, and various aspects of physical and economic geography such that  $A_{i,t} = A(Z_{i,t})$ .

Estimating a linear specification if the true function is non-linear is analogous to imposing the invalid restriction that non-linearities are not present. Suppose the true growth equation is an additively separable but non-linear function of the productivity shifters,  $f(Z_{i,t})$ ,

$$\gamma_i = A_0 + \beta \ln y_0 + \beta_k \ln s_{k,i} + \beta_h \ln s_{h,i} + \beta_g \ln(n_i + g + \delta) + f(Z_{i,t}) + v_i. \tag{3}$$

which can be rewritten

$$\gamma_i = A_0 + \beta \ln y_0 + \beta_k \ln s_{k,i} + \beta_h \ln s_{h,i} + \beta_g \ln(n_i + g + \delta) + \beta'_z Z_{i,t} + h(Z_{i,t}) + v_i. \tag{4}$$

where  $h(Z_{i,t}) \equiv f(Z_{i,t}) - \beta'_z Z_{i,t}$  contains the higher-order terms of  $Z$ . Estimating (2) by OLS is equivalent to invalidly imposing the restriction that  $h(Z_{i,t}) = 0$ . Even if  $f(\cdot)$  is independent of the Solovian growth factors, estimates of  $\beta_z$  will be inconsistent. Moreover, the sign of the bias depends on the sign of the covariance of  $Z_{i,t}$  with the omitted term which is generally unclear. There is thus no reason to believe that estimating  $\beta_z$  using Eq. (2) will provide useful indicators of the marginal effect of changing the productivity shifters.

### 3.1. Misspecified linear regressions do not deliver the sample average effect

Is there a meaningful interpretation to the coefficients estimated from the (misspecified) linear specification? Some authors have suggested that these estimates provide the average effect of changing the explanatory variable over the sample of countries. For example, [Helman \(2004\)](#) has argued that “estimates that exploit cross-country variations are best interpreted as average effects of trade policies on growth,” while [Temple \(2001\)](#) writes that “growth regressions are best thought of as picking up an average effect of schooling.” If this were true, then estimating a linear model may be a reasonable guide to evaluating the expected effects of changes in policies or institutional and structural reforms.

This is correct when the true model is linear but the marginal effect is (exogenously) heterogeneous across countries. However, the parameter estimates from a linear specification will usually not converge to the average partial effect when the true model is non-linear. This point was made by [White \(1980\)](#), who established that the linear estimate of an arbitrary non-linear function will converge to the *linear approximation* of the function. The linear approximation is the closest linear function in the least-squares norm to the non-linear function. The properties of this approximation are in themselves the subject of a body of mathematical literature known as approximation theory and will generally depend on the distribution of the explanatory variable (see e.g., [Rice, 1964](#) or [Rivlin, 1969](#)). In particular, the coefficients of an OLS regression will only converge to the average partial effect of an arbitrary non-linear function when the explanatory variable is normally distributed.

**Proposition 1.** *Let  $y$  be generated by the true model  $y_i = f(x_i) + \varepsilon_i$ ,  $i = 1, \dots, n$ , where  $f(x_i)$  is an arbitrary non-linear measurable function of  $x_i \in \mathfrak{R}^1$ , and  $x_i$  is distributed according to the distribution function  $H(x)$  with mean normalized to 0 and variance  $\sigma_x^2$ . Let  $E(\varepsilon_i) = 0$  and  $E(\varepsilon_i^2) = \sigma_\varepsilon^2 < \infty$ ,  $E(x_i \varepsilon_i) = 0$ ,  $E(f(x_i) \varepsilon_i) = 0$  and  $E(f(x_i)^2) = \sigma_f^2 < \infty$ . Let  $\beta = \{\beta_0, \beta_1\}$  be the vector of coefficients from an OLS regression of  $y$  on  $\{1, x\}$ . Then  $\beta_1 \xrightarrow{a.s.} E(\frac{\partial f(x)}{\partial x})$  for any function  $f(x)$  only if  $H(x)$  is the normal distribution.*

**Proof.** See [Appendix A](#)  $\square$

We can illustrate this result through a simple example. Suppose we try to estimate the function  $f(x) = x^2$  using a linear model. The partial derivative is  $f'(x) = 2x$  and the expected partial derivative will be  $E(\partial f(x)/\partial x) = 2E(x) = 2\mu$ . [Table 1](#) shows the results of a basic Monte Carlo simulation in which we have used the quadratic function  $f(x)$  as the data generating process but have estimated it with a linear specification. Column (1) shows the results of estimating it when we draw  $x$  from a normal distribution (thus satisfying the conditions of [Proposition 1](#)) while column (2) shows the results of estimating it with a log-normal distribution with exactly the same mean (1.65) and variance (4.67).

In both cases the expected partial derivative is  $E(\partial f(x)/\partial x) = 2\mu \approx 3.297$ . As [Table 1](#) shows, in the case of the normal distribution, the average slope estimator of the linear regression in 1000 replications each with sample size 100,000 is the expected 3.297. In the non-symmetric case, in contrast, the average linear coefficient estimate is 16.448, substantially higher than the expected partial derivative.

Intuitively, the asymmetry of the log-normal distribution makes observations with large values relative to the mean much more frequent than observations with small values relative to the mean. For a given mean (and thus a given average marginal effect of  $x$  on  $y$ ), these outliers, while infrequent, are strongly penalized by least-squares and thus pull the least squares line upwards relative to the true expected marginal effect. As a result, the linear specification does not accurately deliver the average marginal effect of  $x$  on  $y$  when the true model is non-linear and  $x$  is non-normally distributed.

Columns (3) and (4) show that using instrumental variables does not solve the problem. Columns (5) through (8) repeat the exercise for two other data generating processes (DGP). First, we use a spline function to generate the data and then run a

**Table 1**Monte Carlo simulation: Estimating linear models for non-linear data generating processes under different distributions of  $x$ .

	1	2	3	4	5	6	7	8
	Quadratic		Quadratic		Spline		Interaction	
Data generating process	$y = x^2 + \varepsilon$		$y = x^2 + \varepsilon + u$ $x = z + u$		$y = \varepsilon$ if $x < E[x]$ $x + \varepsilon$ if $x \geq E[x]$		$y = x + w + xw + \varepsilon$	
Estimated equation	$y = a_0 + a_1x + e$		$y = a_0 + a_1x + e$ $x = b_0 + b_1z$		$y = a_0 + a_1x + e$		$y = a_0 + a_1x + a_2w + e$	
Method of estimation	OLS		IV		OLS		OLS	
Distribution of $x^a$	Normal	Log-normal	Normal	Log-normal	Normal	Log-normal	Normal	Log-normal
Mean of $x$	1.649	1.649	1.649	1.649	1.649	1.649	1.649	1.649
Std. dev of $x$	2.161	2.161	2.161	2.161	2.161	2.161	2.161	2.161
$E(dy/dx)$	3.297	3.297	3.297	3.297	0.500	0.308	2.649	2.649
Mean $\hat{a}_1$	3.297	16.448	3.297	16.505	0.500	16.621	2.649	0.353
Sample size	100,000	100,000	100,000	100,000	100,000	100,000	100,000	100,000
Replications	1000	1000	1000	1000	1000	1000	1000	1000

Data on  $x$  are drawn from either a normal or a lognormal distribution. Corresponding data from  $y$  are generated according to the quadratic process. Then a linear model is estimated using either OLS or IV. When the data are normally distributed, the linear specification correctly estimates the average slope of the underlying non-linear function ( $\sim 3.3$ ). When the data are log-normally distributed, the least-squares estimator excessively penalizes the long-right tail, pulling up the regression line. As a result, the average slope is vastly overestimated (16.4). This is equally true when using IV. In columns 5–6, the data generating process is a spline function and the estimation remains ignorant of the kink. Once again, the OLS slope estimate equals the average partial effect only if the explanatory variable is normally distributed. In columns 7–8, the data generating process includes an interaction between two independent variables that is omitted from the estimated model. Whereas in previous examples, the OLS coefficient overestimated the average partial effect when the explanatory variables are log-normally distributed, here it underestimates.

