Uncovering the (ir)relevance of finance on economic growth

Daniel J. Henderson†
Department of Economics
State University of New York at Binghamton

Chris Papageorgiou‡
Research Department
International Monetary Fund

Christopher F. Parmeter§
Department of Agricultural and Applied Economics
Virginia Polytechnic Institute and State University

November 17, 2008

Abstract

This paper takes a fresh look at the impact of financial development on economic growth. We allow for parameter heterogeneity and nonlinearities in a completely nonparametric fashion using recently developed generalized kernel methods. Our findings suggest that financial development (an appropriate index of it) impacts growth in a nonlinear fashion. We find that financial development has a larger and significant impact for countries with higher initial incomes, lower population growth rates and higher levels of investment. Further, this impact of financial development has grown over time.

Keywords: Economic Growth, Finance, Generalized Kernel Regression, Irrelevant Variables.

JEL Classification: C14, G0, O47, O50.

---

*We thank Patricio Venezuela for excellent research assistance. The views expressed in this study are the sole responsibility of the authors and should not be attributed to the International Monetary Fund, its Executive Board, or its management.

†Daniel J. Henderson, Department of Economics, State University of New York, Binghamton, NY 13902-6000, 607-777-4480, Fax: 607-777-2681, E-mail: djhender@binghamton.edu.

‡Chris Papageorgiou, Research Department, International Monetary Fund, Washington, DC 20431. Phone: 202-623-6107, Fax: 202-589-6107, E-mail: cpapageorgiou@imf.org.

§Christopher F. Parmeter, Department of Agricultural and Applied Economics, Virginia Polytechnic Institute and State University, Blacksburg, VA 24061-0401. Phone: 540-231-0770, Fax: 530-231-7417, E-mail: parms@vt.edu.
1 Introduction

Uncovering the impact of financial development on economic growth has been an interesting aspect of cross-country growth regressions over the last decade. Results suggest that improving the operating environment and mitigating the level of regulation can result in a significant growth performance. We add to this literature by addressing several empirical issues that have been levied against growth regressions in general to provide a robust perspective on how financial development affects economic growth.

The methods that we employ to produce robust conclusions about financial development stem from recent advances in the nonparametric regression literature. Contemporary methods have shown how to improve estimator efficiency in the face of continuous and discrete regressors. Individual and joint tests have been developed for significance testing of regressors. Tests of appropriate parametric specification are now available. Finally, we are able to handle the presence of irrelevant variables that are mistakenly included in an empirical analysis. All of these methods in combination provide the ability to assemble an empirical exercise that is robust to model misspecification, allows for parameter heterogeneity across covariates, and allows us to make valid inferential statements.\(^1\)

The remainder of this paper is organized as follows. Section 2 provides an overview of past research on the impact of financial development on economic growth and convergence. Section 3 gives an intuitive description of generalized nonparametric regression and the bandwidth selection method that removes irrelevant variables and can detect linearity. Section 4 briefly discusses the data used for this study while Section 5 contains our results. Extensions and conclusions are in Section 6. Technical aspects of the estimation and inference procedures used are contained in the appendix.

2 Finance as a growth theory

2.1 Findings from linear regressions

2.2 Accounting for nonlinearities

In this and the following subsection we use term \textit{nonlinear} to describe a variable that enters into a model in a nonlinear fashion independently of any other variables in the model. \textit{Parameter heterogeneity} is defined as a variable that, conditional on the level of the other variables in the model, has a partial effect that is constant, but can be different as the level changes.

\(^1\)Despite recent advances in nonparametric regression methods, one shortcoming of nonparametric methods that still remains is the ability to handle/account for endogeneity. Given that attempts to handle endogeneity nonparametrically are still in their infancy we eschew discussions of the likely endogeneity present in our study. While this is a serious issue that deserves attention, our methods can still provide valuable correlations between financial development and economic growth. This is an important distinction as, noted by Durlauf, Johnson, and Temple (2005), any growth study is likely to be flawed by regressor endogeneity and the instruments used to ‘correct’ for endogeneity are themselves likely flawed. See Ashley (2007) for methods to determine the impact of using a flawed instrument in an IV setting.
The first study to allow financial development to nonlinearly enter a growth regression is Rioja and Valev (2004a). They study nonlinearities of the impact of financial development by introducing dummies corresponding to the level of financial development. They use a dynamic panel estimator with a system of moment conditions and find that, indeed, the impact of financial development varies considerably across levels of income.

Ignoring nonlinearities, Rioja and Valev (2004a) find that the impact of private credit is 0.037 and significant at the 1% level. Interacting the financial development proxy with the ‘regions’ of income, however, suggest that the low region has an impact of -0.045 (but insignificant) while the middle region has an impact of 0.061 (and significant at the 1% level and the high region has an impact of 0.041 which is again significant at the 1% level. These findings suggest that increases in already high levels of financial development will result in increased growth, while increases in existing low levels of financial development will have no impact on growth. 2

Ketteni, Mamuneas, Savvides, and Stengos (2007) is perhaps the first paper to explore nonparametrically the impact of financial development on growth. They use an additive semiparametric partly linear specification that allows certain variables to enter into the model in a linear fashion and assumes others enter into the model in an additively separable, nonlinear fashion. They use the same data as Levine et al. (2000) and find that finance has a linear effect on growth. This suggests that one would expect to find the same impact of financial development on growth, no matter the level of financial development, which is at odds with Rioja and Valev (2004a). The reason for these stark differences may be because Rioja and Valev (2004a) do not allow for nonlinearities in initial income and human capital while Ketteni et al. (2007) do. This model misspecification could be the impetus for the findings in Rioja and Valev (2004a). Alternatively, because, Ketteni et al. (2007) do not allow for interactions between any of the variables in their model(s), their results may also be driven by model misspecification.

2.3 Parameter heterogeneity

Perhaps the first study to allow for a heterogeneous relationship between growth and financial development was Deidda and Fattouh (2002). Using the King and Levine (1993) dataset, Deidda and Fattouh (2002) use endogenous thresholding regression to allow for the impact of financial development to vary based on other variables in the model. Their results suggest that while financial development ignoring interactions has a positive and significant impact on growth, once thresholds are allowed only a positive and significant impact is found for those countries above the threshold (their threshold is defined via income per capita).

Rioja and Valev (2004b) study parameter heterogeneity by introducing dummies corresponding to the level of income. They use a dynamic panel estimator with a system of moment conditions and find that, indeed, the impact of financial development varies considerably across levels of income.

2Instead of arbitrarily splitting the financial development variable into three regions, Rioja and Valev (2004a) also try a total of 64 combinations of where the low and high regions should begin, using every 5th percentile above and below the median. Their qualitative results do not change much across these different pairings. Also, instead of using private credit they incorporate liquid liabilities (defined as currency plus demand and interest-bearing liabilities of banks and non-bank financial intermediaries) and an index of commercial bank assets divided by commercial plus central bank assets.
Ignoring interactions between income levels and financial development, Rioja and Valev (2004b) find that the impact of private credit is 0.010 and significant at the 1% level. Interacting the financial development proxy with the ‘regions’ of income, however, suggest that the low region has an impact of -0.001 (but insignificant) while the middle region has an impact of 0.012 and is significant at the 1% level and the high region has an impact of 0.020 and is also significant at the 1% level. These findings suggest that the effect of an increase in financial development for developed and nearly developed countries has a positive and significant impact on growth, while for underdeveloped countries, improvements in financial development have a negative and insignificant impact on economic growth.\(^3\)

Loayza and Ranciere (2006) use a pooled mean group (PMG) estimator to allow for heterogeneity in growth regressions that include financial development indicators. They estimate a model that includes both the current level of financial development as well as lagged changes in this development to pick up both long run and short term impacts to attempt to reconcile differences between findings in the literatures on empirical growth and banking crises. The estimation strategy that they employ restricts the coefficients on the long run variables (level of financial development) to be equal across countries, but allows for parameter heterogeneity in the short run variables (change in financial development over time). In their models they use the change in financial intermediation as well as two lags of this change as their short run financial variables. They test the restriction of poolability of the long run impacts and fail to reject a Hausman test. Thus, their findings suggest parameter heterogeneity in the short term, but parameter homogeneity in the long run impacts.

Rousseau and Wachtel (2006) use several regression estimators to examine parameter heterogeneity for financial development over time. They find that for their measure of liquid liabilities that the impact of financial development is positive and significant in the 1960-1989 period but positive and insignificant in the period 1990-2003. Decomposing these two time intervals down further, Rousseau and Wachtel (2006) find that financial development has a positive and significant (at the 5% level) impact on growth in the periods 1965-1969, 1970-1974, 1975-1979, and 1980-1984, a positive and insignificant (at the 5% level) impact on growth in the periods 1960-1964, 1985-1989, and 2000-2003. A negative and insignificant impact is found in the 1990-1994 and 1995-1999 periods. They draw almost identical conclusions using an alternative measure of financial development.\(^4\)

Schiavo and Vaona (2008) use the same data as Levine et al. (2000) and fit a linear model while accounting for the panel structure. Their interest lies with the poolability of the impact of a financial proxy in a growth regression. That is, they allow for the parameter on the financial development indicator to vary across countries, but staying constant over time and test whether these country specific parameters are equivalent across countries.

\(^3\)Rioja and Valev (2004b) also try using liquid liabilities (defined as currency plus demand and interest-bearing liabilities of banks and non-bank financial intermediaries) and a index of commercial bank assets divided by commercial plus central bank assets as alternative measures of financial development. Again, their findings suggest negative and insignificant impacts of financial development proxies for underdeveloped countries and positive and significant impacts for developed and nearly developed countries.

\(^4\)The results with private sector credit as a percent of GDP only provide evidence of a positive and significant impact on growth in the 1985-1989 period with all other periods being insignificant.
They use three different measures of financial development, each one appears individually in the growth regression
and they test for poolability of each within their whole sample, developed countries and developing countries.
They use four of the available tests for poolability which lead to 36 tests of poolability.

In 21 of 36 tests they find evidence of poolability. However, for the entire sample and the developed countries
samples Schiavo and Vaona (2008) find that one can assume a constant coefficient on finance in 8 out of 12 tests
for each sample. For developing countries they find evidence of a constant coefficient on financial development
in only 4 out of the twelve test. From their results it appears that the poolability found for the entire sample is
being driven by the apparent poolability in the developed country sample.

