8-2013

Failure to Launch? The Role of Land Inequality in Transition Delays

Andros Kourtellos

Ioanna Stylianou

Chih Ming Tan

University of North Dakota, chihming.tan@UND.edu

[How does access to this work benefit you? Let us know!](#)

Follow this and additional works at: <https://commons.und.edu/ef-fac>

Part of the [Economics Commons](#)

Recommended Citation

Andros Kourtellos, Ioanna Stylianou, and Chih Ming Tan. "Failure to Launch? The Role of Land Inequality in Transition Delays" (2013). *Economics & Finance Faculty Publications*. 7.
<https://commons.und.edu/ef-fac/7>

This Article is brought to you for free and open access by the Department of Economics & Finance at UND Scholarly Commons. It has been accepted for inclusion in Economics & Finance Faculty Publications by an authorized administrator of UND Scholarly Commons. For more information, please contact und.common@library.und.edu.

Failure to Launch? The Role of Land Inequality in Transition Delays^{*}

Andros Kourtellos[†] Ioanna Stylianou[‡] Chih Ming Tan[§]

April 25, 2013

Abstract

This paper provides empirical support for one theory of transition delays: initial land inequality. Using a new historical data set for land inequality (Frankema (2009)) we employ duration analysis to investigate whether higher levels of land inequality lead to longer delays in the extension of primary schooling. Our findings suggest that land inequality is a key determinant of delays in schooling.

Keywords: growth takeoffs; schooling; duration analysis; model uncertainty.

JEL Classification Codes: C59, O40, Z12.

^{*}We thank Oded Galor, Yannis Ioannides, Michael Klein, and Enrico Spolaore as well as seminar participants at the Minerva Summer School in Economic Growth for valuable comments.

[†]Department of Economics, University of Cyprus, P.O. Box 537, CY 1678 Nicosia, Cyprus, e-mail: andros@ucy.ac.cy.

[‡]Department of Economics, University of Cyprus, P.O. Box 537, CY 1678 Nicosia, Cyprus, e-mail: sioanna@ucy.ac.cy.

[§]Department of Economics, Clark University, 950 Main Street, Worcester, MA 01610, USA, e-mail: CTan@clarku.edu.

1 Introduction

The transition from economic stagnation to growth and the associated phenomenon of the great divergence has been the subject of intensive research in the growth literature. In particular, there has been a large body of work that is concerned with the issue of economic take-offs. This work describes the transition of economies from a state of economic stagnation to a modern industrial economy with positive growth rates. Notable examples include Galor and Weil (2000), Hansen and Prescott (2001), and Desmet and Parente (2009).

Recent work in the growth literature has focused on the effect of fundamental theories (associated with slow moving determinants) such as geography and institutions on variations in long-run economic performance across countries. In this paper, we are interested in whether these fundamental determinants are also important explanations for delays in countries achieving economic take-offs. Specifically, we focus on the empirical support for one theory of transition delays - initial land inequality.

Many researchers have highlighted the role that initial land inequality plays in terms of delaying the onset of economic take-off. In particular, the theory has highlighted the deep connection between land inequality and human capital accumulation. In Galor, Moav, and Vollrath (2009), henceforth GMV, land inequality negatively affects the implementation of educational reforms that lead to the extension of educational opportunities to the general population.¹ In particular, due to the low complementarity of human capital and land (see also, Galor and Moav (2006)), an increase in the level of human capital increases productivity in industry more than the agricultural sector, causing a decrease in the returns to land and a rise in wages. Consequently, political elites who initially derive most of their income from land have no incentive to support educational reforms. However, since productivity growth in the industrial sector outstrips that in the agricultural sector, the returns from the capital holdings of political elites increase as a proportion of their total income as the economy advances. Their objection to education reform therefore declines over time such that a critical time is reached whereby human capital-enhancing policies (e.g., compulsory schooling) are enacted.

While GMV posit a direct effect of land inequality on transition delays, other work in the

¹Several other works have also documented the relationship between land inequality and the lower provision of other forms of public goods (including financial development), such as Banerjee and Iyer (2005) and Rajan and Ramcharan (2010).

literature also propose an indirect effect whereby land inequality influences the evolution of political institutions, and it is these institutions that then determine the delays in transition. Parente and Prescott (2000), Acemoglu, Johnson, and Robinson (2006), and Engerman and Sokoloff (2002) have all pointed out the important role that land inequality plays in determining the evolution of political institutions. The difference between these works and that of GMV is the emphasis on an independent role for political institutions and their persistence in determining delays in enacting human capital promoting initiatives. For example, as Acemoglu, Johnson, and Robinson (2006) point out, if there are rents to staying in power, then, the politically powerful landed aristocracy would have a strong incentive to block the introduction of new technologies and institutions in order to protect their power and profits, delaying at the same time the industrialization process. The suggestion here is that the autonomous nature of political institutions may require direct reforms to these institutions in order for welfare enhancing outcomes to be achieved. In contrast, in GMV's framework, economic progress automatically leads to a shift in incentives faced by the elites, and to their willingness to adopt human capital enhancing policies.

Our contribution in this paper is twofold. First, we ask the question of what factors determine the delay of a country in achieving a particular education penetration rate (e.g. 50% primary schooling enrollment). Specifically, do higher levels of land inequality lead to longer delays? We exploit a new historical data set for land inequality by Frankema (2009) to investigate this question in the context of hazard rate models. This is a departure from the standard empirical work that is carried out in the growth literature. Methodologically, empirical work in the growth literature focuses on the effects of various covariates on long-run per capita income or growth. In this paper, we focus instead on a more direct prediction of the theory - what are the effects of various fundamental determinants on delay in schooling?

Second, we explicitly address the issue of model uncertainty in investigating how fundamental determinants, such as land inequality, affect the extension of schooling opportunities. As Brock and Durlauf (2001) and others have argued, a characteristic of the new growth theories is their inherent open-endedness. In our context, for example, the hypothesis that land inequality is a determinant of schooling delays does not automatically preclude an alternative theory, such as institutions, from also being a possible explanation. The inherent open-endedness of new growth theories presents unique challenges to researchers in exploring their quantitative consequences on growth. Because the inclusion of one set of growth theories says nothing about whether other possible growth theories should be included

(or not) in the model, growth researchers face substantial model uncertainty in their work. The fear is that the inclusion or exclusion of growth variables may significantly alter the conclusions one had previously arrived at for, say, the effect of land inequality on delays in schooling based on a particular model in the model space.