<sup>a</sup> In columns 7 and 8,  $x$  and  $w$  are drawn iid from either the normal (7) or log-normal (8) distribution.

linear regression that is unaware of the kink in the true DGP. As with the quadratic case, a log-normal distribution of the independent variable results in an OLS coefficient that greatly overestimates the average partial effect. Second, we consider a DGP with two independent variables that enter separately and through an interaction. We then run a linear regression containing the correct set of variables but missing the interaction term. This time log-normally distributed explanatory variables lead to OLS coefficients that underestimate the average partial effect.

Table 3 shows the results of skewness, kurtosis, and joint normality tests of 12 common explanatory variables in growth empirics (these variables are discussed in Section 4).<sup>8</sup> According to these standard tests, in all 12 cases we reject the normality hypothesis, and in 11 of them we do so with  $p$ -values lower than 0.01. Thus we cannot expect the linear model to yield reasonable estimates of the average marginal effect when the data-generating process is non-linear.

### 3.2. The problem with instruments

Considering misspecification bias arising from ignored non-linearities in (2.21) as a special case of omitted variable bias suggests the use of instrumental variables. Regrettably, this will generally not be possible. Any candidate instrument that is correlated with  $Z$  is also likely to be correlated with  $h(Z_{10}, \dots, Z_{p0})$ . Since the misspecified regression treats  $h(\cdot)$  as part of the disturbance term, the instrument will be correlated with the error term in the structural equation, rendering 2SLS estimates inconsistent.

For the same reason, omitted non-linearities will generally make instruments that would be valid in a linear framework yield biased estimates. To demonstrate the point, we extend our Monte Carlo simulation from the previous subsection. As before, consider a linear regression of  $y$  on  $x$  when the true relationship is quadratic. Suppose that we have an instrument,  $z$ , which is valid in a linear context. That is, the instrument,  $z$ , is correlated with the dependent variable,  $y$ , only through the variable  $x$ . The results are reported in columns (3) and (4) of Table 1. As before, the coefficient estimates are consistent when the explanatory variable,  $x$ , is normally distributed but biased when the distribution of  $x$  is non-normal.

This result is particularly important in light of frequent use of instrumental variables techniques to study the causes of cross-country variations in growth and income. Authors tend to argue for the validity of their instrument based on its causal relationship with the endogenous variable, and its lack of a direct effect on the dependent variable (e.g. Frankel and Romer, 1999; Acemoglu et al., 2001). However, these exercises generally do not discuss the effect of omitted non-linearities on the validity of their estimates. To the extent that the relation between trade, institutions, and development is characterized by strong non-linearities, these exercises may yield biased estimates of the average growth effects of their explanatory variables.<sup>9</sup>

<sup>8</sup> These are the same variables we will use in our empirical analysis of Section 3. Their definitions are presented in Table 2.

<sup>9</sup> This is a variation on the point made by Brock and Durlauf (2001) who argue that instruments for endogenous growth factors may be pre-determined and yet invalid if they are nonetheless correlated with omitted growth factors.

#### 4. The data

We use a standard cross-sectional data set of economy-wide measures of growth and its potential determinants for the 1960–2007 period. Despite the recent expansion of use of panel-data methods in cross-country growth analysis, we restrict attention to the cross-sectional framework for several reasons.<sup>10</sup> We use the Penn World Table (PWT) 6.3 and World-Bank PPP-adjusted per capita GDP Growth Rates from the *World Development Indicators* (WDI) as our dependent variables. WDI data are available for the 1975–2007 period, while PWT data are available for the 1960–2007 period. Given that a number of explanatory variables are not available for the sixties or early seventies, the shorter period may be preferable – for this reason we also present regression results where the PWT data are restricted to the 1975–2007 period.

As production–function shifters,  $Z$ , we use combinations of 12 commonly used measures, as well as three summary indicators made up of subgroups of these. The sample attempts to cover the three key dimensions that have played relevant roles in the analysis of growth empirics: policies, institutions and economic structure (including geography). Within the limitations of parsimony and symmetry (an identical but relatively small number of indicators in each group), we have tried to choose widely used indicators capturing influential theories of growth. To measure policy distortions, we use government consumption as a percent of GDP, the average tax on imports and exports, the log of one plus the inflation rate, and the log of the black market premium. Inflation and government consumption were suggested by Barro (1996) and have seen extensive use since (e.g. Easterly and Levine, 2003). Easterly and Levine (2003) likewise use an overvalued exchange rate as a third potential policy distortion (we use the black market premium). There is an extensive literature on the effects of trade openness following seminal work by Sachs and Warner (1995) generally finding that trade openness aids growth and technological diffusion (e.g. Lucas, 2009). To capture the role of institutions, we use four familiar indicators: a measure of the rule of law, a measure of political instability, an index of the effectiveness of government spending, and an index of economic freedom. The first three are a subset of those suggested by Kaufman et al. (1999) which have then seen extensive use (again see Easterly and Levine, 2003). Economic freedom speaks to an orthogonal set of institutional constraints on effective use of inputs. Our structural measures of the level of social development and economic modernization include the share of primary exports in GDP in 1970, the ratio of liquid liabilities to GDP, the average years of life expectancy, and the absolute latitude. These variables are similarly broadly used: the share of primary goods in exports addresses the natural resource curse; absolute latitude was made famous by Hall and Jones (1999) and Barro and Sala-i-Martin (2004) highlight the role of life expectancy; and the importance of the depth of financial markets, measured by liquid liabilities as a share of GDP, has been highlighted by Beck et al. (2000) among others. The summary indicator for each group is calculated as the unweighted normalized average of the relevant indicators. A full description of the variables and their sources is provided in Table 2.

As is common in the literature, we estimate (3) with a restricted subset of the variables available in the data set in order to economize on degrees of freedom and reduce the potential for multicollinearity arising from the fact that several indicators may be capturing closely related phenomena. Given that the results can be sensitive to the choice of indicators, we present estimates for all possible specifications with one policy indicator ( $z_{p,i}$ ), one institutional variable ( $z_{i,i}$ ), and one measure of structural characteristics ( $z_{s,i}$ ) plus the Solovian growth factors.

$$\gamma_i = A_0 + \beta \ln y_0 + \beta_k \ln s_{k,i} + \beta_h \ln s_{h,i} + \beta_g \ln(n_i + g + \delta) + f(z_{p,i}, z_{i,i}, z_{s,i}) + v_i. \quad (5)$$

With five choices for each of these three spots, each specification is estimated 125 times. We report the fraction of specifications for which a given hypothesis is rejected. This is not the same as Bayesian model averaging because we simply weight all models equally rather than calculating priors and updating them. Given the complexities of our specification tests, Bayesian model averaging is not feasible. Under this restriction, we believe our method is a reasonable second-best approach which retains the spirit of agnosticism regarding the true model.