3 Nonparametric estimation

Nonparametric kernel regression is becoming an increasingly popular method of estimation in applied economic
milesus. The main perceived benefit is that it allows for consistent estimation when the underlying functional
form of the regression function and/or errors are unknown. While this is true, there are many other benefits
which may prove to be just as useful in our context. In this section we will discuss nonparametric regression, and
address the issue of bandwidth selection which can expose irrelevant covariates and detect linearity of others.

3.1 Nonparametric regression

Arguably the most popular regression model in the growth literature is the linear parametric model

\[ y_i = \alpha + \beta x_i + u_i, \quad i = 1, 2, \ldots, n, \]

where \( y_i \) is the dependent variable, \( x_i \) is a vector of \( q \) regressors, \( \alpha \) and \( \beta \) are unknown parameters to be estimated
and \( u_i \) is the random disturbance. Estimation of this model requires that all relevant regressors are included in
\( x_i \) and the functional form is correctly specified. However, when either of these two assumptions do not hold, the
estimates the model produces will most likely be inconsistent. While non-linear functional forms are possible in
a parametric framework, the data generating process still must be assumed \textit{a priori}.

Kernel methods have the ability to alleviate many of the restrictive assumptions necessary in the parametric
framework. Consider the nonparametric regression model

\[ y_i = m(x_i) + u_i, \quad i = 1, 2, \ldots, n, \]

where \( m(\cdot) \) is an unknown smooth function and the remaining variables are the same as before. Here, \( m(\cdot) \) is
interpreted as the conditional mean of \( y \) given \( x \). Note that in the (linear) parametric setting above it is implicitly
assumed that \( E(y_i|x_i) = \alpha + \beta x_i \). Further note that the linear model is a special case of our nonparametric
estimator and thus, if the true data generating process is indeed linear, then the nonparametric estimator will
give results consistent with that model.

One popular method for estimation of the unknown function is by local-constant least-squares regression. The local-constant least-squares (LCLS) estimator of the unknown function is given as

\[
\hat{m}(x) = \frac{\sum_{i=1}^{n} y_i \prod_{s=1}^{q} K \left( \frac{x_i - x_s}{h_s} \right)}{\sum_{i=1}^{n} \prod_{s=1}^{q} K \left( \frac{x_i - x_s}{h_s} \right)} = \sum_{i=1}^{n} y_i w_i(x),
\]

where \( \prod_{s=1}^{q} h_{-1}^{-1} K \left( (x_i - x_s) / h_s \right) \) is the product kernel and \( h_s \) is the smoothing parameter or bandwidth for a particular regressor \( x_s \) (see Pagan and Ullah 1999). The intuition behind this estimator is that it is simply a weighted average of \( y_i \). It is also known as a local average, given that the weights change depending upon the location of the regressors. We estimate the conditional mean function by locally averaging those values of the dependent variable which are ‘close’ in terms of the values taken on by the regressors. The amount of local information used to construct the average is controlled by the bandwidth.

The choice of the kernel function is somewhat arbitrary, but it must satisfy several properties. For relevant regressors, for large values of \( (x_i - x_s) / h_s \), \( K ((x_i - x_s) / h) \) should be small; less weight is assigned to data points which are far away from a particular value of \( x \). In other words, observations ‘close’ to \( x \) should be more helpful in prediction than observations further away from \( x \). In this context \( h \) determines how ‘close’ two observations are. A large \( h \) connotes that more points are considered close while a smaller \( h \) implies that fewer points are close to \( x \). Further, the kernel function must integrate to unity. Given these requirements, kernels are frequently chosen to be well-known density functions, with the standard normal being one of the most popular.

### 3.2 Bandwidth Selection

After choosing the kernel function, we can focus on the choice of bandwidths. Because it is believed that the choice of the continuous kernel function matters little in the estimation of the conditional mean (see Härdle 1990), selection of the bandwidths is considered to be the most salient factor when performing nonparametric estimation. As indicated above, the bandwidth controls the amount by which the data are smoothed. Large bandwidths will lead to large amounts of smoothing, resulting in low variance, but high bias. Small bandwidths, on the other hand, will lead to less smoothing, resulting in high variance, but low bias. This trade-off is well known in applied nonparametric econometrics, and the ‘solution’ is most often to resort to automated determination procedures to estimate the bandwidths. Although there exist many selection methods, we utilize the popular least-squares cross-validation (LSCV) criteria. This method has been studied extensively and additionally has known desirable properties regarding the smoothing out of irrelevant variables. Specifically, the bandwidths are chosen to minimize

\[
CV(h) = \frac{1}{n} \sum_{j=1}^{n} (y_j - \hat{m}_{-j}(x_j))^2,
\]
where \( \hat{m}_{-j}(x_j) \) is the commonly used leave-one-out estimator of \( m(\cdot) \). Obviously, as the sample size grows and/or the number of regressors increases, computation time increases dramatically. However, it is highly recommended that one use a bandwidth selection procedure as opposed to a rule of thumb selection, especially in the presence of discrete data as no rule of thumb selection criteria exists.

As an aside, we note that an even simpler bandwidth selection procedure, the ‘ocular’ method, is not appropriate once the number of covariates is larger than two. As the number of regressors exceeds two, visual methods to investigate the fit of the model are cumbersome and uninstructive. With a large dimension for the number of regressors, it is suggested that cross-validation techniques be used as opposed to either ocular or rule of thumb methods.

### 3.2.1 Irrelevant Regressors

The bandwidths, by affecting the degree of smoothing, are not just a means to an end; they also provide some indication of how the dependent variable is affected by the regressors. For instance, Hall, Li and Racine (2007) show that with LCLS, when the bandwidth on any regressor reaches its upper bound, the regressor is essentially ‘smoothed out’. In other words, the cross-validation procedure determines the bandwidths which predict the dependent variable the best, and thus chooses weights such that irrelevant variables have no impact on the prediction of the dependent variable.

An obvious question is whether or not in practice, these observations actually hit their upper bounds. For the continuous variables this is apparent. Computationally any cross-validation procedure cannot give bandwidths equal to their upper bound of infinity. Thus one must make a decision on how large a bandwidth must be until it is considered irrelevant. Hall, Li and Racine (2007) suggest that when the bandwidth exceeds a few standard deviations of the regressor, that the variable be deemed irrelevant. For discrete regressors, their upper bounds are quite obtainable. However, one may also want to deem regressors irrelevant when they are ‘close’ to their upper bounds. In practice it is preferable to use a formal test. This is the approach we take.

### 3.2.2 Detecting Linearity

In the local-linear least-squares (LLLS) framework discussed in Appendix A, as the bandwidth on a continuous regressor becomes large, the weight given to each observation becomes equal. In other words, as \( h \rightarrow \infty \), the implication is that the regressor enters linearly. The logic is the following: as the bandwidth becomes infinitely large, the local-linear regression fits a linear line using all the points in the neighborhood of \( x \). However, when the bandwidth is very large, then all the points are used and the fit is exactly the same for any \( x \) in the sample. Hence the estimate is linear.

This emphasizes the importance of obtaining a separate bandwidth for each regressor. If one regressor enters linearly we would expect that the cross-validation procedure should select large values of \( h \) for that regressor and
relatively small values for regressors that enter nonlinearly. In practice, Li and Racine (2004) suggest that any bandwidth which is more than two standard deviations of the regressor be deemed a variable that enters linearly.

However, linearity in a particular regressor does not mean that we should necessarily switch to a semiparametric model for the sake of efficiency. In the multi-variable case, it may be that there are important interactions between the ‘linear’ regressor and the remaining variables in the model, implying that the partial effect of the ‘linear’ regressor may still vary across $x$. Moreover, linearity should be more formally assessed, when feasible, using statistical tests. There is no formal statistical test for linearity of a specific regressor, but there are methods to test for specific parametric structure, for example linearity, which will be employed later.

4 Data

In this section we briefly discuss the data used in our estimation. Our data come from two sources. The typical growth variables motivated by the Solow (1956) model are from Durlauf, Kourtellos and Tan (2008) and include per capita real GDP, investment defined as the ratio of average investment to GDP, education defined as the average percentage of working age population (population between the age of 15 and 64) in secondary education, and the average growth rate of the working age population.\(^5\)

The financial development variables used are from the Financial Structure Dataset (revised version September 14, 2006) based on Beck, Demirgüç-Kunt and Levine (2000). Although we have considered several proxies for financial development our baseline results are based on the following four variables:

- **DBACBA** defined as “Deposit Money Bank Assets / (Deposit Money + Central) Bank Assets”.
- **DBAGDP** defined as “Deposit Money Bank Assets / GDP”.
- **PCRDMGD** defined as “Private Credit by Deposit Money Banks / GDP”.
- **BDGD** defined as “Bank Deposits / GDP”.

Merging data on the typical growth variables with data on these four financial development variables obtains an unbalanced 5-year non-overlapping panel dataset. The total number of country-year observations is 676 from 101 countries starting in 1960 and ending in 2000. The entire dataset used in the empirical estimation is available by the authors upon request.\(^6\)

\(^5\)The Durlauf, Kourtellos and Tan (2008) data set contains data for the traditional Solow model (initial income, investment rate, human capital, population growth) as well as variables that compose several of the contending growth theories being debated today: fractionalization, institutions, demographics, geography, religion, and macroeconomic policy. We thanks Steven Durlauf for sharing these data with us.

\(^6\)For a detailed description on these data see Beck, Demirgüç-Kunt and Levine (2000) and http://go.worldbank.org/X23UD9QUX0.
5 Results

Our discussion of the results can be broken into several categories which we highlight in turn. We first discuss baseline parametric results from both pooled cross-sections and traditional fixed effects panel estimation. We then discuss the implications of our bandwidth estimates for several specifications involving our financial market proxies. We then proceed to discuss the interpretation of our ‘parameter’ estimates from our nonparametric models.

5.1 Parametric Estimates

Table 1 presents OLS estimates of three linear growth models which form the basis of our following nonparametric analysis. We estimate a bevy of standard growth regressions. Models I, II, and V pool the data and include region and time fixed effects, Models III and VI employ least squares dummy variables to incorporate country specific fixed effects and models IV and VII use a three-way panel structure to account for country and time specific effects. Our primary financial variables, $\ln(DBACBA)$ is positive and significant in all four panel specifications, but when use in the pooled regression as a stand along for financial market development it is insignificant. Our panel results for initial income are to be expected as past studies have found that the ‘convergence coefficient’ is much larger when switching from cross-sectional methods to panel (Islam 1995, Casselli, Esquivel and Lefort 1997, Durlauf, Johnson and Temple 2005, sec. VI(ii)). As expected none of the additional proxies for financial market development are significant at even the 10% level suggesting that $\ln(DBACBA)$ is the best suited of our measures to attempt to uncover correlations.