One approach taken by the literature has been to move away from drawing inferences that are contingent on a particular model. Instead, researchers first determine a space of plausible models and then present inference conditional on the model space. Hence, our analysis does not assume that the GMV theory is necessarily the true one but rather it provides findings that are robust to alternative theories and their proxies. More precisely, we employ a Bayesian model averaging technique that aggregates the findings across different plausible model specifications using the posterior evidence as weights for each model; see for example Durlauf, Kourtellos, and Tan (2008).

Our findings are consistent with the theory proposed by GMV. We find that increased levels of land inequality lead to more delays in the extension of schooling opportunities to the general population. This result is robust to variations in the specification of the hazard model, and holds true for a range of primary schooling penetration rates as well as for the dates when compulsory schooling were legally introduced across the set of countries. Interestingly, initial values of political institutions (as measured by an executive constraint variable) do not appear to be important in determining delays in schooling.

The paper is organized as follows. Section 2 presents the econometric methodology, discusses our data, and presents the results for the hazard analysis. Section 3 provides a robustness analysis and Section 4 concludes and discusses future work.

2 Delay in Schooling

2.1 Implementation

In this paper, we are interested in the probability of the event that a country moves from a low education state to a high education state, which we define as a take-off. In GMV's theory, these two states correspond to a state of economic stagnation and a state of sustained economic growth, respectively.

Assume that countries $i = 1, \dots, n$ enter the low education state at time zero and each country experiences a single spell in this state. Let the delay in schooling be the duration of this spell denoted by the random variable T_i and let its realization be denoted by t_i , which takes a value in $\{1, 2, \dots, \bar{t}\}$, where $\bar{t}$ is the maximum value of the delay in the data. Denote the cumulative distribution and probability density functions of T_i conditional on a $p \times 1$ vector of time-invariant country-specific characteristics/regressors, X_i , by $F_{T_i|X_i}(t_i) = P(T_i \leq t_i|X_i)$ and $f_{T_i|X_i}(t_i)$, respectively. Conditional on X_i the survival function of T_i is defined as $S_{T_i|X_i}(t_i) = 1 - F_{T_i|X_i}(t_i)$, which gives the probability that a take-off has not occurred by duration t_i .

We focus on the rate at which a country leaves the state at duration t_i given that it has not done so yet. Specifically, conditional on X_i , the probability that the spell of low education is completed at t_i given it has not been completed before t_i is defined by the hazard function

$$\psi(t_i|X_i) = P(T_i = t_i|T_i \geq t_i, X_i) = \frac{f_{T_i|X_i}(t_i)}{S_{T_i|X_i}(t_i)}. \quad (2.1)$$

The value of the hazard function for specific t_i is called the hazard rate. The hazard rate is a measure of risk in the sense that higher hazard rates correspond to higher risks of transitioning out of the low education/stagnation state therefore implying shorter delays in the extension of schooling opportunities.

For a country with regressors X_i , a particular parametric specification for the hazard function of the random variable T_i evaluated at the duration t_i is given by

$$\psi(t_i|X_i) = \psi_0(t_i)\exp(X_i'\theta), \quad (2.2)$$

where $\psi_0(t_i)$ is the time dependent part known as the baseline hazard function that describes the countries' risk for transitioning if their risk was independent of their characteristics. This specification allows only the level of the hazard function to differ across countries. The term $\exp(X_i'\theta)$ is called the "systematic part" of the hazard. This particular specification of the hazard regression is known as the static Cox proportional hazard (Cox-PH) model; Cox (1972). θ is the $p \times 1$ vector of parameters in the Cox-PH model estimated by partial likelihood. The Cox-PH model is a semi-parametric model in the sense that while it makes no assumption about the form of the function $\psi_0(t_i)$, it assumes a parametric form for the effect of the predictors on the hazard.²

²For robustness purposes we also investigated alternative parametric hazard models such as the

2.2 Data

In terms of the data, we construct a cross-sectional dataset based on historical data spanning from 1700 to 1998 for a sample of 53 countries. A detailed description of the data and our sources is given in the Data Appendix A1. Table 1 lists the countries in our sample.

The delay in schooling, T_i , is measured in two ways. The first measure is the time it takes for each country to first reach a penetration rate in primary schooling enrollment, minus the time it took the first country to pass that critical value. For example, the United States was the first country to pass the 50% primary school enrollment penetration rate in 1831. Therefore, t_{US} is normalized at 1. In contrast, by the end of the span of the dataset, Mozambique had not achieved that penetration rate yet; hence, $\bar{t}$ equals 164. The reason for constructing the delay variable as a measure that relates primary schooling enrollment in one country relative to the first country to pass the penetration rate is so as to overcome the left censoring problem. Since all other countries achieve the penetration rate at later dates than the first country to do so, left censoring is eliminated. The penetration rates we consider have to fulfill two conditions: (i) they have to be high enough to capture the GMV idea of a large scale extension of public schooling opportunities to the population, and (ii) they have to be low enough so that enough countries attain the level within the sample so that we do not have too many instances of right censoring in the data.

The actual construction of the primary schooling data follows Comin and Hobijn (2004) who construct historical primary schooling data for 23 industrialized countries, measured as the number of students in primary school as a fraction of the population, in the age range 5-14. We extend the primary schooling data set to a larger set of 53 countries for this analysis. The number of students enrolled is taken from Banks (1999), while the population in the age range 5-14 is taken from Mitchell (1998).

The above delay in primary schooling measure is ideal because it directly measures actual penetration rates that were historically achieved. As a robustness check, however, we consider a second measure, which is the delay in introducing a compulsory primary schooling law. Similar to the delay in primary schooling measure, the delay in compulsory primary schooling is defined as the time it takes for each country to introduce compulsory schooling compared to the first country to introduce compulsory schooling. In this case we only have a sample of 50 countries for which the first country was Austria in 1774 while the US (ranked

Exponential, Log Logistic, and Weibull without finding substantial differences.

4th) only passed compulsory primary schooling legislation in 1852. The countries with the longest delays in legislating compulsory primary schooling include Guatemala, Kenya, and Pakistan, all of which had yet to do so by the end of the sample period.

We now discuss the set of explanatory variables, X_i . For our analysis to correspond closely with the theory, we imagine that countries always existed, but have different structural characteristics and historical experiences that influence when they achieve a particular penetration rate in schooling. These factors then explain why a particular country experienced a delay in schooling attainment vis-à-vis the US experience. We think of these factors as controlling for two kinds of country-specific heterogeneity.

The first type of country-specific heterogeneity corresponds to factors that are invariant to the particular political elites that are in power at the time when schooling policy decisions are made. These factors largely correspond to country-specific fixed effects as well as the time it took for the relevant political elites; that is, the political elites who would make policy decisions about schooling and who would see these through, to come to power.