#### 5. WADE and monotonicity tests

##### 5.1. The weighted average derivative estimator (WADE)

Calculating the WADE of  $y$  with respect to a single variable  $x$  consists of three steps. First, non-parametric methods are used to estimate the local conditional mean  $E[y|x]$ . Second, the local derivative of  $E[y|x]$  with respect to  $x$ ,  $\delta(x)$ , is estimated analytically by calculating the kernel regression estimate at  $x + \lambda/2$  and  $x - \lambda/2$  and dividing the difference by  $\lambda$ . Finally, the

<sup>10</sup> First, the cross-sectional approach is still broadly used including several important contributions. Some examples are Frankel and Romer (1999), Acemoglu et al. (2001) and Sala-i-Martin et al., 2004. The first two articles use a levels specification, whereas the third uses the growth specification that we reproduce here. For a critique of the levels approach, see Sachs (2005). Second, relevant methodological questions remain about the applicability of the panel data approach to study questions of long-term economic growth. For example, it is not clear that segmenting the data into 10 or 5-year intervals is appropriate when the phenomenon of interest is development over long periods of time, and most existing panel methods used require the introduction of fixed effects, impeding the analysis of the effect of potential growth determinants, such as institutions or geography, which exhibit little or no variation over time. Standard random effects estimators require the random effect to be uncorrelated with the explanatory variables, which is by construction not the case in a growth regression. See Durlauf et al., 2005 for a discussion. Third, the theory behind the specification tests presented in this section is not at this moment fully developed for its application to a panel context.

**Table 2**  
Data sources.

Policy indicators	
1. Trade openness	$(1 + t_m)(1 + t_e) - 1$ where $t_m$ ( $t_e$ ) is the ratio of import (export) tax revenue to total imports (exports)
2. Log of black market premium	Dollar and Kraay (2002)
3. Government consumption as a percentage of GDP	World Bank (2005a,b)
4. Log of (1 + inflation rate)	World Bank (2005a,b)
5. Policy summary indicator	Sum of 1–4, normalized over unit interval
Institutional indicators	
6. Rule of law	Dollar and Kraay (2002)
7. Political instability	Average variation in POLITY score (Polity IV)
8. Effectiveness of government spending	Glaeser et al. (2004)
9. Economic freedom index	Heritage foundation
10. Institutions summary indicator	Sum of 6–9, normalized over unit interval
Economic structure indicators	
11. Share of primary exports in total exports	World Bank (2005a,b)
12. Absolute latitude	World Bank (2005a,b)
13. Life expectancy	World Bank (2005a,b)
14. Liquid liabilities as a share of GDP	IMF (2004)
15. Structure summary indicator	Sum of 11–14, normalized over unit interval

WADE is the sample average of this derivative. Härdle and Stoker (1989) and Rilstone (1991) provide many of the details for implementing WADE including the calculation of standard errors. We use Xplore's implementation of Härdle and Stoker's indirect average derivative estimator with smoothing parameter (akin to kernel bandwidth)  $m = 1.5$ .

The WADE has two important advantages over other estimators in the growth context. First, it does not require parametric assumptions which might lead to omitted variables bias. Second, it is  $\sqrt{n}$ -consistent and thus not subject to the curse of dimensionality (Stone, 1980). As Rilstone explains, "with non-parametric point estimation, inferences are effectively based only on those observations lying within a small window-width of the point of interest, estimates tend to have high standard errors and hypothesis tests tend to have low power for moderately sized samples. . . On the other hand, by averaging these point estimates, one uses all the information in the sample and. . . the result is a  $\sqrt{n}$ -consistent estimate." (Rilstone, 1991, p. 210) At the same time, because the derivative is estimated locally and non-parametrically, the WADE does not suffer from the problems documented in Section 3.1: it is consistent for any sample distribution of  $x$ .

We estimate the weighted average derivative for an unrestricted specification in which we use a flexible Fourier series to model  $f(z_{p,i}, z_{l,i}, z_{s,i})$ .<sup>11</sup>

$$\gamma_i = A_0 + \beta \ln y_0 + \beta_k \ln s_{k,i} + \beta_h \ln s_{h,i} + \beta_g \ln(n_i + g + \delta) + \sum_{i \in \{P, I, S\}} b_i z_i + \sum_{ij \in \{P, I, S\}} c_{ij} z_i z_j + \sum_{k \in K} \sum_{l=1}^L \{ u_l \cos(l \cdot k'z) + v_l \sin(l \cdot k'z) \} + v_i. \tag{6}$$

The results are reported in Table 6 column two. For comparison, column one displays the results from OLS estimations of an additively separable linear specification.

$$\gamma_i = A_0 + \beta \ln y_0 + \beta_k \ln s_{k,i} + \beta_h \ln s_{h,i} + \beta_g \ln(n_i + g + \delta) + \beta_p z_{p,i} + \beta_l z_{l,i} + \beta_s z_{s,i} + v_i. \tag{7}$$

This is not a classic regression table: each cell represents a separate regression targeting a particular production shifter. The production shifter in question (row) is matched with the indices from the other two groups to form a particular set  $z_p, z_l, z_s$  which are then used, along with the Solovian growth factors in the given specification (column). The cell then reports (only) the coefficient and standard error for the production shifter in question.

For example, the second cell of column one represents OLS estimation of Eq. (7) where  $z_p, z_l, z_s$  are the black market premium, and the indices for institutions and economic structure. The cell reports the coefficient on black market premium. The first column shows that the OLS estimate of the partial effect of the black market premium on growth is negative but insignificant; the WADE estimate remains negative, is smaller in magnitude, but strongly significant. Obviously, one cannot recover the local partial effect. But one possible explanation is that the partial effect of the black market premium is non-linear and varies from weakly to strongly negative but with the bulk of the data in the weak zone. In this case, the OLS estimate would be relatively imprecise and the local derivatives that are averaged to form the WADE would be more precise. Both procedures would more heavily weight the weak relationship because that is where the bulk of the data lies, but as we showed in Section 3, OLS may be excessively skewed by the extreme values at far right so as to avoid large errors.

Thus each row in the table shows the difference between using OLS and WADE to estimate the sample average treatment effect of a particular explanatory variable, when controlling for the indices of the other variables. We chose to focus on this

<sup>11</sup> See Appendix B for further discussion of the flexible Fourier approximation.