It is interesting to point out that the perceived importance of human capital is heavily dependent on model choice. In our pooled cross-section results we have positive significance but when we employ panel methods we lose significance and sign. Previous studies have found similar impacts and have also attempted to explain why/when human capital may be expected to have positive and significant impacts on growth (Benhabib and Spiegel 1995). We note that our measure of human capital is evolving slowly over time and as such estimation methods which use fixed effects are likely to wipe-out almost all of the expected impact human capital is likely to have in a cross-sectional setting.

Before proceeding we note that the addition of the extra financial market proxies in the pooled setting (Model V) suggest that our main financial measure, $\ln(DBACBA)$ is highly positively significant. None of the other measures are significant however and the sample size has dropped. This is suggestive that the countries we dropped, all developing countries, have an influence on the results we found in Model II, similar to the findings of Rioja and Valev (2004a) that suggest at low-levels of development, financial market amelioration impedes, or at least does not improve, growth. The addition of these types of countries to increase our sample size for Models I-IV could harm our significance results.
5.2 Bandwidth Estimates

Prior to discussing the results, Table 2 presents both local constant and local linear cross-validated bandwidths for three distinct growth models. Following our discussion of the financial development variables in section 4 we consider a Solow-type model with no financial variables, a model with only one financial proxy as well as a model that incorporates all of our available financial proxies.

The first column in Table 2 gives the standard deviation for each continuous regressor along with the upper bound for the bandwidth for each categorical regressors. We choose to list the standard deviation instead of the upper bound of infinity for the continuous regressors as this upper bound is infeasible to obtain in practice. We instead use the rule of thumb analysis suggested by Li and Racine for a first-round analysis, but use more formal tests below. The second column of numbers are the bandwidths from the LCLS regression. We again note that when a bandwidth (for either type of regressor) reaches its upper bound in the LCLS framework, that variable is deemed irrelevant. Here we see that corresponding bandwidths for the schooling and population growth variables are very large. This suggests that both of these variables may be irrelevant in predicting output growth. This may be expected because we often find these variables to be insignificant in standard parametric growth regressions. At the same time, the bandwidth for OECD hits its upper bound. This is also expected as we have additionally included a variable for region.

The bandwidths for Model I which are obtained by LSCV of the LLLS regression are given as the third column in Table 2. Recall that for the LLLS regression, when a continuous regressor hits its upper bound, it is considered to enter the regression linearly (holding all else constant). Here we see that the bandwidth for the population growth variable is well beyond two times the standard deviation of the regressor and thus it is very possible that this variable enters in linearly. However, recall that this linearity is not synonymous with constant partial effects. There could be important interactions which would prohibit us from switching to a semiparametric model. Finally, we see similar results for the bandwidths for the categorical regressors. The OECD variable is deemed irrelevant while both the region and time variables are below their upper bounds.

The bandwidths in the single finance variable (Model II) case reveal three salient points. The results from the LSCV procedure of the LCLS estimator show that the bandwidth for ln(Pop.Growth) far exceeds two times its standard deviation. Since the relevance of the variable disappears as the bandwidths approach infinity, this suggests that population growth is irrelevant in predicting output growth. At the same time, the bandwidth on the OECD variable equals its upper bound of 0.5. This shows that this variable plays no role in the prediction of output growth. Again, this result should not be surprising here because we have a separate variable for region. When this variable is excluded from the analysis, the OECD variable is deemed relevant. The remaining bandwidths are much smaller than their respective upper bounds, implying that these variables are relevant in the model.

For the bandwidths selected via LSCV for the LLLS estimator, only population growth has a bandwidth
which is larger than two times its standard deviation. This suggests that this variable enters the model linearly. However, note that its bandwidth for the LCLS case showed that it was irrelevant in the prediction of output growth. Further, note that the remaining continuous variables all have bandwidths that are relatively small and thus a simple linear model would not be suggested here. This includes the financial development proxy. This is in contrast to the result found in Ketteni et al. (2007) who suggest that this variable enters in linearly. One possible reason the results differ is that their model did not allow for interactions with the financial proxy. This is confirmed by the Hsiao, Li and Racine (2007) test which rejects the parametric model at the 1% level (p-value = 0.001). In other words, assuming a linear parametric model would most likely lead to inconsistent estimates and incorrect inference.

The bandwidths for the model with four finance variables (Model III) differ somewhat in terms of their implications. First, the LCLS bandwidths for the three additional finance variables are shown to be far larger than two times their standard deviation. What this says is that each is irrelevant in the prediction of output growth. Another way to interpret this is that the first finance variable sufficiently explains the variation in the dependent variable and thus the other finance variables are not necessary for prediction purposes. A second difference is that now the bandwidth on \( \ln(\text{School}) \) is now larger than twice its standard deviation, and as the Solow model suggests, this variable may be irrelevant as well. Third, the LCLS bandwidth for OECD is now equal to 0.000. This implies that observations with different values of this covariate are given (essentially) no weight in the estimation; the kernel reduces to an indicator function. In other words, the estimation is equivalent to dividing the sample into OECD and non-OECD samples and performing separate regressions.

LSCV of the LLLS estimator now provides several bandwidths which are much larger than two times their standard deviations. Here, \( \ln(\text{School}) \), \( \ln(\text{Inv}) \), \( \ln(\text{Pop.Growth}) \) and \( \ln(\text{DBAGDP}) \) each have bandwidths which suggest linearity. However, recall that linearity is not synonymous with homogeneous effects of the covariates. The variables may enter linearly, but there may also exist important interactions which simple, linear in parameters models may not account for. Also, not all the variables enter linearly (specifically the two main variables of interest, \( \ln(Y_0) \) and \( \ln(DBACBA) \)) and thus a simple linear specification, even with interactions, may not be appropriate. Again, this is confirmed by the Hsiao, Li and Racine (2007) test which rejects the parametric model at the 1% level (p-value = 0.000). Finally, note that the OECD and time bandwidths each hit their upper bounds of 0.5 and 1.00 respectively. This suggests that in this model neither is important in the prediction of output growth.

In sum, examination of the bandwidths suggest that some variables enter the model nonlinearly and some variables enter the model linearly. However, this does not mean we should automatically switch to a semiparametric estimation procedure. Linearity is not synonymous with homogeneous effects of the covariates. Consequently, while the first assumption made in the majority of empirical analyses, linearity, receives the most attention,
heterogeneity may be just as problematic. At this point it is premature to determine which specification is more appropriate, but the model with four variables suggests that the additional three of them were irrelevant, whereas the in both models II and III, the bandwidths suggest that ln (DBACBA) is relevant. We now turn to the actual results, as well as more formal statistical tests.

5.3 Basic Parameter Estimates

Table 3 presents LLLS estimates of the partial effects for the continuous covariates across all three models. We present nonparametric estimates corresponding to the 25th, 50th, and 75th percentiles of the estimated parameter distributions (labelled Q1, Q2, and Q3) along with their corresponding bootstrap standard errors.

The results for the Solow model are listed first. The coefficients on the initial income variable are negative and significant for the first quartile and median. Perhaps more interesting than the significance is the large amount of variation in the parameter estimates. Note that the result at the first quartile is nearly six times smaller than the coefficient at the third quartile. Further, the results for the schooling variable are insignificant as expected and the results for the investment variable are generally positive and significant. For the last variable, population growth, the coefficients at the first quartile and median are negative and significant, while the coefficient at the upper quartile is negative, but insignificant. Finally, the $R^2 = 0.630$ is nearly three times as large as in the corresponding parametric model. This suggests that there may be a lot to gain in terms of prediction by employing the nonparametric model.

The signs for Model II are as expected. First, the coefficients on ln ($Y_0$) are negative but insignificant at the median and Q3. Second, the coefficients on ln (School) are small and insignificant for all three metrics. Next, the coefficients on ln (Inv) are all positive but again insignificant across metrics. Fourth, the coefficients on ln (Pop.Growth) are negative and insignificant. Finally, for ln (DBACBA), the results are positive but insignificant across measures.

Although we note that generally the signs are similar to those of parametric studies and the nonparametric model I, we find that many of our results are insignificant. The insignificant results are alarming at first glance. However, as we will soon witness, the significance of a particular variable depends upon the characteristics of the country. Here we point out that there is significant variation in the parameter estimates for a single variable. Figure 1 shows kernel density estimates for the parameter estimates across countries for each variable. For instance, the partial effect at the third quartile for the finance variable is nearly eight times larger than at the first quartile. In addition, we formally test that the finance variable is irrelevant in the estimation of the dependent variable by employing the Lavergne and Vuong (2002) test. We reject the null that the finance variable is irrelevant using either the bandwidths from the LCLS regression (p-value = 0.000) or the LLLS

---

9We choose to present the LLLS estimates because they have been shown to be estimated with less bias. In addition, estimation of the unknown function and its derivative (partial effects) are calculated simultaneously.

10A detailed description of the test is given in Appendix D.
regression (p-value = 0.000).

The results for Model III differ modestly. The coefficient estimates for the same five variables in Model II are qualitatively similar. A key difference is the change in significance of $\ln(Y_0)$ and $\ln(Inv)$ at the median. The additional three finance covariates offer little in terms of statistical or economic significance. Here we see that none of the coefficient estimates are significant. Again, this is not surprising given the findings of the previous sub-section. From this, there appears to be no reason to include the additional finance variables outside of $\ln(DBACBA)$. This is somewhat confirmed by the Lavergne and Vuong (2000) joint significance test of the three additional finance variables. We fail to reject the null that they are irrelevant using the LCLS regression bandwidths (p-value = 0.282), but reject the null using the LLLS regression bandwidths (p-value = 0.008).\textsuperscript{11} These findings are intuitive as the LCLS bandwidths can show irrelevance while the LLLS bandwidths can show linearity.

A traditional parametric analysis usually draws to a close at this point. Fortunately, in our case, we have a parameter estimate for each variable, for each country in each time period. Therefore we can examine further the coefficient estimates across suitably defined groups.