One reason why a country might have experienced a delay in schooling attainment vis-à-vis the US may be because of its colonial history. We do have information about whether a country was historically a European colony. To the extent that we can think of the initial conditions of a colony as being substantially influenced by the European metropolis, we can control for country-specific heterogeneity by including colonial dummy variables (specifically, whether a country was a British colony, a Spanish or Portuguese colony, a French colony, or Other European colony).

Another reason why a country may take more time than the US to attain a particular schooling penetration rate may be that the relevant elites took longer to attain power and therefore control over schooling policies. To control for the variation in the time it took a country's elites to attain autonomy over policies relative to the US, we include an Independence variable that measures the additional years it took for each country to declare independence relative to the US, who declared independence in 1776. This variable takes the value zero for metropolis countries, and positive integers for colonies.

We also control for the elites' hold on power by including a measure of Political Instability due to Miller (2011). The idea is that elites who cannot secure their hold on power may have less ability to influence policy outcomes (or, alternatively, face different incentives in enacting particular policies) hence leading to variation in delays in achieving particular schooling

penetration rates. Political Instability is measured as the average of the first differences (in absolute values) of the Polity2 variable from Polity IV. The Polity2 variable is a measure of the degree of democracy in a country with a score of +10 representing most democratic and -10 signifying most autocratic. The averages of the first differences are calculated as follows: for colonies we average values of the (absolute) year-to-year changes in the Polity2 variable from the year of independence to the year the colony achieves the schooling penetration rate, while for non-colonies, we take the corresponding average values from the earliest available observation until the year the country achieves the schooling penetration rate.

The second set of variables corresponds to factors that influence the incentives of political elites to extend primary schooling opportunities to the population according to the theory. As detailed in the Introduction, our main aim is to investigate how land inequality affects the transition from economic stagnation to the sustained growth era, through the human capital channel. To do so, we use land inequality data from Frankema (2009). The variable is expressed in Gini coefficients, and it is compiled on the basis of the decile distribution of the total number of land holdings (farms), and the total amount of agricultural land (nation-wide), excluding communal pastures and forests.³ Here, a holding refers to “all agricultural land assigned to a “holder” that is one or two persons, but no group, community or state, or to a distinct “management unit”, i.e. a farm. The total agricultural area includes all land that is part of a holding, i.e. arable land, land under permanent crops, land under permanent meadows and pastures, wood and forest land and a category of all other land. In the case of shifting cultivation the total area of the holding consists of the total area under crops and the area that is prepared for cultivation [Frankema, 2006, p. 3]”. The primary data sources that Frankema uses to calculate the land distribution data comes from the IIA and FAO World Census of Agriculture. For our analysis, we use the earliest available land Gini observation for each country.

One concern with using land Gini as a variable is that it may be proxying for other forms of wealth or income inequality. Some forms of wealth inequality may in fact imply dramatically different theoretical outcomes from those of GMV. For example, if inequality was a result of inequality in capital holdings and not of land holdings by elites, it may be the case that elites would prefer higher levels of schooling for the population since human capital is complementary to physical capital. However, if we fail to include a proxy

³The land Gini coefficient is defined as, $G = \sum_{i=1}^n \sum_{j=1}^n |z_i - z_j| / 2n^2\mu$, where z_i and z_j are the percentage shares of land of n deciles ($n = 10$) and $\mu = 1/n$.

variable for capital holdings inequality then the estimates for the effects of land inequality on schooling outcomes are likely to be biased. Alternatively, the precise nature of the inequality responsible for lower schooling levels may be misspecified. For example, land Gini may be proxying for income inequality (instead of land inequality) which has also been shown to be associated with poor education outcomes across countries. In fact, Goldin and Katz (1997) find evidence that supports this proposition for the case of the US. To safeguard against these possibilities, we consider a new dataset of global inequality (BFLZ Gini Index) that has recently been introduced by Van Zanden, Baten, Földvari, and Van Leeuwen (2011). This new dataset is available for a large set of countries spanning from 1820 to 1995 and improves the Bourguignon and Morrisson (2002) dataset in several ways. In particular, it is calculated using a much larger number of observations of within country inequality and it is based on the new 2005 PPPs of the World Bank’s ICP project, which gives a more accurate picture of disparities in GDP per capita than the previous ICP rounds.⁴

GMV also theorize that land abundance that would benefit agriculture in the early stages of development would lead the landowning elite to be more reluctant to enact human capital enhancing policies that disproportionately benefit capitalists and workers. We include therefore a measure of land abundance, the log of arable land (absolute) in hectares, in 1700, and investigate whether more land abundance leads to greater delays in schooling. We also control for total land and total population by including the log of total land in hectares and the log of total population in 1700.

GMV’s theory also requires that we control for other developmental differences between countries. The reason is that for a given level of land inequality, all else equal, higher levels of economic development corresponds to capital holdings constituting a larger proportion of the asset portfolio of elites. Since elites in more developed economies would derive a higher portion of their income from the industrial sector, they would be more willing to enact human capital-friendly policies. We control for initial development differences between countries using the log of GDP per capita (Initial Income; Maddison (2009)). For non-colonies, we take the average of log GDP per capita values from the earliest possible data point until 1831, while for colonies, we use the data on independence day or, if this is unavailable, the earliest data point after independence. We should note that in all cases, the income data occurs prior to the country achieving the schooling penetration rate. The timing of the

⁴We are very grateful to Bas van Leeuwen for very kindly sharing the global inequality data from Van Zanden, Baten, Földvari, and Van Leeuwen (2011) with us.

variable is meant to capture the level of development that was relevant to the elites that are in power at the time when schooling policies are enacted.

The main alternative theory for schooling delays, as noted in the Introduction, is political institutions. We proxy initial political institutions using historical executive constraints data from Polity IV (Initial Executive Constraints). This variable lies between zero and one, with higher values indicating more constraints on the power of the executive. Similar to the Initial Income variable above, we take the average of executive constraints values from the earliest possible data point until 1831 for non-colonies and use the data on independence day or, if unavailable, the earliest data point after independence for colonies. In all cases, the data for executive constraints occurs prior to the country achieving the schooling penetration rate, and is meant to capture the relevant degree of executive constraints that apply to elites empowered to determine schooling outcomes.