**Table 3**  
Skewness, kurtosis and normality tests of some common explanatory variables.

Explanatory variable	Number of observations	Skewness <i>p</i> -value	Kurtosis <i>p</i> -value	Joint normality test	
				$\chi^2$	<i>p</i> -Value
Log(1 + inflation)	137	0.00	0.00	129.95	0.00
Black market premium	132	0.00	0.00	133.65	0.00
Government consumption	133	0.00	0.32	11.43	0.00
Tariff rate	150	0.00	0.00	137.14	0.00
Rule of law	112	0.06	0.01	11.58	0.00
Political instability	121	0.00	0.00	103.43	0.00
Economic freedom index	124	0.02	0.06	9.13	0.01
Index of government effectiveness	133	0.02	0.28	6.51	0.04
Primary exports in 1970	112	0.00	0.47	14.95	0.00
Absolute latitude	133	0.00	0.17	10.82	0.00
Life expectancy	137	0.02	0.00	48.79	0.00
Liquid liabilities/GDP	126	0.00	0.00	118.91	0.00

All values refer to the D'Agostino et al. (1990) tests for skewness, kurtosis and normality. The joint test rejects the hypothesis that the variable is normally distributed for all 12 growth factors at the 5% significance level and 11 of the 12 at the 1% significance level. All 12 variables display considerable skewness and eight of the 12 are either meso- or leptokurtic.

**Table 4**  
Linearity tests, 2nd and 3rd order polynomial expansions.

Explanatory variable	Penn world table, 1975–2007	World bank, 1975–2007	Penn world table, 1960–2007
<i>Second-order polynomial</i>			
Median <i>F</i> -statistic	3.07	3.42	3.66
Median <i>p</i> -value	0.0110	0.00497	0.00454
Number significant (/125)	84	94	86
Percent significant	67.2%	75.2%	68.8%
<i>Third-order polynomial</i>			
Median <i>F</i> -statistic	7.69	5.04	7.64
Median <i>p</i> -value	7.54E–09	2.57E–06	3.80E–08
Number significant (/125)	125	118	122
Percent significant	100.0%	94.4%	97.6%

For each of the 125 combinations of ( $z_p, z_i, z_s$ ), we take a Taylor expansion (2nd or 3rd order) around the means. We then test whether the interaction terms in the expansion are jointly zero using a conventional *F* test. This is the hypothesis that the linear model is correct. We report the fraction of these 125 regressions for which the hypothesis is rejected at the 5% significance level. The linear hypothesis is rejected in 2/3–3/4 of the cases for the second order expansion and over 94% of the cases in the third order expansion, confirming the linear model is not a close approximation of the data.

subset rather than report all 125 possible combinations of institutional shifters as before because our primary goal in this section is to compare different methods of estimation in the face of unknown non-linearities. Significance tests are based on bootstrapped standard errors with 100 replications. The WADE estimates are generally much more precise: the standard errors of the reported WADE estimates range from 15% to 64% of the standard errors of their OLS counterparts. At the same time, the point estimates are, in most cases, considerably different.

On the whole, compared to the OLS estimates, our results show much weaker effects for policy variables and the political environment (inflation, black market premium, government consumption, political stability, government effectiveness) and much more robust effects for structural variables and the economic environment (rule of law, economic freedom, geography, life expectancy, depth of financial markets). We will discuss these results further in Section 6 after presenting the monotonicity tests.

### 5.2. Testing for monotonicity

Formally, testing for monotonicity is implemented by testing the null hypothesis:

$$H_0 : z_i > z'_i, z_{-i} = z'_{-i} \rightarrow f(z_1, \dots, z_m) \geq f(z'_1, \dots, z'_m), \tag{8}$$

against the alternative:

$$H_1 : \exists(z_1, \dots, z_m), (z'_1, \dots, z'_m) | z_i > z'_i, z_{-i} = z'_{-i}, f(z_1, \dots, z_m) < f(z'_1, \dots, z'_m). \tag{9}$$

Using the framework described in Section B.2, we test by imposing monotonicity as a restriction on the estimated  $\hat{f}(\cdot)$  and calculating the *HW* statistic. We explicitly calculate the derivative of the Fourier expansion and directly impose the restriction

**Table 5**  
Separability tests: Series expansions.

	Taylor	
	Polynomial	Fourier series
Penn world table, 1975–2007		
Median statistic	3.9203224	1.9032264
Number significant (/125)	100	64
Percent significant (5%)	80.00%	51.20%
World bank, 1975–2007		
Median statistic	2.7087586	1.3581369
Number significant (/125)	81	58
Percent significant (5%)	64.80%	46.40%
Penn world table, 1960–2007		
Median statistic	3.4333341	1.1350211
Number significant (/125)	85	50
Percent significant (5%)	68.00%	40.00%

For each of the 125 combinations of  $(z_p, z_f, z_s)$ , we take an expansion (Taylor or flexible Fourier) around the means. We then test the joint restriction that all interactions between distinct productivity shifters are zero. This is the hypothesis of separability. We report the fraction of these 125 regressions for which the hypothesis is rejected at the 5% significance level. The separability hypothesis is rejected in 2/3–4/5 of the cases for the second order Taylor expansion and 2/5–1/2 of the cases in the flexible Fourier expansion, suggesting the true model is probably not separable. Test for the Taylor expansion are a standard  $F$  test. Tests for the flexible Fourier are the Hong-White statistic of Eq. (10). Confidence intervals are built using bootstrapped test statistics with 100 replications.

$$\frac{\partial \hat{f}}{\partial z_i}(\cdot) \geq 0. \quad (10)$$

at all observed values of  $z_i$ . The sum of squared residuals is minimized subject to (10) to obtain the restricted estimator  $\hat{f}_r(\cdot)$  and calculate the  $HW$  statistic. This is a non-linear optimization problem subject to an inequality restriction that can be solved numerically. The continuity of the first derivatives of the Fourier representation ensures that imposing (8) on all observations will deliver a reasonably smooth function.<sup>12</sup>

Table 7 displays the results of the Fourier monotonicity tests. For each variable, we test for monotonicity in both directions: monotonically positive effect on growth and monotonically negative effect on growth. But for each variable, there is a conventional wisdom from the previous literature that suggests one of these directions. The conventional wisdom says government distortions are detrimental for growth, while certain institutions and structural changes are beneficial. We have put this conventional wisdom in the left column and the opposite on the right.

## 6. Discussion

### 6.1. Interpreting the results

Combining the monotonicity tests with the weighted average derivative estimate suggests that some factors have strong and consistent growth effects that are likely to apply to the majority of the sample while other factors display weak or variable effects. Strikingly, 3 of the 4 structural factors and 3 of the 4 institutional factors display clear effects along the lines of the conventional wisdom in the growth literature, while only 1 of the 4 policy variables does so.

For rule of law, political stability, economic freedom, absolute latitude, and life expectancy the average derivative estimates are all positive. The WADE coefficients tend to be smaller than the OLS coefficients but because the standard errors are smaller, the effects remain strongly significant. Moreover, we cannot reject positive monotonicity. One of the policy-variables—the black market premium—displays a similarly clear picture. The WADE is significant and negative and we cannot reject negative monotonicity. These results are thus consistent with the conventional wisdom that improvements in these factors are likely to be growth enhancing for any country.

The effect of deeper financial markets is slightly more complicated. The WADE supports a significant positive average marginal effect on growth of deeper financial markets. However, monotonicity is rejected in both directions, suggesting that there are some cases in which the marginal effect is negative.