5.3.1 Parameter Estimates based on Sample Splitting

Even though we found that many of the results were insignificant for the single finance variable model (Model II), we choose to focus on this particular model for several reasons. First, the Lavergne and Vuong (2000) test was unable to reject the null that the finance variable $\ln(DBACBA)$ was irrelevant. Second, the same test failed to reject the null that the three additional finance variables were irrelevant in the LCLS case. This was further emphasized by their respective bandwidths in Table 2. Finally, given the curse of dimensionality in nonparametric regression, it is important to try to limit the number of continuous regressors to only those that are relevant in predicting the dependent variable.

Table 4 gives the nonparametric estimates corresponding to the median of the distributions for the estimated LLLS coefficients for every continuous variable in model II across splits on every variable’s median in the model. That is, we find the median coefficient on the logarithm of initial income, say, for all countries with population growth above the sample median and below the sample median. What this type of analysis will allow us to do is determine the possibility of nonlinearities and parameter heterogeneity at an aggregated level above point estimates.

\textsuperscript{11}Finally, we test the null that all four finance variables are irrelevant. We fail to reject the null that they are jointly irrelevant using the bandwidths from the LCLS regression (p-value = 0.124), but reject the null of joint significance with bandwidths from the LLLS regression (p-value = 0.000). The former result is peculiar because this null was rejected with a single variable. This lack of power is likely due to the relatively small sample. Our model with all four finance variables is nearly 20% smaller than our model with just $\ln(DBACBA)$.
5.3.2 Initial Income

The results across different covariates are interesting, but the main purpose of most growth studies is to examine the coefficient estimate on the initial income variable. In most studies, a single coefficient is obtained for this variable and its sign determines whether or not convergence exists across the sample. Here we obtain a separate coefficient for each country in each time period. Thus, we can examine the partial effect among pre-specified groups.

The results for the first column of numbers correspond to partial effect of the initial income variable. The partial effects are negative and significant at the median for those observations where the initial income was above the median, the level of investment was above the median and the finance variable was above the median. Not surprising is that when the level of schooling was above the median, the median coefficient on initial income was smaller at the median than when the level of schooling was below the median. However, this partial effect is insignificant. At the same time, when the level of population growth was above the median, the coefficient on the initial income variable was smaller on average than the coefficient on initial income for an observation with population growth below the median. Figure 2 shows the plots of each of these distributions of parameter estimates. In each case we can see that the means of the distributions differ as well as the variances. The differences are confirmed by the Li (1996) test\textsuperscript{12} for equality of two unknown distributions. In each case we reject the null that the distributions of the comparisons are equal. Specifically, each test has a p-value which is zero to three decimal places.

The results are also broken down for the initial income partial effect by geographical region in Table 5. Recall that the results from Table 3 showed that the coefficients were significant only at the first and second quartiles. Thus, it makes sense that some groups may have insignificant results. Here we see that the coefficients are significant at the median solely for OECD and North African/Middle Eastern countries. The other groups of countries have insignificant results for the median partial effect for their particular group.

5.3.3 The Conditioning Variables

Regardless of the split, Table 4 shows us that the median marginal effect of additional average schooling is insignificant. This is not surprising since our bandwidth estimates from this model suggest that ln(School) is ‘smoothed out’\. This is not the case when looking at the median marginal effects for either ln(Inv) and ln(PopGrowth). We see that positive and significant impacts related to capital investment appear for countries with above median initial income and years of schooling. Splits based according to median population growth and financial market access reveal no impact at the median for the logarithm of investment.

Focusing our attention on population growth impacts on the growth process we see negative and significant impacts almost across the board. For initial income, investment and population growth, the median effect of

\textsuperscript{12}A detailed description of the test is given in Appendix E.
population growth has a significant adverse affect on economic growth above and below the median with the above median estimates being smaller for initial income and investment. The opposite impact is found with respect to the population growth split. There countries with above median population growth witness a more than double impact on economic growth for an increase in the rate of increase in the population as opposed to those countries currently below the median level of population growth. A truly Malthusian effect.

For countries with above median levels of financial market access, we see that all three of our conditional variables from the Solow model are insignificant at the median. This is not suggestive, however, that there exist no interactions between the classical Solow conditioning variables and financial market access; the median marginal effect of $\ln(DBACBA)$ is positive and significant for above median levels of investment and below median levels of population growth. What these results do imply however is that across a broad range of splits, the median impact of the additional Solow controls have varied insights. This suggests that a one-size-fits all linear regression is inappropriate for modelling the growth process.

5.3.4 Financial Development Proxy

The results for the breakdown of the financial development variable are in the final column of Table 4. Here we see that those observations where the initial income is above the median have larger returns to the finance variable on average. The coefficient is significant at the median for the higher initial income group and insignificant for the lower initial income group. This is similar to the result found in Deidda and Fattouh (2002). The same magnitude differences occur for observations with above median levels of schooling and investment, but the median effect is insignificant for the median breakdown by schooling and marginally significant for the median result from the investment breakdown. The opposite result is true for those with above median population growth rates. There is a significant coefficient on the finance variable at the median for countries which have lower population growth rates. Finally, we looked at the median coefficient on the finance proxy for observations where $\ln(DBACBA)$ was above the median. We find this result to be larger than the median result for observations where the financial variable was below the median, but neither are significant. Thus, we unable to confirm or refute the finding in Rioja and Valev (2004a) that the former is larger.

These variations are visually seen in Figure 3. The differences are confirmed by the Li (1996) test. Again, each test has a p-value which is zero to three decimal places. The main departure from the previous case is that the (unreported) first quartile for the ‘lesser’ groups is negative in each case. For all scenarios the negative coefficients are insignificant. However, what this shows us is that in each of these comparisons, while the coefficients for the ‘greater’ groups are generally significantly positive at the median, over a quarter of the coefficients are (insignificantly) negative for the ‘lesser’ groups. In other words, for a large portion of the sample, the impact of financial development does not have a significant impact on output growth.

Table 5 breaks down the partial effects on the financial variable by geographical region. First, we see a
similar pattern in that now OECD countries have a larger effect of the finance variable as opposed to non-OECD countries. Further, the effect is significant for OECD economies at the median, but insignificant for non-OECD economies. This means that well developed economies are able to exploit financial markets for economic gain while lesser developed countries are not in a position to exploit their financial markets either because they are in a primitive stage of development or they lack the proper financial infrastructure. The same seems to hold true for each of the other groupings. Each are positive, but insignificant at their median values.

5.3.5 Time Variation in the Marginal Effects

Perhaps the most interesting results of the paper come from examining the time dimension of the estimated coefficients. In Table 6 we present the median partial effects for the each of the continuous variable as in Table 4, but here the groups correspond to the years under consideration. Most of the Solow variables do not seem to show a very strong trend. However, it is obvious that the median coefficient is increasing with time for the finance variable. The same holds true for the (unreported) upper and lower quartiles. Notice that the coefficient in 1960 is roughly four times smaller and insignificant at the median than the coefficient in 2000. This shows that the returns to financial markets have been increasing over time.

These large changes in the impact of financial development over time correspond to a similar impact that economic globalization has had over the same time period. Having access to a well developed financial market has also provided businesses with unabridged access to world markets over time. These connections can dramatically impact the economic outcomes of countries. Thus, while financial development appears to impact economic growth for specific types of countries, it also appears that having the ability to exploit a financial market is paramount. See Galor (2007) for the implications of how develop stages can dictate which economic variables are relevant for determining growth.

6 Conclusion

This paper has shed new light on the impact of financial development on economic growth using measures of financial development. We have assembled an array of results that suggests financial development impacts growth but is very specific to the particular stage a country is in along the development path. Countries deemed to be fully developed benefit the most from stable financial markets while underdeveloped countries are ill-equipped to benefit from financial markets.

These results were determined via contemporary nonparametric regression and inference results. These methods allow a researcher to avoid model misspecification, construct heterogenous parameter estimates for each observation and can invoke dimension reduction avenues easily. Taken together these methods represent the current apex of conducting a robust nonparametric analysis.

Future extensions of this work involve tracking continuing changes over time as more data (and countries)
become available for inclusion. The issue of endogeneity still remains and future studies should attempt to deal with the likely simultaneity that exists between financial development and economic growth. It would also be worth thinking about other types of groups to construct when considering the magnitude/significance of the estimated coefficients on both initial income and financial development. Applying these methods to other growth theories would also add to the discussion of model specification and parameter heterogeneity that has become a popular topic in the growth empirics literature.
References


A  Generalized Kernel Regression

The discussion above assumes that all the regressors are continuous. This assumption is not reasonable for most economic data sets. Previously, in the presence of ‘mixed’ data, kernel users had to resort to semiparametric methods. Often authors would assume that the categorical regressors entered the model linearly for ease of estimation. Fortunately, recent advances in nonparametric estimation allow for estimation of both continuous and discrete (order and unordered) variables. In a series of papers, Li and Racine (2004) and Racine and Li (2004) show that the unknown function can include both types of data. The nonparametric model in (2) is rewritten as

\[ y_i = m(x_i^c, x_i^u, x_i^o) + u_i, \quad i = 1, 2, \ldots, n, \]  

where \( x_i^c \) is a vector of continuous regressors (for example, initial income), \( x_i^u \) is a vector of unordered categorical regressors (for example, geographic region) and \( x_i^o \) is a vector of ordered categorical regressors (for example, year). All other variables are as previously described.