Following the empirical growth literature, we also control for a set of new growth theories that have potential implications for human capital accumulation. The first such theory argues that a detrimental climate may have negative effects on human capital accumulation (see, Sachs, Gallup, and Mellinger (1999)). We proxy climate using a variable (Tropics) which measures the percentage of a country’s land area that is classified as tropical or subtropical. Finally, another theory requires that we account for the effects of ethnic heterogeneity on delays in schooling. Alesina, Baqir, and Easterly (1999) suggest that higher levels of ethnic heterogeneity potentially result in political disagreements over the provision of public goods (such as schooling), and its subsequent under-provision. To control for the effect of ethnic heterogeneity on delays in schooling, we include a measure of ethnic fractionalization due to Alesina, Devleeschauwer, Easterly, Kurlat, and Wacziarg (2003) in X_i .

Table 2 presents summary statistics while Table A1 of the Appendix provides a detailed descriptions of all the variables.

2.3 Model averaging of hazard models

Standard duration analysis estimates a baseline PH-Cox model in equation (2.2), which is closest to the theory in question and then reports a few robustness exercises that include some additional controls. An alternative approach to evaluate the relative evidentiary support of competing theories includes a large number of variables and those variables that prove to be

significant are then rendered as the important determinants. This approach is often referred to as a ‘kitchen sink’ approach.

However, both approaches ignore the issue of model uncertainty that arises because of theory open-endedness (as described in the Introduction). Both approaches above essentially present inference that is contingent on a single benchmark model. The evidence therefore relies on strong priors on the part of the econometrician regarding the correct specification of the benchmark model. How can we obtain robust conclusions about the effect of land inequality in equation (2.2) and more generally about the structural parameters θ that do not condition on the model choice?

To systematically address the issue of model uncertainty, we employ a Bayesian Model Averaging (BMA) approach by constructing estimates conditional not on a single model, but on a model space whose elements span an appropriate range of determinants suggested by a large body of work. In the context of this paper, we consider two specifications for the model space, denoted by $\mathcal{M} = \{M_1, \dots, M_K\}$. In the first specification, we are totally agnostic about whether any of the particular growth regressors described in Section 2.2 are in the true model, and therefore assign each regressor a 0.5 prior probability of being in the true model. The model space $\mathcal{M}$, in this case, is composed of the set of all possible permutations of the above growth regressors. This first specification, therefore, deals very generally with the issue of model uncertainty and is consistent with approaches taken in the growth literature such as Fernandez, Ley, and Steel (2001) and Sala-i Martin, Doppelhofer, and Miller (2004). For robustness, we also consider a second specification where either (i) land inequality (Land gini) or (ii) the set of variables associated with GMV’s theory (Land gini, Initial Income, Arable Land, Total Land, and Total Population) are included in every model in the model space.⁵ This second specification addresses a narrower form of model uncertainty and is consistent with the approach of Levine and Renelt (1992). The aim, in this case, is to specifically test the maintained GMV hypothesis that land inequality affects schooling delays while treating the other growth regressors as purely nuisance variables.

A number of recent papers have documented the advantages of using BMA in constructing robust estimates primarily in the context of the linear model.⁶ Our BMA approach is closest to Volinsky, Madigan, Raftery, and Kronmal (1997) who employ BMA

⁵We thank an anonymous referee for this suggestion.

⁶See for example, Brock and Durlauf (2001), Fernandez, Ley, and Steel (2001), Sala-i Martin, Doppelhofer, and Miller (2004), Durlauf, Kourtellos, and Tan (2008), and Masanjala and Papageorgiou (2008).

in the context of Cox-PH models to study the risk factors for stroke. Model averaging “integrates out” the uncertainty over models by taking the weighted average of model-specific estimates, where the weights reflect the evidentiary support for each model given the data, D , and which are constructed to be analogous to posterior model probabilities. Then the posterior distribution of θ given the data, D , is given by

$$\hat{\mu}(\theta|D) = \sum_{k=1}^K \hat{\mu}(\theta|M_k, D) \hat{\mu}(M_k|D), \quad (2.3)$$

where $\hat{\mu}(\theta|M_k, D)$ is the empirical posterior distribution of θ given a particular model M_k , and $\hat{\mu}(M_k|D)$ is the empirical posterior probability of model M_k , which is an element of the model space $\mathcal{M}$. The former is a standard Bayesian object, which does not have a closed form expressions in the case of Cox-PH models. Following Volinsky, Madigan, Raftery, and Kronmal (1997) we use the maximum likelihood approximation $\hat{\mu}(\theta|M_k, D) \approx \hat{\mu}(\theta|M_k, \hat{\theta}_{D,k}, D)$, where $\hat{\theta}_{D,k}$ is the partial likelihood estimator of the model parameter θ_k for a particular model M_k .⁷

As for the model weights, $\hat{\mu}(M_k|D)$ we use the Bayes’ rule, so that each weight is the product of the integrated likelihood of the data given a model, $\hat{\mu}(D|M_k)$, and the prior probability for a model, $\mu(M_k)$:

$$\hat{\mu}(M_k|D) \propto \hat{\mu}(D|M_k) \mu(M_k). \quad (2.4)$$

As standard in the literature, we assume a uniform model prior so that the prior probability that any variable is included in the true model is taken to be 0.5. The integrated likelihood of model M_k is approximated by the Bayesian information criterion (BIC), $\log \hat{\mu}(D|M_k) = \log \hat{\mu}(D|\hat{\theta}_{D,k}, M_k) - (p/2) \log n + O(1)$, where n is the total number of uncensored cases. The BIC approximation to the integrated likelihood implicitly defines that the parameter prior is the unit information prior, which can be viewed as a special case of the Zellner’s (fixed) g-prior that contains information approximately equal to that contained in a single observation; see Kass and Wasserman (1995) and Raftery (1995). Our choice of the priors follows Eicher, Papageorgiou, and Raftery (2011) who found that the unit information prior combined with a uniform prior over the model space generally outperformed competing

⁷This posterior refers to the following integral $\mu(\theta|M_k, D) = \int \mu(\theta|\theta_k, M_k, D) \mu(\theta_k|M_k, D) d\theta_k$, where $\mu(\theta|\theta_k, M_k, D)$ is the likelihood and $\mu(\theta_k|M_k, D) d\theta_k$ is the prior density of θ_k .

priors.