For inflation and government consumption, the OLS coefficient is negative but the WADE coefficient, despite smaller standard errors, is insignificant. Monotonicity cannot be rejected in either direction. Together, these suggest that inflation and

<sup>12</sup> All problems were solved using the CONOPT solver in GAMS; the code is available from the authors upon request. For the Fourier series expansion, the sum of squared residuals was minimized subject to the explicit constraint that the analytic first-derivative has the prescribed sign. For the residual regression and differencing tests (unreported, available on request), we first obtain a non-parametric estimate of  $\hat{f}(\cdot)$  by penalized thin-plate regression spline estimation using the `gam` command in R; we then calculate the restricted estimate by finding the closest set of points to the fitted function that satisfy the monotonicity constraints in GAMS. See Mammen (1991) for a discussion of this method for construction of monotonic estimates.

**Table 6**  
Using WADE to relax the linearity and separability assumptions.

	$z_p + z_l + z_s$	$h(z_p, z_l, z_s)$
Specification for productivity shifters	OLS	WADE
Method of estimation	OLS	WADE
Log(1 + inflation)	−0.0182 <sup>***</sup> (0.0067)	−0.0033 (0.0028)
Black market premium	−0.0244 (0.0162)	−0.0051 <sup>**</sup> (0.0025)
Government consumption	−0.0223 <sup>***</sup> (0.0078)	−0.0049 (0.0031)
Tariff rate	0.0109 (0.0197)	−0.0025 (0.0036)
Policy index	−0.0258 <sup>***</sup> (0.0067)	−0.0072 <sup>*</sup> (0.0041)
Rule of law	−0.0030 (0.0065)	0.0081 <sup>***</sup> (0.0031)
Political stability	0.0214 <sup>**</sup> (0.0086)	0.0094 <sup>**</sup> (0.0042)
Economic freedom index	0.0250 <sup>**</sup> (0.0113)	0.0115 <sup>***</sup> (0.0036)
Index of government effectiveness	0.0170 <sup>*</sup> (0.0100)	0.0038 (0.0037)
Institutions index	0.0216 <sup>*</sup> (0.0111)	0.0104 <sup>***</sup> (0.0040)
Primary export share in 1970	0.0024 (0.0065)	0.0035 (0.0042)
Absolute latitude	0.0112 (0.0076)	0.0101 <sup>*</sup> (0.0052)
Life expectancy	0.0328 <sup>**</sup> (0.0103)	0.0104 <sup>**</sup> (0.0050)
Liquid liabilities/GDP	0.0214 (0.0168)	0.0167 <sup>***</sup> (0.0049)
Economic structure index	0.0226 <sup>**</sup> (0.0094)	0.0137 <sup>***</sup> (0.0043)

This is not a standard regression table. Each cell reports one coefficient from a separate regression (so there are 30 regressions here). In turn, we choose one of the 15 indicators, pair it with the indices from the other two dimensions and run two separate regressions with that choice of  $(z_p, z_l, z_s)$  plus the Solovian growth factors  $(s_k, s_n, n)$ . The first uses these three indicators in the separable, linear form standard to growth regressions and is estimated using OLS. The second imposes neither linearity nor separability and is estimated using the WADE. We repeat for each of the 15 indicators. For each cell we report the coefficient on the indicator at left (row) for the regression form at top (column). OLS: Standard errors in parenthesis. WADE: The semi-parametric estimate is obtained using Xplore's implementation of Härdle and Stoker's (1989) indirect average derivative estimator, with  $m = 1.5$ . Bootstrapped standard errors, with 200 replications, in parenthesis.

<sup>\*</sup> Significance at  $-10\%$ .

<sup>\*\*</sup> Significance at  $-5\%$  level.

<sup>\*\*\*</sup> Significance at  $-1\%$  level.

government spending are either never a strong growth determinant or their effect is non-monotonic. Either way, the implication of the least squares results—that greater government spending and higher inflation have a consistent and significant negative effect on growth—requires reexamination. The same is true for the average tariff rate: further opening to trade is not always and everywhere a recipe for significant growth.

Our results also cast doubt on the panacea of effective government. While the least squares coefficient is positive and significant, the WADE coefficient is indistinguishable from zero despite much greater precision. Moreover, we strongly reject monotonicity in both directions. The resulting conclusion is that improving the effectiveness of government can be associated with either faster or slower growth. Perhaps this is because the marginal effect of government effectiveness on growth depends on whether or not the policies more effectively implemented are good for growth.

Finally, the share of the primary sector in exports is also complex. The WADE estimate is statistically indistinguishable from zero. While we can reject monotonically decreasing, we cannot reject monotonically increasing. Thus having a large primary sector is either never terribly relevant for growth, or it is sometimes a help and sometimes a hindrance. These results dovetail with the resource curse literature, which has generally found that primary resource dependence can be either good or bad for growth depending on how it is used and that different institutional frameworks seem to be more or less capable of channeling resource wealth productively.

In general, our results point to several growth determinants whose benefits are likely conditioned on the state of other policies and institutions. Our result that the benefits of trade openness are conditional has already been examined by [Chang et al. \(2009\)](#), who find that openness to trade is beneficial to growth only if a country has a degree of labor market flexibility. We agree with [Easterly and Levine \(2003\)](#) that macroeconomic policies are less clearly consequential than institutions. On the other hand, while much of the recent literature on growth regimes has downplayed the importance of geography (e.g.

**Table 7**  
Monotonicity tests.

Null hypothesis	Hong-White Fourier series test	
	Conventional Wisdom	Opposite
	Decreasing	Increasing
Log(1 + inflation)	0.86	0.46
Black market premium	0.47	0.10
Government consumption	0.86	0.25
Tariff rate	0.20	0.10
Combined policy index	0.53	0.00
Null hypothesis	Increasing	Decreasing
Rule of law	0.21	0.15
Political stability	0.10	0.03
Economic freedom index	0.14	0.02
Index of government effectiveness	0.02	0.00
Combined institutions index	0.10	0.00
Null hypothesis	Increasing	Decreasing
Primary exports in 1970	0.33	0.06
Absolute latitude	0.07	0.01
Life expectancy	0.15	0.00
Liquid liabilities/GDP	0.03	0.01
Combined structure index	0.22	0.00

The table reports  $p$ -values for the null hypothesis of monotonicity. The Hong-White test is based on the HW statistic in Eq. (B4). Confidence intervals are built using bootstrapped test statistics with 100 replications.

Paap et al., 2005; Owen et al., 2009; Tan, 2010), we find an important role for geographic and structural variables. We believe our other results are leads on fruitful areas for similar studies of policy complementarities.

## 6.2. Conclusion

We have shown that it is possible to reach conclusions that place significant restrictions on the shape of the growth function without imposing unfounded restrictions. Monotonicity tests suggest that several institutional and structural reforms are beneficial for at least a number of countries, but that changes to an important subset of variables – such as government effectiveness and the depth of financial markets – can prove either beneficial or detrimental to growth depending on the country in question. The weighted average derivative and monotonicity tests also suggest that higher inflation and larger government are not necessarily growth inhibiting, having either insubstantial or ambiguous effects. And overall, our evidence suggests that structural and institutional factors are of stronger and clearer importance than monetary, fiscal, and trade policies and should be given more weight in reforms.