Estimation of this model by local-constant least-squares is quite similar. The main departure is in the construction of the product kernel. The product kernel, as the name suggests, is the product of the kernel functions for each variable. Here, one type of kernel function is used for continuous regressors, another is used for unordered discrete regressors and a third is used for ordered discrete regressors. The estimator is given as

\[
\hat{m}(x) = \frac{\sum_{i=1}^{n} y_i \prod_{s=1}^{q_c} K \left( \frac{x_{si}^c - x_{sh}^c}{h_c} \right) \prod_{s=1}^{q_u} l^u \left( x_{si}^u, x_s^u, \lambda_s^u \right) \prod_{s=1}^{q_o} l^o \left( x_{si}^o, x_s^o, \lambda_s^o \right)}{\sum_{i=1}^{n} \prod_{s=1}^{q_c} K \left( \frac{x_{si}^c - x_{sh}^c}{h_c} \right) \prod_{s=1}^{q_u} l^u \left( x_{si}^u, x_s^u, \lambda_s^u \right) \prod_{s=1}^{q_o} l^o \left( x_{si}^o, x_s^o, \lambda_s^o \right)} = \sum_{i=1}^{n} y_i w_i(x),
\]

where \( l^u \left( x_{si}^u, x_s^u, \lambda_s^u \right) \) is the kernel function for a particular unordered discrete regressor with bandwidth \( \lambda_s^u \) and again shows that our estimate is a weighted average of the \( y_i \)s. Similarly, \( l^o \left( x_{si}^o, x_s^o, \lambda_s^o \right) \) is the kernel function for a particular ordered discrete regressors with bandwidth \( \lambda_s^o \). For the continuous regressors, we choose the Gaussian kernel function

\[
K \left( \frac{x_{si}^c - x^c_s}{h_s} \right) = \frac{1}{\sqrt{2\pi}} \exp \left[ -\frac{1}{2} \left( \frac{x_{si}^c - x^c_s}{\lambda_s^c} \right)^2 \right];
\]

where the bandwidth ranges from zero to infinity. A variation of the Aitchinson and Aitken (1976) kernel function for unordered categorical regressors is given as

\[
l^u \left( x_{si}^u, x_s^u, \lambda_s^u \right) = \begin{cases} 1 - \lambda_s^u & \text{if } x_{si}^u = x_s^u \\ \frac{\lambda_s^u}{d_s - 1} & \text{otherwise} \end{cases};
\]

where the bandwidth is constrained to lie in the range \([0, (d_s - 1) / d_s]\), where \( d_s \) is the number of unique values the unordered variable will take. For example, for the case where the unordered variable is a traditional ‘dummy variable’, the upper bound will be 0.50. Finally, the Wang and Van Ryzin (1981) kernel function for ordered
categorical regressors is given by

\[
l^o(x^0_{oi}, x^0_{os}, \lambda^o_s) = \begin{cases} 
1 - \lambda^o_s & \text{if } x^o_{oi} = x^0_{os} \\
\frac{1 - \lambda^o_s}{2} (\lambda^o_s) |x^o_{oi} - x^0_{os}| & \text{otherwise}
\end{cases},
\]

where the bandwidth ranges from zero to unity.

Beyond the benefit of being able to incorporate categorical regressors, Li and Racine (2004) show that the rate of convergence of the conditional mean is only dependent on the number of continuous regressors. This is extremely important given the curse of dimensionality that is one of the criticisms levied against the use of nonparametric methods. In essence, the addition of discrete regressors need not require additional observations to achieve the same level of precision as the inclusion of additional continuous regressors would.

Another popular estimation procedure is local-linear least-squares estimator. This estimation procedure generally estimates the unknown function with more precision than the local-constant estimator. Additionally it estimates the marginal effects simultaneously. Again, consider the nonparametric regression model in (2). Taking a first-order Taylor expansion of (2) with respect to \(x^o_j\) yields

\[
y_i \approx m(x^o_j) + (x^o_i - x^o_j)\beta(x^o_j) + \varepsilon_i
\]

where \(\beta(x^o_j)\) is defined as the partial derivative of \(m(x^o_j)\) with respect to \(x^o\). For example, if \(y\) and \(x\) are both expressed in logarithmic form, then \(\beta(x_j)\) is interpreted as an elasticity. The estimator of \(\delta(x_j) \equiv [m(x_j) \quad \beta(x_j)]'\) is given by

\[
\hat{\delta}(x) = [X'K(x)X]^{-1} X'K(x)y
\]

where \(X = [1 \quad (x_i - x)]\) and \(K(x)\) is an \(n \times n\) diagonal matrix where the \(i^{th}\) diagonal element is \(K(h^{-1}(x_i - x))\). The intuition behind this estimator is that it fits a line through \(x\) based on the points ‘close’ to \(x\). This is repeated for each \(x\) and the slope and intercept of the lines do not have to be equal for different \(x\). Each of these lines are connected to produce the estimate of the unknown function.

**B The Inclusion of Irrelevant Variables**

In standard nonparametric regression, it is assumed that the bandwidth for a particular continuous regressor goes to zero as the sample size tends towards infinity. Here, when the variable is irrelevant, the cross-validated smoothing parameters converge in probability to the upper extremities of their respective ranges. In addition to improving prediction, this attenuates the curse of dimensionality by removing these variables from the analysis.

More formally, consider the estimator in (5), but say we add an additional \(p > 0\) irrelevant regressors for each
particular type of variable. The estimator of the conditional mean becomes

\[
\hat{m}(x) = \frac{\sum_{i=1}^{n} y_i \prod_{s=1}^{q_i} K \left( \frac{x_i - x_s}{h_s} \right) \prod_{s=q_i+1}^{q} K \left( \frac{x_i - x_s}{h_s} \right) \prod_{s=1}^{q} \hat{l}^{\nu}(x_i, x_s, \lambda^\nu_y)}{\sum_{i=1}^{n} \prod_{s=1}^{q_i} K \left( \frac{x_i - x_s}{h_s} \right) \prod_{s=q_i+1}^{q} K \left( \frac{x_i - x_s}{h_s} \right) \prod_{s=1}^{q} \hat{l}^{\nu}(x_i, x_s, \lambda^\nu_y)}
\]  

(10)

The idea is that the cross-validation procedure will recognize that each of these \(p\) irrelevant regressors for each variable type are in fact irrelevant and thus should not be used in the prediction of \(y\). For the continuous regressors, the upper bound is infinity. When this bandwidth takes its upper bound, the kernel function becomes \(K((x_i - x)/\infty) = K(0)\). For the unordered and ordered categorical kernels, the upper bounds are \((d_s - 1)/d_s\) and unity respectively. A closer examination of (8) and (9) show that when the bandwidth hits its upper bound, the weights given to observations equal to \(x_i^u\) and \(x_i^o\), respectively, are equal to the weights when the observations are different from \(x_i^u\) and \(x_i^o\), respectively. Therefore, when each of these irrelevant regressors are given their appropriate (upper bound) bandwidth, the kernel functions for the irrelevant regressors in (10) cancel out and we are left with (6), i.e. the second fraction in (10) is 1.

\section{Consistent Specification Testing}

To assess the correct estimation strategy, we utilize the Hsiao, Li and Racine (2007) specification test for mixed categorical and continuous data. The null hypothesis is that the parametric model \((f(x_i, \beta))\) is correctly specified \((H_0 : \Pr[E(y_i|x_i) = f(x_i, \beta)] = 1)\) against the alternative that it is not \((H_1 : \Pr[E(y_i|x_i) = f(x_i, \beta)] < 1)\). The test statistic is based on \(I \equiv E \left( E(u|x)^2 f(x) \right)\), where \(u = y - f(x, \beta)\). \(I\) is non-negative and equals zero if and only if the null is true. The resulting test statistic is

\[
T_n^2 = \frac{n \sqrt{h_1 h_2 \cdots h_q \hat{F}_n}}{\hat{\sigma}_n^2} \sim N(0, 1),
\]  

(11)

where

\[
\hat{F}_n = \frac{1}{n(n-1)} \sum_{i=1}^{n} \sum_{j=1, j \neq i}^{n} \hat{u}_i \hat{u}_j K_{\hat{h}, \hat{\lambda}^u, \hat{\lambda}^o},
\]

\[
\hat{\sigma}_n^2 = \frac{2h_1 h_2 \cdots h_q}{n(n-1)} \sum_{i=1}^{n} \sum_{j=1, j \neq i}^{n} \hat{u}_i^2 \hat{u}_j^2 K_{\hat{h}, \hat{\lambda}^u, \hat{\lambda}^o}^2
\]

with \(\hat{u}_i = y_i - f(x_i, \hat{\beta})\) the residual from the parametric model, \(K_{\hat{h}, \hat{\lambda}^u, \hat{\lambda}^o}\) is the product kernel discussed previously, \(q\) is the number of continuous regressors, and \(\hat{h}, \hat{\lambda}^u, \hat{\lambda}^o\) are the bandwidths obtained via LSCV. If the null is false,
$T^n$ diverges to positive infinity. Unfortunately, the asymptotic normal approximation performs poorly in finite samples and a bootstrap method is generally suggested for approximating the finite sample null distribution of the test statistic. Formally, the steps involved in computing the wild bootstrap statistic are as follows:

1. For $i = 1, 2, \ldots, n$, generate the two-point wild bootstrap error $u_i^* = \left( (1 - \sqrt{5}) / 2 \right) \hat{u}_i$, where $\hat{u}_i = y_i - f \left( x_i, \hat{\beta} \right)$ with probability $r = (1 - \sqrt{5}) / 2\sqrt{5}$ and $u_i^* = \left( (1 + \sqrt{5}) / 2 \right) \hat{u}_i$ with probability $1 - r$.

2. Create $y_i^* = f \left( x_i, \hat{\beta} \right) + u_i^* (i = 1, 2, \ldots, n)$. The resulting sample $\{x_i, y_i\}_{i=1}^n$ is called the bootstrap sample.

3. Obtain bootstrap residuals $\hat{u}_i^* = y_i^* - f \left( x_i, \hat{\beta}^* \right)$ ($i = 1, 2, \ldots, n$), where $\hat{\beta}^*$ is the parametric estimator of $\beta$ estimated from the bootstrap sample.

4. Use the bootstrap residuals to compute the test statistic $T_{n}^{a*} = n \left( h_1 h_2 \cdots h_q \right)^{1/2} \tilde{I}_{n}^{a*} / \hat{\sigma}_{n}^{a*}$, where $\tilde{I}_{n}^{a*}$ and $\hat{\sigma}_{n}^{a*}$ are the same as $I_{n}^{a*}$ and $\hat{\sigma}_{n}^{a*}$ except that $\hat{u}_i$ is replaced by $\hat{u}_i^*$.

5. Repeat steps (1-4) a large number ($B$) of times and then construct the empirical distribution of the $B$ bootstrap test statistics, $\{T_{n}^{a*}\}_{b=1}^B$. This bootstrap empirical distribution is used to approximate the null distribution of the test statistic $T_{n}^{a*}$. We reject $H_0$ if $T_{n}^{a*} > T_{n(aB)}^{a*}$, where $T_{n(aB)}^{a*}$ is the upper $\alpha$-percentile of $\{T_{n}^{a*}\}_{b=1}^B$.

Steps 2 through 4 heuristically ensure that conditional on the random sample, the bootstrap sample is generated by the null model. Conditional on $\{x_i, y_i\}_{i=1}^n$, $u_i^*$ has zero mean and the bootstrap statistic obtained in step 3 approximates the null distribution of the test statistic whether the null hypothesis is true or not.