The model averaging estimator of θ is given by the posterior mean defined by

$$\hat{\theta}_{D,\mathcal{M}} = \sum_{k=1}^K \hat{\theta}_{D,k} \hat{\mu}(M_k|D). \quad (2.5)$$

The notation $\hat{\theta}_{D,\mathcal{M}}$ emphasizes the dependence of the estimator on data D and model space $\mathcal{M}$ instead of individual model M_k . We also compute the corresponding model averaging standard errors using the posterior variance of θ

$$\hat{V}_{D,\mathcal{M}}^\theta = \sum_{k=1}^K \hat{V}_{D,k}^\theta \hat{\mu}(M_k|D) + \sum_{k=1}^K (\hat{\theta}_{D,k} - \hat{\theta}_{D,\mathcal{M}})^2 \hat{\mu}(M_k|D), \quad (2.6)$$

where $\hat{V}_{D,k}^\theta$ is the model-specific posterior variance of the partial likelihood estimator. The first term in equation (2.6) is the average of the posterior variances within models and the second term is the variance of the posterior means across models (i.e. weighted average of the squared deviations of the model-specific from the model averaged estimates).⁸

Following the literature one can conduct inference on each covariate's effect θ_j on the delay of schooling in the context of BMA in a number of ways. A standard approach uses the posterior probability of inclusion (PIP) for each covariate. More precisely, this is the posterior probability that a regression coefficient for a variable is nonzero (posterior effect probability) and is computed as the sum of posterior probabilities of the models which contain that variable:

$$\mathbb{P}(\theta_j \neq 0|D) = \sum_{k \in \{s: \theta_j \neq 0 \text{ in } M_s\}} \mathbb{P}(M_k|D) \quad (2.7)$$

The larger the probability of the nonzero effect, the larger the evidence in favor of the covariate j being part of the true theory.⁹

An alternative way to conduct inference is to view equations (2.5) and (2.6) above as a

⁸Following Volinsky, Madigan, Raftery, and Kronmal (1997) we employ the “leaps and bounds algorithm” to choose top K models.

⁹Eicher, Henn, and Papageorgiou (2012) following Kass and Raftery (1995) proposed a classification rule to sort the posterior inclusion probability into a categorical measure of the strength of evidence of the covariate's effect: PIP < 50% indicates lack of evidence for an effect; 50% < PIP < 75% indicates weak evidence for an effect; 75% < PIP < 95% indicates positive evidence for an effect; 95% < PIP < 99% indicates strong evidence for an effect; 99% < PIP < 100% indicates decisive evidence for an effect.

“hybrid” approach to model averaging in the sense that frequentist probability statements about observables given unobservables are mixed with Bayesian probability statement about unobservables given observables. Sala-i Martin, Doppelhofer, and Miller (2004) proposed this approach in the context of growth regressions and argued that the weighting scheme for their “hybrid” model average estimator can be derived as a limiting case of a standard Bayesian analysis as the prior information becomes dominated by the data. For a similar approach, see Durlauf, Kourtellos, and Tan (2011). Papers that interpret equations (2.5) and (2.6) in this way typically report BMA posterior t-statistics for coefficient estimates and interpret them in the classical sense.

However, inference based on a BMA posterior t-statistic must be interpreted with caution because its asymptotic distribution is a mixture of Normal distributions, which is often characterized by irregular shapes, far away from Normal, and thereby rendering inference based on classical interpretations invalid. In the Appendix, Figures A1 and A2 present the BMA posterior probability distributions for the coefficients of the covariates for the hazard models using the delay in primary and compulsory schooling, respectively. These figures also show the posterior probability that each coefficient is zero by a solid line at zero, with height equal to the probability. The nonzero part of the distribution is scaled so that the maximum height is equal to the probability that the coefficient is nonzero.

In the discussion of our findings below, we focus on PIPs for covariates but we also report the corresponding posterior t-statistics and interpret them in the classical (“hybrid”) sense subject to the aforementioned caveats.¹⁰

2.4 Hazard results for delay in schooling

We present our benchmark findings for the Cox PH model in equation (2.2) in Table 4. The dependent variable, delay in schooling, is the time it takes for each country to first reach a particular penetration rate in primary schooling enrollment, minus the time the first country (the US, in this case) passed that level. For the penetration rate to be consistent with the GMV idea of a substantial extension of schooling opportunities to the population, we considered penetration rates ranging from 30% to 65%.

¹⁰Raftery (1995) provides some guidance on how to interpret posterior t-statistics given various sample sizes. In our context, Raftery suggests that a t-statistic greater than 2.43 would correspond to a “positive” statement in support of rejecting the null hypothesis while a t-statistic greater than 3.15 would constitute “strong” evidence for doing so.

When the penetration rate is low (essentially for all penetration rates below 45%), almost all countries successfully attain the penetration rate with very little difference in the time it took to do so, so that there is not enough variation in the data to properly identify the effects of land inequality on schooling delays. However, when the penetration rate is high (above 55%), the number of right censored countries becomes large. Table 3 shows the countries that failed to reach various primary schooling enrollment penetration rates; i.e., countries that are right censored. Right censoring reduces the observed variation in schooling delays, and makes it difficult to identify the effects of land inequality on delays.

For conciseness, we only report full results for the 50% penetration rate in Table 4. The results for land gini as well as the other covariates for penetration rates between about 40% and 50% do not differ substantively. This can be seen from Figure 1, which shows the PIPs for the land gini variable as a function of penetration rates in primary schooling enrollment. Full results for all other covariates are available upon request.

The first three columns of Table 4 present the results from our model averaging analysis. Here, we employ the more general specification for model uncertainty (as discussed in Section 2.2) and assume that each regressor has a 0.5 prior probability of being in the true model. In the next subsection, we consider the case when the GMV variables are always included in all models. The first column shows the posterior probability that each of the covariates is included in the true model for the hazard rate, while the second and third columns present the BMA posterior means and standard errors for each covariate. The remaining four columns show, respectively, the coefficient estimate and standard error for each covariate for the posterior mode model from the BMA analysis, and the largest model in the model space considered in the BMA analysis.

Our reason for reporting the results from the posterior mode and largest model is to provide the reader with the ability to compare findings via model selection - using the best model (in terms of posterior weights) or a low-bias model (at the cost of reduced efficiency) with potentially many irrelevant covariates - with those obtained via model averaging (BMA). Finally, we also note that the posterior means are interpreted as the marginal effect of each covariate on the risk of crossing the 50% primary schooling penetration rate. Therefore, positive estimates imply that the marginal contribution of the corresponding covariate is to reduce the delay in schooling for countries.

Our BMA results are consistent with the theoretical predictions of GMV. As GMV

argued, for given levels of economic progress, land inequality implies a higher reliance of political elites on income derived from landholdings leading them to delay the implementation of human capital enhancing policies, which primarily benefit capitalists and workers. Similarly, the greater the abundance of arable land, all else equal, the greater the importance of agriculture in the elites' portfolio, the higher their subsequent reliance on returns from landholdings, and the greater their reluctance to expand schooling opportunities. However, for given levels of land inequality and arable land, economic progress results in a rebalancing of the portfolio returns of landholding elites away from income derived from land holdings to returns from capital holdings resulting in elites being more willing to extend schooling to the population.