More generally, our approach suggests that the empirical growth literature may be better served by analyzing general, “deeper” hypotheses about the growth process than in trying to reach an exact characterization of its form. Testing restricted specifications can give guidance on functional forms and suggest which growth factors are likely complements. Weighted average derivative estimators can then deliver a consistently estimated sample-average marginal effect despite a complex true model and non-normally distributed growth factors. Together, these tools can, in spite of complex functional forms and limited data, further our understanding of the growth process.

## Appendix A. Proof of Proposition 1

**Proposition 1.** Let  $y$  be generated by the true model  $y_i = f(x_i) + \varepsilon_i$ ,  $i = 1, \dots, n$ , where  $f(x_i)$  is an arbitrary non-linear measurable function of  $x_i \in \mathbb{R}$  and  $x_i$  is distributed according to the distribution function  $H(x)$  with mean normalized to 0 and variance  $\sigma_x^2$ . Let  $E(\varepsilon_i) = 0$  and  $E(\varepsilon_i^2) = \sigma_\varepsilon^2 < \infty$ ,  $E(x_i \varepsilon_i) = 0$ ,  $E(f(x_i) \varepsilon_i) = 0$  and  $E(f(x_i)^2) = \sigma_f^2 < \infty$ . Let  $\beta = \{\beta_0, \beta_1\}$  be the vector of coefficients from an OLS regression of  $y$  on  $\{1, x\}$ . Then  $\beta_1 \xrightarrow{a.s.} E(\frac{\partial f(x)}{\partial x})$  for any function  $f(x)$  only if  $H(x)$  is the normal distribution:

**Proof.** By White (1980, Theorem 2),  $\beta \xrightarrow{a.s.} \beta^*$ , the vector which uniquely solves  $\text{Min}_\beta \int [f(x) - x\beta]^2 dH(x)$ . This vector is the coefficient vector on  $(1, x)$  in the linear projection of  $f(x)$  on  $(1, x)$ . Particularly,

$$\beta_1^* = \frac{\text{Cov}(x, f(x))}{\text{Var}(x)} = \frac{E(xf(x))}{\sigma_x^2},$$

where we have used the normalization  $E(x) = 0$ . Now let us approximate  $f(x)$  by an  $n$ th order Taylor expansion at  $x = 0$ , that is:

$$f(x) \approx P(f(x)) \equiv \alpha_0 + \sum_{k=1}^m \alpha_n x^n,$$

where:

$$\alpha_0 = f(0)$$

$$\alpha_n = \frac{\partial^n}{\partial x^n} f(0).$$

Since the linear projection is a linear operator (see Wooldridge, p. 32) then:

$$\beta_1^* = \sum_{i=1}^n \alpha_n \frac{E(x \cdot x^n)}{\sigma_x^2}.$$

Therefore  $\beta_1^*$  will equal  $E\left(\frac{\partial f(x)}{\partial x}\right)$  for any  $\alpha_n$  only if:

$$\beta_1^* = \frac{1}{\sigma_x^2} E(x^{n+1}) = nE(x^{n-1}).$$

Since  $E(x) = 0$ , (2) implies  $E(x^q) = 0$  for any  $q > 1$  odd. For  $q \geq 2$  even, repeated substitution gives  $E(z_d^q) = \sigma_{dd}^{q/2} (q - 1)(q - 3) \dots 3 \cdot 1$ . These are exactly the moments of the normal distribution (see, e.g., Rohatgi, 1976, pp. 220–221). Since the normal distribution is  $M$ -determinate (Stoyanov, 2000), then  $H(x)$  is the normal distribution.  $\square$

## Appendix B. Tests of linearity and separability

### B.1. Testing the linear, additively separable specification

The null hypothesis of linearity can be tested within the framework of traditional OLS estimation. Under the null that growth is a linear function of its determinants, non-linear terms should not enter significantly into the regression. A simple test of linearity can be carried out by augmenting the linear specification with a series approximation and testing for the joint excludability of the higher-order terms. We estimate

$$\gamma_i = A_0 + \beta \ln y_0 + \beta_k \ln s_{k,i} + \beta_h \ln s_{h,i} + \beta_g \ln(n_i + g + \delta) + \beta_p z_{p,i} + \beta_l z_{l,i} + \beta_s z_{s,i} + h(z_{p,i}, z_{l,i}, z_{s,i}) + v_i. \tag{B1}$$

where  $z_p, z_l, z_s$  stand, respectively, for the policy, institutional and structural indicators and  $h(z_p, z_l, z_s)$  groups the higher-order components of the series approximation. We first approximate  $h(z_p, z_l, z_s)$  with a 2nd or 3rd order Taylor expansion around the means of  $z_p, z_l, z_s$  and then test the null hypothesis that all terms in  $h(\cdot)$  are jointly zero.

Table 4 reports the percentage of specifications (out of the 125 regressions generated by alternative combinations of the  $z$  variables) for which the null hypothesis is rejected as well as the median  $F$ -statistic and associated  $p$ -value for rejection of the null hypothesis.

The results of Table 4 present relatively strong rejections of the linearity hypothesis. Between two thirds and three fourths of the specifications that use a 2nd order Taylor expansion and more than 94% of the specifications that use a 3rd order Taylor expansion reject the separable, linear specification. The fact that the rejections become stronger as the order of the polynomial is increased is not automatic:  $F$ -tests for joint significance can and often do become weaker when new variables are added. This pattern suggests that the higher-order polynomial terms may be playing an important role in explaining growth differences across countries. Moreover, these rejections remain similarly strong when we use more general non-parametric methods to approximate possible non-linearities.<sup>13,14</sup>

### B.2. Testing additive separability with potential non-linearity

The results of the preceding section tell us that the linearity assumption can be rejected in existing cross-country data sets, but leave us little clue as to the form of the actual non-linearity. In the spirit of “testing down”, we start out from a very general form of the non-linearity and study the restrictions that can be imposed on it without excessive loss of fit. One key issue in estimation of a non-linear multi-dimensional function is whether it can be taken to be additively separable.

We start by testing additive separability of  $h(Z)$  in our partially linear specification. This is done by testing the additively separable specification:

<sup>13</sup> It is possible that rejections of the separable, linear specification spring from incorrectly imposing equal productivity growth rates. We have repeated the analysis of this section under the assumption that productivity growth is a linear function of the change in our growth factors,  $z_p, z_l, z_s$ . The rejections of the separable, linear model are even stronger. Results are available on request.

<sup>14</sup> Thus far we have focused on nonlinearities in the production shifters while maintaining linearity in the Solovian growth factors. It is possible that nonlinearities in the Solovian growth factors, of which several authors have provided evidence including Durlauf and Johnson (1995), Hansen (2000), Durlauf et al., 2001, Canova (2004) and Minier (2007a), lessen the impact of nonlinearities in the production shifters which are the focus of this paper. Future research on the interaction of production shifters and Solovian factors may be productive.

$$\gamma_i = A_0 + \beta \ln y_0 + \beta_k \ln s_{k,i} + \beta_h \ln s_{h,i} + \beta_g \ln(n_i + g + \delta) + \beta_p z_{p,i} + \beta_l z_{l,i} + \beta_s z_{s,i} + f_p(z_{p,i}) + f_l(z_{l,i}) + f_s(z_{s,i}) + v_i. \quad (B2)$$

against the more general model of (B1). In order to do this, we carry out two different tests that are broadly used in the literature on estimation of non-parametric and semi-parametric methods. These are briefly described in what follows.