## D Nonparametric Significance Testing

To determine whether or not a set of variables are jointly significant, we utilize the Lavergne and Vuong (2000) test modified by Racine, Hart and Li (2006) to allow for mixed categorical and continuous data. Consider a nonparametric regression model of the form

$$y_i = m (w_i, z_i) + u_i.$$ 

Here we discuss the case where all the variables in $z$ are continuous, but $w$ may contain mixed data. Let $w$ have dimension $r$ and $z$ have dimension $q - r$. The null hypothesis is that the conditional mean of $y$ does not depend on $z$.

$$H_0 : E (y|w, z) = E (y|w)$$

Define $u = y - E (y|w)$. Then $E (u|x) = 0$ under the null and we can construct a test statistic based on

$$E \left\{ u f_w (w) E [u f_w (w) | x] f (x) \right\}$$
where \( f_w(w) \) and \( f(x) \) are the pdf’s of \( w \) and \( x = (w, z) \), respectively. A feasible test statistic is given by

\[
\hat{T}_n^b = \frac{1}{n(n-1)} \sum_{i=1}^{n} \sum_{j=1, j \neq i}^{n} (y_i - \hat{y}_i) \hat{f}_w(w_i) (y_j - \hat{y}_j) \hat{f}_w(w_j) W(x_i, x, h, \lambda^o, \lambda^u) \tag{12}
\]

where \( W(x_i, x, h, \lambda^o, \lambda^u) = \prod_{s=1}^{q} K \left( \frac{x_i - x_s}{h_{s}} \right) \prod_{l=1}^{q} l^n(x_{s1}, x_{sl}, \lambda^o) \prod_{l=1}^{q} l^o(x_{s1}, x_{sl}, \lambda^o) \) is the product kernel mentioned in Section 3 and

\[
\hat{f}_w(w_i) = \frac{1}{n-1} \sum_{j=1, j \neq i}^{n} W(w_i, w, h_w, \lambda_w^o, \lambda_w^u)
\]

is the leave-one-out estimator of \( f_w(w_i) \). The leave one out estimator of \( E(y_i|w_i) \) is

\[
\hat{y}_i = \frac{1}{(n-1)f_w(w_i)} \sum_{j=1, j \neq i}^{n} y_j W(w_i, w, h_w, \lambda_w^o, \lambda_w^u)
\]

Under the null we have that

\[
T_n^b = (nh_1h_2\cdots h_q)^{1/2} \frac{\hat{T}_n^b}{\hat{\sigma}_n^b} \to N(0, 1)
\]

where

\[
\hat{\sigma}_n^{b2} = \frac{2h_1h_2\cdots h_q}{n^2} \sum_{i=1}^{n} \sum_{j=1, j \neq i}^{n} (y_i - \hat{y}_i)^2 \hat{f}_w(w_i)(y_j - \hat{y}_j)^2 \hat{f}_w(w_j) W(x_i, x, h, \lambda^o, \lambda^u)
\]

Again, the asymptotic distribution does not work well for finite samples. A bootstrap procedure is suggested instead. The bootstrap test statistic is obtained via the following steps:

1. For \( i = 1, 2, \ldots, n \), generate the two-point wild bootstrap error \( u_i^* = \left[ (1 - \sqrt{5}) / 2 \right] \hat{u}_i \), where \( \hat{u}_i = y_i - \hat{y}_i \) with probability \( r = (1 - \sqrt{5}) / 2\sqrt{5} \) and \( u_i^* = \left[ (1 + \sqrt{5}) / 2 \right] \hat{u}_i \) with probability \( 1 - r \).

2. Use the wild bootstrap error \( u_i^* \) to construct \( y_i^* = \hat{y}_i + u_i^* \), then obtain the kernel estimator of \( E^*(y_i^*|w_i) f_w(w_i) \) via

\[
\hat{y}_i^* \hat{f}_w(w_i) = \frac{1}{n-1} \sum_{j=1, j \neq i}^{n} y_j^* W(w_i, w, h_w, \lambda_w^o, \lambda_w^u)
\]

\[
\hat{y}_i^* = \frac{1}{(n-1)f_w(w_i)} \sum_{j=1, j \neq i}^{n} y_j^* W(w_i, w, h_w, \lambda_w^o, \lambda_w^u)
\]

The estimated density-weighted bootstrap residual is

\[
\hat{u}_i^* \hat{f}_w(w_i) = (y_i^* - \hat{y}_i^*) \hat{f}_w(w_i)
\]

\[
= y_i^* \hat{f}_w(w_i) - \hat{y}_i^* \hat{f}_w(w_i)
\]

3. Compute the standardized bootstrap test statistic \( T_n^{b*} \) where \( y^* \) and \( \hat{y}^* \) replace \( y \) and \( \hat{y} \) wherever they occur.

4. Repeat steps 1-3 a large number \( (B) \) of times and obtain the empirical distribution of the \( B \) bootstrap
test statistics. Let $T_n^{bs(\alpha)}$ denote the the $\alpha$-percentile of the bootstrap distribution. We will reject the null hypothesis at significance level $\alpha$ if $T_n^b > T_n^{bs(\alpha)}$.

### E Testing Equality of Two PDFs

To test whether two vectors of data $\{x_i\}_{i=1}^{n_1}$ and $\{z_i\}_{i=1}^{n_2}$ are drawn from the same distribution we employ the Li (1996) test. The Li (1996) test, which tests the null hypothesis $H_0 : f(x) = g(x)$ for all $x$, against the alternative $H_1 : f(x) \neq g(x)$ for some $x$, works with either independent or dependent data. The test statistic used to test for the difference between the two unknown distributions (which Fan and Ullah 1999 show goes asymptotically to the standard normal), predicated on the integrated square error metric on a space of density functions, $I(f, g) = \int_x (f(x) - g(x))^2 \, dx$, is

$$T_n = \frac{(n_1n_2h_1\cdots h_q)^{1/2} \hat{\tau}_n^c}{\hat{\sigma}_n^c} \sim N(0, 1), \quad (13)$$

where

$$\hat{\tau}_n^c = \frac{1}{n_1n_2} \sum_{i=1}^{n_1} \sum_{j=1, j \neq i}^{n_1} K_{h, ij}^x + \frac{1}{n_1n_2} \sum_{i=1}^{n_2} \sum_{j=1, j \neq i}^{n_2} K_{h, ij}^z - \frac{2}{n_1n_2} \sum_{i=1}^{n_1} \sum_{j=1, j \neq i}^{n_2} K_{h, ij}^{xz},$$

and

$$\hat{\sigma}_n^2 = \frac{h_1h_2\cdots h_q}{n_1n_2} \left\{ \sum_{i=1}^{n_1} \sum_{j=1, j \neq i}^{n_1} \left[ K_{h, ij}^x / n_1/n_2 \right]^2 + \sum_{i=1}^{n_2} \sum_{j=1, j \neq i}^{n_2} \left[ K_{h, ij}^z / n_2/n_1 \right]^2 + 2 \sum_{i=1}^{n_1} \sum_{j=1, j \neq i}^{n_2} \left[ K_{h, ij}^{xz} \right]^2 \right\},$$

where $K_{h, ij}^x = \prod_{s=1}^{q} h_s^{-1} K((x_{is} - x_{js})/h_s)$, $K_{h, ij}^z = \prod_{s=1}^{q} h_s^{-1} K((z_{is} - z_{js})/h_s)$, and $K_{h, ij}^{xz} = \prod_{s=1}^{q} h_s^{-1} K((x_{is} - z_{js})/h_s)$.

Again, if the null is false, $T^c$ diverges to positive infinity. Unfortunately, the asymptotic normal approximation performs poorly in finite samples and a bootstrap method is generally suggested for approximating the finite sample null distribution of the test statistic. Formally, this is accomplished by randomly sampling with replacement from the pooled data. The steps are as follows:

1. Randomly draw $n_1 + n_2$ observations with replacement from the pooled data set. Call the first $n_1$ observations $\{x^*_i\}_{i=1}^{n_1}$ and the remaining $n_2$ observations $\{z^*_i\}_{i=1}^{n_2}$.

2. Use the bootstrap data to compute the test statistic $T_n^{c*} = (n_1n_2h_1h_2\cdots h_q)^{1/2} \hat{\tau}_n^{c*} / \hat{\sigma}_n^{c*}$, where $\hat{\tau}_n^{c*}$ and $\hat{\sigma}_n^{c*}$ are the same as $\hat{\tau}_n^c$ and $\hat{\sigma}_n^c$ except that $\{x_i\}_{i=1}^{n_1}$ and $\{z_i\}_{i=1}^{n_2}$ are replaced by $\{x^*_i\}_{i=1}^{n_1}$ and $\{z^*_i\}_{i=1}^{n_2}$, respectively.

3. Repeat steps (1-2) a large number ($B$) of times and then construct the empirical distribution of the $B$ bootstrap test statistics, $T_n^{c*}_{b=1}$. This bootstrap empirical distribution is used to approximate the null
distribution of the test statistic $T_n^c$. We reject $H_0$ if $T_n^c > T_{n(\alpha B)}^{c*}$, where $T_{n(\alpha B)}^{c*}$ is the upper $\alpha$-percentile of $\{T_n^c\}_{b=1}^B$. 
Table 1: OLS and Panel estimates for several parametric specifications