Consistent with the theory, we find that higher levels of land inequality (higher values for Land Gini), greater abundance of Arable Land, and lower Initial Income result in lower risks of exceeding the 50% schooling penetration rate, thereby implying greater delays in the expansion of schooling opportunities. More precisely, the posterior inclusion probabilities of Land Gini, Arable Land, and Initial Income are all very high at 99%, 90.7%, and 98%, respectively - well above the 50% prior inclusion probability. The corresponding posterior means for all three variables are also strongly significant at the 1% level for Initial Income and Land Gini and at the 5% level for Arable Land. In particular, a 1% decrease in land gini yields approximately a 6.5% increase in the relative hazard risk of taking off from stagnation. The BMA findings are confirmed by the results from the posterior mode model and the largest model. Interestingly, the posterior mode model includes all the model averaging covariates with statistically significant posterior mean. Note, however, the posterior model probability for the mode model is 0.165, whereas the largest model has posterior model probability of 0.000 suggesting that the latter is a rather poor model choice.

Table 4 also makes clear that it is inequality in land ownership specifically, and not other (non-land) forms of inequality (as proxied for by BFLZ Gini Index) that is important in determining schooling delays. The posterior inclusion probability for BFLZ Gini Index is at 10.1% whereas the posterior mean is not significant. Noticeable is also the fact that the variable is not included in the posterior mode model.

Two sets of factors that can be interpreted as country fixed effects are shown to be strongly significant. The first is the delay in a country gaining independence relative to the US. We find, predictably, that countries that took more time to gain independence, so that the relevant elites required more time to attain autonomous control over policies, also

faced longer delays in achieving an extension of schooling opportunities to the population, all else equal. The posterior inclusion probability of the Independence variable is 100% and significant at the 1% level.

Along with gaining autonomy over a country’s policies, the level of Political Instability (elites’ hold over power) is also important (with posterior inclusion probability of 100%) and highly significant at the 1% level. Our BMA findings (consistent with those of the other reported models) indicate that a greater degree of Political Instability, all else equal, leads to longer delays in reaching the 50% schooling penetration rate. An interesting related finding is that there is no evidence that initial institutions (as measured by Initial Executive Constraints) affects schooling delays. The posterior inclusion probability for Initial Executive Constraints is well below 50% at 12.9%, and the posterior mean is not significant. Initial Executive Constraints also does not appear in the posterior mode model. One reason for this negative finding may be that Political Instability is already a de facto constraint on executive power, and therefore the effect of institutions on schooling delays flows through this channel.

Furthermore, being either a British colony or some other colony that is not French, Spanish, or Portuguese results in a shorter delay in achieving schooling take-off. The posterior inclusion probabilities for the British colony and Other colony dummies are both very high at 96.2% and 100%, respectively, and the corresponding coefficient estimates are strongly significant. None of the other growth theories appear to be either significant or important (in terms of posterior inclusion probabilities) explanations for delays in achieving the schooling penetration rate.

As a final note, a careful look into the individual posterior model probabilities suggests that the posterior mode model is not a dominant model but rather the posterior mass is spread evenly, and over larger models, resulting in a high share of important covariates. For example, beyond the posterior mode, the next best four models carry probabilities 0.150, 0.057, 0.049, and 0.036. Figures A3 and A4 in the Appendix show the cumulative posterior probabilities of the top 25 models. This evidence provides some reassurance that the results do not suffer from the supermodel effect behavior, which refers to the spurious concentration of posterior mass on a single or a few models. In particular, the concern is that the fixed prior structure (Zellners g-prior) may exert non-negligible influence on the posterior model probabilities and hence on inference; see for example Feldkircher and Zeugner (2009).¹¹

¹¹To overcome this problem some authors proposed, in the context of the linear regression, the use of flexible

In sum, our findings appear to provide strong support for the hypothesis that schooling delays are entirely explained by variables suggested by GMV’s theory. More precisely, we find that higher levels of land inequality lead to longer delays in the introduction of a compulsory primary schooling law.

3 Robustness

We provide two complementary sets of robustness exercises. First, we consider an alternative measure for schooling delays. Instead of using the 50% penetration rate in primary schooling enrollment as a measure of schooling delays, we now estimate a hazard model for delays in the introduction of a compulsory primary schooling law. The dependent variable here is the time it takes for each country to introduce compulsory schooling compared to the first country to introduce compulsory schooling. According to Table 5, and consistently with the theory and the earlier findings, higher levels of land inequality results in increased delays in introducing a compulsory schooling law. More precisely, the posterior inclusion probability of Land Gini is again very high at 90.6% - well above the 50% prior inclusion probability. The corresponding posterior mean is also negative and significant at the 5% level. The BMA results are also confirmed by the results from the posterior mode model.

Second, as discussed in Section 2.3, we present results for a more restrictive model averaging exercise. Table 6 shows the various cases where we include either Land gini or the entire set of GMV variables in every model in the model space. We then treat the other regressors as nuisance variables and allow them to vary across models. We report results for both our main dependent variable; i.e., the time it took to achieve the 50% penetration rate in primary schooling enrollment, as well as the delay in the introduction of compulsory primary schooling variable as described above. As with the previous robustness exercise, we find that our results confirm the benchmark findings in Section 2.4. In all cases, the coefficient to Land gini is negative and highly significant at the 1% level. The magnitude of the partial effect is also slightly larger in this case.

prior structures, a hyper-g prior whose data-dependent shrinkage adapts posterior model distributions to data quality, but the evidence appears to be mixed or preliminary. For example, Liang, Paulo, Molina, Clyde, and Berger (2008) were not able to find strong evidence to demonstrate that the hyper-g prior is better than the fixed g-prior. In contrast, in a complementary paper, Feldkircher and Zeugner (2009) demonstrate that the use of hyper-g prior distributes the posterior mass more evenly than fixed g-priors.

We conclude therefore that our benchmark findings are robust to alternative specifications for schooling delays. They are also consistent with the results from a narrower test of the GMV hypothesis that maintains the GMV variables in the regression equation while accounting for model uncertainty induced by the presence of auxiliary growth theories.

4 Conclusion and Discussion

In this paper, we empirically investigate the direct predictions of the theory of Galor, Moav, and Vollrath (2009) that higher levels of land inequality result in delays in the implementation of human capital enhancing policies. Using new historical data by Frankema (2009), we test the importance of land inequality as a determinant of delays in the extension of schooling opportunities against alternatives theories. Using BMA analysis in hazard models, we find that higher levels of land inequality result in delays in countries reaching particular schooling penetration rates. Moreover, the BMA analysis provides a useful way to account for model uncertainty by summarizing evidence from all possible models rather than focusing on a single model.