The first approach mimics the previous section: model  $h(z_p, z_l, z_s)$  with a Taylor series approximation and test the joint restriction that all interactions between distinct production shifters are zero. The results of this  $F$ -test are reported in the first column of Table 5. Rejection rates for the separability hypothesis at the 5% level of significance oscillate between 64.8% and 80.0%.

The second test of separability is conceptually similar but we model  $h(z_p, z_l, z_s)$  with a flexible Fourier series instead of a Taylor series expansion.

$$\gamma_i = \alpha X + \sum_{i \in \{P, L, S\}} b_i z_i + \sum_{i, j \in \{P, L, S\}} c_{ij} z_i z_j + \sum_{k \in K} \sum_{l=1}^L \{u_l \cos(l \cdot k'z) + v_l \sin(l \cdot k'z)\}. \quad (B3)$$

The flexible Fourier is essentially a polynomial expansion in quadratic and trigonometric terms. The basic benefit is the greater ability of the trigonometric expansion to approximate highly non-linear functions. Moreover, whereas a Taylor series expansion is a local approximation, the flexible Fourier fits the entire sample. There is an extensive econometric literature studying the properties of these estimators (Gallant, 1981, 1982; Geman and Hwang, 1982 and Gallant, 1985).

We have written the parametric part of the equation compactly as  $\alpha X$  to focus on the expansion. The trigonometric terms are indexed first by frequency: higher values of  $l$  refer to higher frequencies. Choosing a higher cutoff  $L$  delivers a more precise approximation by use of a greater range of frequencies, albeit at the cost of using degrees of freedom. Trigonometric terms are also distinguished by which variable forms the argument. In a multivariate context one can take a trigonometric expansion of a linear combination of the variables. We write these linear combinations as  $k'z$ , where  $z$  is the vector  $z_p, z_l, z_s$ . The  $k$  are known as *multi-indices* and are vectors whose elements are positive integers with values summing to a pre-specified value  $K^*$ . We denote  $K$  as the set of all such vectors for a given  $K^*$ . Thus, for  $K^* = 1$ ,  $K = \{[1, 0, 0], [0, 1, 0], [0, 0, 1]\}$  and the trigonometric expansion is separate in each dimension of  $z$ . For  $K^* = 2$ ,  $K = \{[1, 1, 0], [1, 0, 1], [0, 1, 1], [2, 0, 0], [0, 2, 0], [0, 0, 2]\}$  which allows for trigonometric expansions of linear combinations of each pair of  $z$ -s.<sup>15</sup>

Given a value of  $K^*$  and  $L$ , the parameter vector  $\beta = \{b, c, u, v\}$  can be estimated by ordinary least squares. In principle the total number of terms in the expansion is supposed to grow with sample size. Common practice for choosing  $K^*$  and  $L$  focuses on the “saturation ratio”, the ratio of the total number of terms in the expansion  $M$  to the number of observations  $N$ . In practice, saturation ratios between .25–.40 are typical of the applied literature (see Chalfant and Gallant, 1985; Pagan and Ullah, 1999). We have chosen  $K^* = 2$  to allow some interaction between shifters in the trigonometric expansion. We have chosen  $L = 1$  to preserve degrees of freedom and avoid over-fitting. As a result, we have 12 trigonometric terms, 9 quadratic, 3 linear, and 4 Solovian terms for a total of  $M = 28$  and a saturation ratio between .29 to .45 (depending on the sample) which is in line with previously published work.

We can obtain a restricted estimator  $\beta_r$  by restricting the coefficients on the terms involving interactions between different  $z$  variables to equal zero. Let  $e_u$  and  $e_r$  denote respectively the residuals from the restricted and unrestricted estimation. Hong and White (1995) have established that under the null hypothesis that the restrictions are valid, the HW statistic

$$HW = \frac{1}{(2M)^{1/2}} \left[ \frac{1}{N} \frac{\sum_{i=1}^N e_u^2 - \sum_{i=1}^N e_r^2}{\sum_{i=1}^N e_{ui}^2} - M \right] \rightarrow N(0, 1), \quad (B4)$$

where  $M$  is the number of terms in the Fourier expansion.<sup>16</sup> Since our data sets have a relatively small number of observations (between 62 and 96, depending on the specification), asymptotic standard errors may lead to erroneous inferences. We therefore construct bootstrapped test statistics with 100 replications.<sup>17</sup> The second column of Table 5 displays the median test statistic,  $p$ -values, and number of rejections of the null of separability using the Hong-White test.

In high dimensions and with limited information, one is likely to be able to fit many functional forms to the data, including separable and non-separable specifications. The null hypothesis of separability may be difficult to reject not because the world looks particularly separable, but rather because sparse data allows the world to be consistent with many views, including separability. In this light, the greater flexibility of the Fourier expansion likely results in reduced power to reject the null. And indeed, Table 5 shows this is the case: in each data sample, the flexible Fourier is less likely to reject than the Taylor polynomial. Nonetheless, each of these tests suggests that separability can be rejected in a large number of specifications. The Taylor polynomial tests reject separability for 65–80% of the specifications. Even the flexible Fourier manages to reject separability in 40–50% of the specifications. The rejection of separability in a large fraction of the specifications, even when using the rather forgiving flexible Fourier expansion, suggests that we must be wary of imposing this simplifying assumption.<sup>18</sup>

<sup>15</sup> These vectors are normalized to length one so that choice of  $K^*$  does not affect the frequency  $l$ .

<sup>16</sup> Earlier versions of this paper employed a third test of separability based on analysis of the residuals derived from direct estimation of the additively separable specification (8) by penalized spline estimation. This has been removed due to space constraints but is available from the authors on request.

<sup>17</sup> Li and Wang (1998) have found the bootstrap approximation to be superior to the asymptotic one for residual regression tests.

<sup>18</sup> To test the robustness to outliers, we have rerun these linearity and separability tests after dropping one country. All potential outliers have been investigated. A few countries do significantly reduce the number of specifications under which we can reject the restriction but in no case is the overall conclusion overturned.