<table>
<thead>
<tr>
<th>Variable</th>
<th>Model I</th>
<th>Model II</th>
<th>Model III</th>
<th>Model IV</th>
<th>Model V</th>
<th>Model VI</th>
<th>Model VII</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\ln(Y_0)$</td>
<td>-0.043</td>
<td>-0.048</td>
<td>-0.183</td>
<td>-0.143</td>
<td>-0.051</td>
<td>-0.188</td>
<td>-0.138</td>
</tr>
<tr>
<td></td>
<td>(0.000)</td>
<td>(0.000)</td>
<td>(0.016)</td>
<td>(0.017)</td>
<td>(0.000)</td>
<td>(0.023)</td>
<td>(0.021)</td>
</tr>
<tr>
<td>$\ln(\text{School})$</td>
<td>0.033</td>
<td>0.035</td>
<td>0.020</td>
<td>-0.001</td>
<td>0.036</td>
<td>0.000</td>
<td>-0.013</td>
</tr>
<tr>
<td></td>
<td>(0.003)</td>
<td>(0.001)</td>
<td>(0.016)</td>
<td>(0.019)</td>
<td>(0.009)</td>
<td>(0.024)</td>
<td>(0.026)</td>
</tr>
<tr>
<td>$\ln(\text{Inv})$</td>
<td>0.044</td>
<td>0.400</td>
<td>0.058</td>
<td>0.125</td>
<td>0.035</td>
<td>0.054</td>
<td>0.131</td>
</tr>
<tr>
<td></td>
<td>(0.000)</td>
<td>(0.001)</td>
<td>(0.018)</td>
<td>(0.018)</td>
<td>(0.009)</td>
<td>(0.024)</td>
<td>(0.024)</td>
</tr>
<tr>
<td>$\ln(\text{Pop Growth})$</td>
<td>-0.074</td>
<td>-0.076</td>
<td>-0.007</td>
<td>-0.006</td>
<td>-0.133</td>
<td>-0.076</td>
<td>-0.106</td>
</tr>
<tr>
<td></td>
<td>(0.000)</td>
<td>(0.001)</td>
<td>(0.018)</td>
<td>(0.018)</td>
<td>(0.013)</td>
<td>(0.024)</td>
<td>(0.024)</td>
</tr>
<tr>
<td>OECD</td>
<td>-0.010</td>
<td>-0.009</td>
<td>$-$</td>
<td>$-$</td>
<td>$-$</td>
<td>$-$</td>
<td>$-$</td>
</tr>
<tr>
<td></td>
<td>(0.020)</td>
<td>(0.020)</td>
<td>$-$</td>
<td>$-$</td>
<td>(0.027)</td>
<td>$-$</td>
<td>$-$</td>
</tr>
<tr>
<td>$\ln(DBACBA)$</td>
<td>$-$</td>
<td>0.024</td>
<td>0.075</td>
<td>0.039</td>
<td>0.066</td>
<td>0.135</td>
<td>0.105</td>
</tr>
<tr>
<td></td>
<td>(0.140)</td>
<td>(0.022)</td>
<td>(0.022)</td>
<td>(0.002)</td>
<td>(0.034)</td>
<td>(0.034)</td>
<td>(0.034)</td>
</tr>
<tr>
<td>$\ln(DBAGDP)$</td>
<td>$-$</td>
<td>$-$</td>
<td>$-$</td>
<td>$-$</td>
<td>$-$</td>
<td>$-$</td>
<td>$-$</td>
</tr>
<tr>
<td></td>
<td>(0.452)</td>
<td>(0.039)</td>
<td>(0.040)</td>
<td>(0.039)</td>
<td>(0.040)</td>
<td>(0.040)</td>
<td>(0.040)</td>
</tr>
<tr>
<td>$\ln(\text{PCRDBGDP})$</td>
<td>-0.003</td>
<td>0.015</td>
<td>0.005</td>
<td>(0.882)</td>
<td>(0.029)</td>
<td>(0.029)</td>
<td>(0.029)</td>
</tr>
<tr>
<td>$\ln(\text{BDGDP})$</td>
<td>0.033</td>
<td>0.048</td>
<td>0.021</td>
<td>(0.088)</td>
<td>(0.028)</td>
<td>(0.028)</td>
<td>(0.028)</td>
</tr>
<tr>
<td>Region/Time</td>
<td>Yes</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td>Individual</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
<td>No</td>
</tr>
<tr>
<td>Individual/Time</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
</tr>
<tr>
<td>$R^2$</td>
<td>0.258</td>
<td>0.260</td>
<td>0.251</td>
<td>0.170</td>
<td>0.206</td>
<td>0.247</td>
<td>0.171</td>
</tr>
<tr>
<td>$n$</td>
<td>676</td>
<td>676</td>
<td>676</td>
<td>676</td>
<td>528</td>
<td>528</td>
<td>528</td>
</tr>
</tbody>
</table>
Table 2: Bandwidths for various growth models. In place of the standard errors for OECD, Region, and Time we have given the upper bound that the bandwidth can take in the cross-validation procedure.

<table>
<thead>
<tr>
<th>Variable</th>
<th>Model I</th>
<th></th>
<th></th>
<th>Model II</th>
<th></th>
<th></th>
<th>Model III</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>St Err</td>
<td>LCLS</td>
<td>LLLS</td>
<td>St Err</td>
<td>LCLS</td>
<td>LLLS</td>
<td>St Err</td>
<td>LCLS</td>
</tr>
<tr>
<td>ln(Y0)</td>
<td>1.089</td>
<td>0.188</td>
<td>0.821</td>
<td>1.089</td>
<td>0.223</td>
<td>0.758</td>
<td>1.053</td>
<td>0.612</td>
</tr>
<tr>
<td>ln(School)</td>
<td>0.755</td>
<td>1.7629220</td>
<td>0.249</td>
<td>0.755</td>
<td>0.337</td>
<td>0.220</td>
<td>0.681</td>
<td>12.93</td>
</tr>
<tr>
<td>ln(Inv)</td>
<td>0.610</td>
<td>0.482</td>
<td>0.624</td>
<td>0.610</td>
<td>0.351</td>
<td>0.573</td>
<td>0.579</td>
<td>0.291</td>
</tr>
<tr>
<td>ln(Pop Growth)</td>
<td>0.176</td>
<td>298507</td>
<td>19112</td>
<td>0.176</td>
<td>2337505</td>
<td>99781</td>
<td>0.169</td>
<td>42826</td>
</tr>
<tr>
<td>ln(DBACBA)</td>
<td>0.779</td>
<td>2249826</td>
<td>30455</td>
<td>0.779</td>
<td>2249826</td>
<td>30455</td>
<td>0.779</td>
<td>30455</td>
</tr>
<tr>
<td>ln(PCRDBGDP)</td>
<td>0.843</td>
<td>10651825</td>
<td>0.511</td>
<td>0.843</td>
<td>10651825</td>
<td>0.511</td>
<td>0.843</td>
<td>10651825</td>
</tr>
<tr>
<td>ln(BDGD)</td>
<td>0.691</td>
<td>1996485</td>
<td>0.212</td>
<td>0.691</td>
<td>1996485</td>
<td>0.212</td>
<td>0.691</td>
<td>1996485</td>
</tr>
<tr>
<td>OECD</td>
<td>0.500</td>
<td>0.500</td>
<td>0.500</td>
<td>0.500</td>
<td>0.500</td>
<td>0.500</td>
<td>0.500</td>
<td>0.000</td>
</tr>
<tr>
<td>Region</td>
<td>0.900</td>
<td>0.100</td>
<td>0.698</td>
<td>0.900</td>
<td>0.487</td>
<td>0.727</td>
<td>0.900</td>
<td>0.678</td>
</tr>
<tr>
<td>Time</td>
<td>1.00</td>
<td>0.931</td>
<td>0.913</td>
<td>1.000</td>
<td>0.930</td>
<td>0.730</td>
<td>1.000</td>
<td>0.255</td>
</tr>
<tr>
<td>n</td>
<td>676</td>
<td>676</td>
<td>676</td>
<td>676</td>
<td>676</td>
<td>676</td>
<td>528</td>
<td>528</td>
</tr>
</tbody>
</table>
Table 3: LLLS Quartile coefficient estimates from the three growth models. Bootstrap standard errors are in parentheses below each estimate.

<table>
<thead>
<tr>
<th>Variable</th>
<th>Model I</th>
<th>Model II</th>
<th>Model III</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Q1</td>
<td>Q2</td>
<td>Q3</td>
</tr>
<tr>
<td>ln(Y0)</td>
<td>-0.086</td>
<td>-0.043</td>
<td>-0.014</td>
</tr>
<tr>
<td></td>
<td>(0.014)</td>
<td>(0.018)</td>
<td>(0.022)</td>
</tr>
<tr>
<td>ln(School)</td>
<td>-0.042</td>
<td>0.007</td>
<td>0.054</td>
</tr>
<tr>
<td></td>
<td>(0.038)</td>
<td>(0.034)</td>
<td>(0.039)</td>
</tr>
<tr>
<td>ln(Inv)</td>
<td>0.056</td>
<td>0.082</td>
<td>0.116</td>
</tr>
<tr>
<td></td>
<td>(0.030)</td>
<td>(0.025)</td>
<td>(0.021)</td>
</tr>
<tr>
<td>ln(Pop Growth)</td>
<td>-0.205</td>
<td>-0.108</td>
<td>-0.017</td>
</tr>
<tr>
<td></td>
<td>(0.115)</td>
<td>(0.055)</td>
<td>(0.044)</td>
</tr>
<tr>
<td>ln(DBACBA)</td>
<td>0.016</td>
<td>0.069</td>
<td>0.122</td>
</tr>
<tr>
<td></td>
<td>(0.102)</td>
<td>(0.068)</td>
<td>(0.096)</td>
</tr>
<tr>
<td>ln(DBAGDP)</td>
<td>-0.125</td>
<td>-0.047</td>
<td>0.009</td>
</tr>
<tr>
<td></td>
<td>(0.100)</td>
<td>(0.075)</td>
<td>(0.051)</td>
</tr>
<tr>
<td>ln(PCRDBGDP)</td>
<td>-0.028</td>
<td>0.014</td>
<td>0.063</td>
</tr>
<tr>
<td></td>
<td>(0.060)</td>
<td>(0.078)</td>
<td>(0.051)</td>
</tr>
<tr>
<td>ln(BDGDP)</td>
<td>-0.024</td>
<td>0.043</td>
<td>0.111</td>
</tr>
<tr>
<td></td>
<td>(0.154)</td>
<td>(0.105)</td>
<td>(0.070)</td>
</tr>
<tr>
<td>$R^2$</td>
<td>0.630</td>
<td>0.772</td>
<td>0.839</td>
</tr>
</tbody>
</table>
Table 4: Median coefficient of LLLS estimates from Model II for each continuous regressor across specific groups. Bootstrap standard errors are in parentheses below each estimate.