In future work, we plan to extent the analysis in this paper to investigate the effects of schooling delays on long-run economic growth. An empirically testable long-run implication of GMV that is also generally related to the Unified Growth Theory (UGT); see, Galor and Weil (1999, 2000), Galor and Moav (2002), and Galor (2005), is that the differential timing of take-offs from stagnation to growth ultimately segments economies into convergence clubs. We therefore plan to investigate whether observed values for historical schooling delays may be responsible for segmenting countries into multiple growth regimes.

Figure 1: Posterior inclusion probabilities (PIPs) for the land gini variable as a function of penetration rates in primary schooling enrollment.

Table 1: List of Countries

This table presents the list of countries. In the bracket we note the date of the actual starting date of each country.

Europe	Latin America	Middle East and North Africa
Austria (1919)	Argentina (1895)	Algeria (1962)
Belgium (1860)	Brazil (1872)	Egypt (1951)
Denmark (1882)	Chile (1895)	Iran (1887)
France (1851)	Colombia (1938)	Iraq (1957)
Greece (1870)	Costa Rica (1883)	Morocco (1956)
Ireland (1926)	Dom. Rep (1844)	Tunisia (1956)
Netherlands (1899)	Ecuador (1950)	
Norway (1855)	El Salvador (1930)	South Asia
Portugal (1864)	Guatemala (1950)	India (1947)
Romania (1899)	Honduras (1930)	Pakistan (1961)
Sweden (1882)	Nicaragua (1950)	Sri Lanka (1948)
UK (1860)	Panama (1950)	
	Paraguay (1950)	Central Asia
Offshoots	Peru (1961)	Turkey (1935)
Australia (1946)		
Canada (1865)	East Asia and Pacific	
N. Zealand (1907)	China (1953)	
USA (1830)	Indonesia (1961)	
	Japan (1815)	
Sub-Saharan	Korea Rep. (1960)	
Ghana (1960)	Malaysia (1957)	
Kenya (1969)	Myanmar (1948)	
Mozambique (1975)	Philippines (1948)	
South Africa (1946)	Thailand (1929)	
Zambia (1969)		

Table 2: Descriptive Statistics

This table presents the summary statistics for the 53 countries of our dataset.

	Mean	Std. Dev.	Min	Max
Delay in Primary Schooling	110.6	40.10	1	164
Delay in Compulsory Schooling	162.2	57.21	1	224
Initial Income	7.010	0.482	6.116	8.583
Initial Exec. Cons.	0.370	0.399	0.000	1.000
Political Instability	0.278	0.385	0.000	2.375
BFLZ Gini Index	0.491	0.110	0.276	0.794
Land Gini	0.636	0.144	0.307	0.863
Arable Land	6.934	1.573	3.367	11.27
Total Population	7.343	1.728	4.143	12.01
Total Land	10.60	1.580	7.784	13.77
Independence	84.00	73.10	0.000	199.0
Tropics	0.376	0.417	0.000	1.000
Ethnic Frac.	0.373	0.247	0.012	0.859
British Colony	0.302	0.463	0.000	1.000
French Colony	0.057	0.233	0.000	1.000
Span./Port Colony	0.302	0.463	0.000	1.000
Other Colonies	0.113	0.320	0.000	1.000

Table 3: Primary Schooling Penetration Rate Failure

This table lists the countries that fail to attain a given penetration rate in primary schooling enrollment for various penetration rates.

40%	45%	50%	55%	60%	65%	70%
Morocco	India	Egypt	Egypt	Egypt	China	Algeria
Pakistan	Morocco	India	Guatemala	Ghana	Colombia	China
	Pakistan	Morocco	India	Guatemala	Egypt	Colombia
		Mozambique	Iraq	Honduras	Ghana	Costa Rica
		Pakistan	Morocco	India	Guatemala	Egypt
			Mozambique	Iran	Honduras	El Salvador
			Myanmar	Iraq	India	Ghana
			Nicaragua	Morocco	Iran	Greece
			Pakistan	Mozambique	Iraq	Guatemala
			Turkey	Myanmar	Korea Republic	Honduras
				Nicaragua	Morocco	India
				Pakistan	Mozambique	Iran
				Turkey	Myanmar	Iraq
					Nicaragua	Japan
					Pakistan	Korea Republic
					Thailand	Morocco
					Turkey	Mozambique
					Zambia	Myanmar
						Nicaragua
						Pakistan
						Paraguay
						Thailand
						Turkey
						Zambia

Table 4: Hazard Model for the Delay in Primary Schooling

The table presents BMA results for the Cox-PH duration model. The posterior inclusion probability (PIP) of a variable is the sum of the posterior probabilities models that include that variable. The posterior mean (PM) is the average of the partial likelihood coefficient estimates (COEF) of individual models weighted by posterior probability. The posterior standard error (PSE) is the BMA estimate for the standard error (SE) taking model uncertainty into account. All reported standard errors refer to heteroskedasticity consistent standard errors. ***, **, and * denote significance at 1%, 5%, and 10%, respectively.

	<u>Model Averaging</u>			<u>Posterior Mode</u>		<u>Largest</u>	
	PIP	PM	PSE	COEF	SE	COEF	SE
Land gini	99.0	-6.5499***	2.0889	-6.8442***	1.9240	-6.4240***	2.0255
Initial Income	98.0	1.3818***	0.5067	1.4561***	0.3901	1.2741**	0.5884
Arable Land	90.7	-0.3914**	0.1917	-0.4225***	0.1255	-0.3850*	0.2222
Total Land	8.6	-0.0084	0.0511	-	-	-0.0223	0.1488
Total Population	11.5	-0.0213	0.1038	-	-	-0.0616	0.2005
British Colony	96.2	2.1990**	0.8716	2.3066***	0.6135	2.3614*	1.2712
Span./Port. Colony	8.2	-0.0410	0.2292	-	-	-0.4015	0.7610
French Colony	16.0	0.2377	0.7276	-	-	1.5705	1.5621
Other Colonies	100.0	2.6569***	0.7838	2.6228***	0.6439	3.1831***	1.0227
Independence	100.0	-0.0234***	0.0054	-0.0235***	0.0046	-0.0279***	0.0080
Political Instability	100.0	-2.3603***	0.8898	-2.3131***	0.7411	-2.3003**	1.0115
Initial Exec. Cons.	12.9	0.1204	0.4456	-	-	0.7721	0.7475
Tropics	12.5	-0.0642	0.2512	-	-	0.1920	0.9229
Ethnic Frac.	39.0	-0.5629	0.8587	-	-	-1.1971	0.9417
BFLZ Gini Index	10.1	0.1696	0.7714	-	-	2.1577	2.1915
Wald Statistic				42.80		52.42	
Posterior Model Probability				0.165		0.000	

Table 5: Hazard Model for the Delay in Compulsory Schooling

The table presents BMA results for the Cox-PH duration model. The posterior inclusion probability (PIP) of a variable is the sum of the posterior probabilities models that include that variable. The posterior mean (PM) is the average of the partial likelihood coefficient estimates (COEF) of individual models weighted by posterior probability. The posterior standard error (PSE) is the BMA estimate for the standard error (SE) taking model uncertainty into account. All reported standard errors refer to heteroskedasticity consistent standard errors. ***, **, and * denote significance at 1%, 5%, and 10%, respectively.