## References

- Acemoglu, D., Johnson, S., Robinson, J.A., 2001. The colonial origins of comparative development: an empirical investigation. *American Economic Review* 91 (5), 1369–1401.
- Banerjee, A., Duflo, E., 2003. Inequality and growth: what can the data say? *Journal of Economic Growth* 8, 267–299.
- Barro, R., 1996. Democracy and growth. *Journal of Economic Growth* 1, 1–27.
- Barro, R., Sala-i-Martin, X., 2004. *Economic Growth*, second ed. MIT Press, Cambridge.
- Beck, T., Levine, R., Loayza, N., 2000. Finance and the sources of growth. *Journal of Financial Economics* 58 (1–2), 261–300.
- Bos, J.W.B., Economidou, C., Koetter, M., Kolari, J.W., 2010. Do all countries grow alike? *Journal of Development Economics* 91 (1), 113–127.
- Brock, W.A., Durlauf, S.N., 2001. *World Bank Economic Review* 15 (2).
- Canova, F., 2004. Testing for convergence clubs in income per capita: a predictive density approach. *International Economic Review* 45 (1), 49–77.
- Chalfant, J.A., Gallant, A.R., 1985. Estimating substitution elasticities with the Fourier cost function: some Monte-Carlo results. *Journal of Econometrics* 28, 205–222.
- Chang, R., Kaltani, L., Loayza, N., 2009. Openness can be good for growth: the role of policy complementarities. *Journal of Development Economics* 90, 33–49.
- Cuaresma, J.C., Doppelhofer, G., 2007. Nonlinearities in cross-country growth regressions: a Bayesian Averaging of Thresholds (BAT) approach. *Journal of Macroeconomics* 29, 541–554.
- Dejong, D., Ripoll, M., 2006. Tariffs and growth: an empirical exploration of contingent relationships. *The Review of Economics and Statistics* 88 (4), 625–640.
- Di Vaio, G., Enflo, K., 2011. Did globalization drive convergence? Identifying cross-country growth regimes in the long run. *European Economic Review* 55, 832–844.
- Dollar, D., Kraay, A., 2002. Growth is good for the poor. *Journal of Economic Growth* 7 (3), 195–225.
- Durlauf, S.N., Johnson, P.A., 1995. Multiple regimes and cross-country growth behavior. *Journal of Applied Econometrics* 10, 365–384.
- Durlauf, S., Kourtellos, A., Minkin, A., 2001. The local solow growth model. *European Economic Review* 15, 928–940.
- Durlauf, S., Johnson, P., Temple, J., 2005. Growth econometrics. In: Aghion, P., Durlauf, S. (Eds.), *Handbook of Economic Growth*. North Holland, Amsterdam.
- Easterly, W., Levine, R., 2003. Tropics, germs, and crops: how endowments influence economic development. *Journal of Monetary Economics* 50, 3–39.
- Frankel, J., Romer, R., 1999. Does trade cause growth? *American Economic Review* 89 (3), 379–399.
- Gallant, A.R., 1981. On the bias in flexible functional forms and an essentially unbiased form. *Journal of Econometrics* 15, 211–245.
- Gallant, A.R., 1982. Unbiased determination of production technologies. *Journal of Econometrics* 20, 285–323.
- Gallant, A.R., 1985. Unbiased determination of production technologies. *Journal of Econometrics* 20, 285–323.
- Geman, S., Hwang, C.H., 1982. Nonparametric maximum likelihood estimation by the method of sieves. *Annals of Statistics* 70, 401–414.
- Glaeser, E.L., La Porta, R., Lopez-de-Silanes, F., Shleifer, A., 2004. Do institutions cause growth. *Journal of Economic Growth* 9 (3), 271–303.
- Hall, R.E., Jones, C.L., 1999. Why do some countries produce so much more output per worker than others? *Quarterly Journal of Economics* 114, 83–116.
- Hansen, B., 2000. Sample splitting and threshold estimation. *Econometrica* 68 (3), 575–603.
- Härdle, W., Stoker, T., 1989. Investigating smooth multiple regression by the method of average derivatives. *Journal of the American Statistical Association* 84 (408), 986–995.
- Helpman, E., 2004. *The Mystery of Economic Growth*. Harvard University Press, Cambridge, US.
- Henderson, D.J., Papageorgiou, C., Parmeter, C.F., 2008. Are any growth theories linear? Why we should care about what the evidence tells us. Munich Personal RePEc Archive Working Paper No. 8767.
- Hong, Y., White, H., 1995. Consistent specification testing via nonparametric series regression. *Econometrica* 63, 1133–1160.
- Kalatzidakis, P., Mamuneas, T.P., Stengos, T., 2000. A non-linear sensitivity analysis of cross-country growth regressions. *The Canadian Journal of Economics* 33 (3), 604–617.
- Kaufman, D., Kraay, A., Zoido-Lobaton, P., 1999. *Governance Matters*. World Bank Research Working Paper No. 2196.
- Kourtellos, A., Tan, C.M., Zhang, X., 2007. Is the relationship between aid and economic growth nonlinear. *Journal of Macroeconomics* 29, 515–540.
- Li, Q., Wang, S., 1998. A simple consistent bootstrap test for a parametric regression function. *Journal of Econometrics* 87, 145–165.
- Liu, Z., Stengos, T., 1999. Non-linearities in cross-country growth regressions: a semiparametric approach. *Journal of Applied Econometrics* 14, 527–538.
- Lucas Jr., R.E., 2009. Trade and the diffusion of the industrial revolution. *American Economic Journal: Macroeconomics* 1 (1), 1–25.
- Mammem, E., 1991. Estimating a smooth-monotone regression function. *Annals of Statistics* 19, 724–740.
- Mankiw, N.G., Romer, D., Weil, D.N., 1992. A contribution to the empirics of economic growth. *The Quarterly Journal of Economics* 107 (2), 407–437.
- Minier, J., 2007a. Nonlinearities and robustness in growth regressions. *The American Economic Review* 97 (2), 388–392.
- Minier, J., 2007b. Institutions and parameter heterogeneity. *Journal of Macroeconomics* 29, 595–611.
- Owen, A., Videras, J., Davis, L., 2009. Do all countries follow the same growth process? *Journal of Economic Growth* 14, 265–286.
- Paap, R., Franses, P.H., van Dijk, D., 2005. Does Africa grow slower than Asia, Latin America, and the Middle East? Evidence from a new data-based classification method. *Journal of Development Economics* 77, 553–570.
- Pagan, A., Ullah, A., 1999. *Nonparametric Econometrics*. Cambridge University Press, Cambridge, UK.
- Rice, J.R., 1964. *The Approximation of Functions*. Addison-Wesley Pub. Co., Reading, Mass.
- Rilstone, P., 1991. Nonparametric hypothesis testing with parametric rates of convergence. *International Economic Review* 32 (1), 209–227.
- Rivlin, T., 1969. *An Introduction to the Approximation of Functions*. Blaisdell Pub. Co., Waltham, Mass.
- Sachs, J., 2005. *The End of Poverty: Economic Possibilities for Our Time*. The Penguin Press, New York.
- Sachs, J., Warner, A., 1995. *Natural Resource Abundance and Economic Growth*. NBER Working Paper 5398.
- Sala-i-Martin, X., Doppelhofer, G., Miller, R., 2004. Determinants of long-term growth: a Bayesian averaging of classical estimates (BACE) approach. *American Economic Review* (9), 812–835.
- Sirimaneetham, V., Temple, J., 2009. Macroeconomic stability and the distribution of growth rates. *World Bank Economic Review* 23 (3), 443–479.
- Tan, C.M., 2010. No one true path: uncovering the interplay between geography, institutions, and fractionalization in economic development. *Journal of Applied Econometrics* 25, 1100–1127.
- Temple, J., 2001. Growth effects of education and social capital in the OECD countries. *OECD Economic Studies* 33, 57–101.
- White, H., 1980. Using least squares to approximate unknown regression functions. *International Economic Review* 21 (1), 149–170.
- World Bank, 2005a. *World Development Indicators*. Electronic Database. The World Bank, Washington, DC.
- World Bank, 2005b. *Economic Growth in the 1990s: Learning from a Decade of Reform*. The World Bank, Washington.