<table>
<thead>
<tr>
<th>Split/Variable</th>
<th>$ln(Y_0)$</th>
<th>$ln(School)$</th>
<th>$ln(Inv)$</th>
<th>$ln(Pop\ Growth)$</th>
<th>$ln(DBACBA)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>&gt; median($ln(Y_0)$)</td>
<td>-0.084</td>
<td>0.046</td>
<td>0.070</td>
<td>-0.093</td>
<td>0.080</td>
</tr>
<tr>
<td></td>
<td>(0.071)</td>
<td>(0.123)</td>
<td>(0.027)</td>
<td>(0.025)</td>
<td>(0.034)</td>
</tr>
<tr>
<td>&lt; median($ln(Y_0)$)</td>
<td>-0.035</td>
<td>-0.007</td>
<td>0.054</td>
<td>-0.220</td>
<td>0.046</td>
</tr>
<tr>
<td></td>
<td>(0.037)</td>
<td>(0.088)</td>
<td>(0.039)</td>
<td>(0.045)</td>
<td>(0.031)</td>
</tr>
<tr>
<td>&gt; median($ln(School)$)</td>
<td>-0.087</td>
<td>0.048</td>
<td>0.075</td>
<td>-0.101</td>
<td>0.079</td>
</tr>
<tr>
<td></td>
<td>(0.062)</td>
<td>(0.069)</td>
<td>(0.025)</td>
<td>(0.071)</td>
<td>(0.077)</td>
</tr>
<tr>
<td>&lt; median($ln(School)$)</td>
<td>-0.035</td>
<td>-0.010</td>
<td>0.049</td>
<td>-0.191</td>
<td>0.042</td>
</tr>
<tr>
<td></td>
<td>(0.035)</td>
<td>(0.167)</td>
<td>(0.080)</td>
<td>(0.045)</td>
<td>(0.047)</td>
</tr>
<tr>
<td>&gt; median($ln(Inv)$)</td>
<td>-0.089</td>
<td>0.049</td>
<td>0.090</td>
<td>-0.083</td>
<td>0.075</td>
</tr>
<tr>
<td></td>
<td>(0.031)</td>
<td>(0.229)</td>
<td>(0.031)</td>
<td>(0.046)</td>
<td>(0.042)</td>
</tr>
<tr>
<td>&lt; median($ln(Inv)$)</td>
<td>-0.034</td>
<td>-0.018</td>
<td>0.042</td>
<td>-0.224</td>
<td>0.058</td>
</tr>
<tr>
<td></td>
<td>(0.031)</td>
<td>(0.229)</td>
<td>(0.031)</td>
<td>(0.046)</td>
<td>(0.042)</td>
</tr>
<tr>
<td>&gt; median($ln(Pop\ Growth)$)</td>
<td>-0.036</td>
<td>-0.010</td>
<td>0.048</td>
<td>-0.213</td>
<td>0.046</td>
</tr>
<tr>
<td></td>
<td>(0.040)</td>
<td>(0.052)</td>
<td>(0.036)</td>
<td>(0.035)</td>
<td>(0.037)</td>
</tr>
<tr>
<td>&lt; median($ln(Pop\ Growth)$)</td>
<td>-0.085</td>
<td>0.049</td>
<td>0.075</td>
<td>-0.100</td>
<td>0.083</td>
</tr>
<tr>
<td></td>
<td>(0.152)</td>
<td>(0.092)</td>
<td>(0.047)</td>
<td>(0.039)</td>
<td>(0.029)</td>
</tr>
<tr>
<td>&gt; median($ln(DBACBA)$)</td>
<td>-0.080</td>
<td>0.041</td>
<td>0.069</td>
<td>-0.092</td>
<td>0.088</td>
</tr>
<tr>
<td></td>
<td>(0.027)</td>
<td>(0.053)</td>
<td>(0.056)</td>
<td>(0.275)</td>
<td>(0.087)</td>
</tr>
<tr>
<td>&lt; median($ln(DBACBA)$)</td>
<td>-0.041</td>
<td>0.002</td>
<td>0.055</td>
<td>-0.197</td>
<td>0.046</td>
</tr>
<tr>
<td></td>
<td>(0.034)</td>
<td>(0.078)</td>
<td>(0.041)</td>
<td>(0.093)</td>
<td>(0.091)</td>
</tr>
</tbody>
</table>
Table 5: Median coefficient of LLLS estimates from Model II for each continuous regressor for specific groups of countries. Bootstrap standard errors are in parentheses below each estimate.

<table>
<thead>
<tr>
<th>Classification</th>
<th>$ln(Y_0)$</th>
<th>$ln(\text{School})$</th>
<th>$ln(\text{Ine})$</th>
<th>$ln(\text{Pop Growth})$</th>
<th>$ln(DBACBA)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>OECD</td>
<td>-0.100</td>
<td>0.055</td>
<td>0.075</td>
<td>-0.056</td>
<td>0.093</td>
</tr>
<tr>
<td></td>
<td>(0.053)</td>
<td>(0.191)</td>
<td>(0.057)</td>
<td>(0.032)</td>
<td>(0.025)</td>
</tr>
<tr>
<td>Non-OECD</td>
<td>-0.041</td>
<td>0.000</td>
<td>0.055</td>
<td>-0.186</td>
<td>0.053</td>
</tr>
<tr>
<td></td>
<td>(0.070)</td>
<td>(0.073)</td>
<td>(0.049)</td>
<td>(0.054)</td>
<td>(0.454)</td>
</tr>
<tr>
<td>Sub-Saharan Africa</td>
<td>-0.036</td>
<td>0.016</td>
<td>0.049</td>
<td>-0.149</td>
<td>0.042</td>
</tr>
<tr>
<td></td>
<td>(0.052)</td>
<td>(0.157)</td>
<td>(0.042)</td>
<td>(0.078)</td>
<td>(0.063)</td>
</tr>
<tr>
<td>North Africa/Middle East</td>
<td>-0.127</td>
<td>0.077</td>
<td>0.157</td>
<td>-0.096</td>
<td>0.110</td>
</tr>
<tr>
<td></td>
<td>(0.045)</td>
<td>(0.082)</td>
<td>(0.043)</td>
<td>(0.115)</td>
<td>(0.129)</td>
</tr>
<tr>
<td>Asia</td>
<td>-0.043</td>
<td>-0.019</td>
<td>0.048</td>
<td>-0.247</td>
<td>0.052</td>
</tr>
<tr>
<td></td>
<td>(0.064)</td>
<td>(0.143)</td>
<td>(0.049)</td>
<td>(0.152)</td>
<td>(0.143)</td>
</tr>
<tr>
<td>Latin America</td>
<td>-0.023</td>
<td>-0.010</td>
<td>0.054</td>
<td>-0.360</td>
<td>0.019</td>
</tr>
<tr>
<td></td>
<td>(0.044)</td>
<td>(0.125)</td>
<td>(0.053)</td>
<td>(0.128)</td>
<td>(0.075)</td>
</tr>
</tbody>
</table>
Table 6: Median coefficient of LLLS estimates from Model II for each continuous regressor for each time period. Bootstrap standard errors are in parentheses below each estimate.

<table>
<thead>
<tr>
<th>Time Period</th>
<th>$ln(Y_0)$</th>
<th>$ln(School)$</th>
<th>$ln(Inv)$</th>
<th>$ln(Pop Growth)$</th>
<th>$ln(DBACBA)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1960</td>
<td>-0.062</td>
<td>0.023</td>
<td>0.093</td>
<td>-0.077</td>
<td>0.030</td>
</tr>
<tr>
<td></td>
<td>(0.069)</td>
<td>(0.172)</td>
<td>(0.054)</td>
<td>(0.064)</td>
<td>(0.068)</td>
</tr>
<tr>
<td>1965</td>
<td>-0.052</td>
<td>0.035</td>
<td>0.091</td>
<td>-0.111</td>
<td>0.025</td>
</tr>
<tr>
<td></td>
<td>(0.034)</td>
<td>(0.059)</td>
<td>(0.035)</td>
<td>(0.044)</td>
<td>(0.059)</td>
</tr>
<tr>
<td>1970</td>
<td>-0.058</td>
<td>0.017</td>
<td>0.085</td>
<td>-0.061</td>
<td>0.037</td>
</tr>
<tr>
<td></td>
<td>(0.040)</td>
<td>(0.092)</td>
<td>(0.040)</td>
<td>(0.031)</td>
<td>(0.163)</td>
</tr>
<tr>
<td>1975</td>
<td>-0.070</td>
<td>0.011</td>
<td>0.079</td>
<td>-0.121</td>
<td>0.073</td>
</tr>
<tr>
<td></td>
<td>(0.031)</td>
<td>(0.124)</td>
<td>(0.064)</td>
<td>(0.042)</td>
<td>(0.026)</td>
</tr>
<tr>
<td>1980</td>
<td>-0.072</td>
<td>0.009</td>
<td>0.065</td>
<td>-0.174</td>
<td>0.077</td>
</tr>
<tr>
<td></td>
<td>(0.032)</td>
<td>(0.309)</td>
<td>(0.031)</td>
<td>(0.035)</td>
<td>(0.065)</td>
</tr>
<tr>
<td>1985</td>
<td>-0.044</td>
<td>-0.009</td>
<td>0.062</td>
<td>-0.224</td>
<td>0.066</td>
</tr>
<tr>
<td></td>
<td>(0.067)</td>
<td>(0.111)</td>
<td>(0.050)</td>
<td>(0.050)</td>
<td>(0.153)</td>
</tr>
<tr>
<td>1990</td>
<td>-0.046</td>
<td>0.036</td>
<td>0.047</td>
<td>-0.180</td>
<td>0.090</td>
</tr>
<tr>
<td></td>
<td>(0.060)</td>
<td>(0.092)</td>
<td>(0.059)</td>
<td>(0.062)</td>
<td>(0.026)</td>
</tr>
<tr>
<td>1995</td>
<td>-0.047</td>
<td>0.023</td>
<td>0.012</td>
<td>-0.164</td>
<td>0.097</td>
</tr>
<tr>
<td></td>
<td>(0.031)</td>
<td>(0.096)</td>
<td>(0.067)</td>
<td>(0.050)</td>
<td>(0.025)</td>
</tr>
<tr>
<td>2000</td>
<td>-0.089</td>
<td>0.085</td>
<td>0.039</td>
<td>-0.083</td>
<td>0.139</td>
</tr>
<tr>
<td></td>
<td>(0.094)</td>
<td>(0.172)</td>
<td>(0.042)</td>
<td>(0.045)</td>
<td>(0.045)</td>
</tr>
</tbody>
</table>
Figure 1: Kernel Density Plots of LLLS Estimated Coefficients from Model II
Figure 2: Comparison of LLLS Estimated Coefficients of Initial Income from Model II.
Figure 3: Comparison of LLLS Estimated Coefficients of DBACBA from Model II.