	<u>Model Averaging</u>			<u>Posterior Mode</u>		<u>Largest</u>	
	PIP	PM	PSE	COEF	SE	COEF	SE
Land gini	90.6	-4.6400**	2.2970	-5.1309***	1.2034	-4.5010***	1.3918
Initial Income	91.0	1.1788**	0.5866	1.4142***	0.3954	1.2249**	0.5876
Arable Land	7.8	0.0095	0.0550	-	-	0.0393	0.2424
Total Land	93.3	0.3333**	0.1511	0.4163***	0.1080	0.3895***	0.1126
Total Population	5.7	-0.0004	0.0410	-	-	-0.0600	0.2487
British Colony	9.5	0.0533	0.2462	-	-	0.3713	0.9218
Span./Port. Colony	8.2	-0.0324	0.1758	-	-	-0.1884	0.6721
French Colony	13.4	-0.1261	0.4427	-	-	0.3561	1.2753
Other Colonies	45.8	0.6427	0.8455	1.4122**	0.6902	1.7615*	1.0220
Independence	100.0	-0.0115***	0.0036	-0.0123***	0.0028	-0.0147 ***	0.0056
Political Instability	100.0	-2.8292***	1.0404	-3.3992***	0.9948	-3.5118***	1.3231
Initial Exec. Cons.	7.9	0.0387	0.2125	-	-	0.2396	0.7668
Tropics	10.8	-0.0548	0.2437	-	-	-0.1263	0.5372
Ethnic Frac.	95.0	-2.4218**	1.0129	-2.4756***	0.6493	-2.4497***	0.8252
BFLZ Gini Index	56.4	2.1925	2.3665	4.2561***	1.5141	4.3827**	1.7977
Wald Statistic				53.58		71.44	
Posterior Model Probability				0.104		0.000	

Table 6: Alternative Model Spaces

This model averaging analysis using different model spaces for Cox-PH duration model using both Delay in Primary Schooling and Delay in Compulsory Schooling. Model Space I always keeps in Land gini in the model and Model Space II always keeps all the variables suggested by GMV: Land Gini, Total Land, Arable Land, Initial Income, and Total Population.

	Delay in Primary Schooling						Delay in Compulsory Schooling					
	Model Space I			Model Space II			Model Space I			Model Space II		
	PIP	PM	PSE	PIP	PM	PSE	PIP	PM	PSE	PIP	PM	PSE
Land gini	-	-6.6131***	1.9969	-	-6.6858***	1.9849	-	-5.1224***	1.8314	-	-5.0562***	1.9364
Initial Income	97.9	1.3842***	0.5073	-	1.4385***	0.5280	94.3	1.2460**	0.5492	-	1.2132**	0.5707
Arable Land	91.6	-0.3952**	0.1887	-	-0.4085*	0.2116	7.7	0.0083	0.0523	-	0.1471	0.2253
Total Land	8.7	-0.0085	0.0513	-	-0.0662	0.1383	92.6	0.3351**	0.1540	-	0.3383**	0.1336
Total Population	11.6	-0.0215	0.1043	-	0.0125	0.2050	6.2	-0.0005	0.0431	-	-0.0988	0.2238
British Colony	96.2	2.2068**	0.8701	95.4	2.2915**	0.9554	8.8	0.0431	0.2192	10.3	0.0383	0.2336
Span./Port. Colony	8.3	-0.0414	0.2303	12.2	-0.0668	0.2952	7.3	-0.0232	0.1517	9.0	-0.0205	0.1934
French Colony	16.2	0.2400	0.7308	20.3	0.3326	0.8674	14.8	-0.1392	0.4632	16.9	-0.1471	0.4821
Other Colonies	100.0	2.6555***	0.7846	100.0	2.6946***	0.8202	40.9	0.5443	0.7943	47.9	0.6704	0.8516
Independence	100.0	-0.0234***	0.0053	100.0	-0.0242***	0.0057	100.0	-0.0116***	0.0036	100.0	-0.0120***	0.0034
Political Instability	100.0	-2.3628***	0.8899	100.0	-2.3834***	0.9089	100.0	-2.7748***	1.0332	100.0	-2.9696***	1.1004
Initial Exec. Cons.	13.1	0.1215	0.4476	16.8	0.1537	0.4929	8.0	0.0371	0.2076	9.9	0.0355	0.2207
Tropics	12.6	-0.0648	0.2523	14.4	-0.0704	0.2666	11.9	-0.0605	0.2554	8.6	-0.0128	0.1686
Ethnic Frac.	38.4	-0.5518	0.8519	32.5	-0.4578	0.8129	94.5	-2.3795**	1.0205	99.1	-2.5756***	0.8926
BFLZ Gini Index	10.2	0.1713	0.7749	11.1	0.1673	0.7809	59.6	2.3489	2.3903	54.5	2.0696	2.3265

References

- Acemoglu, A., S. Johnson, and J. Robinson, 2006, Economic Backwardness in Political Perspective, *American Political Science Review* 100, 115–131.
- Alesina, A., R. Baqir, and W. Easterly, 1999, Public goods and ethnic divisions, *Quarterly Journal of Economics* 114, 1243–1284.
- Alesina, A., A. Devleeschauwer, W. Easterly, S. Kurlat, and R. Wacziarg, 2003, Fractionalization, *Journal of Economic Growth* 8, 155–194.
- Banerjee, A., and L. Iyer, 2005, History, Institutions, and Economic Performance: The Legacy of Colonial Land Tenure Systems in India, *American Economic Review* 95, 1190–1213.
- Banks, A., 1999, *Cross-National Time-Series Data Archive* (Binghamton, New York: Databanks International).
- Bourguignon, F., and C. Morrisson, 2002, Inequality among world citizens: 1890-1992, *American Economic Review* 92, 1190–1213.
- Brock, W., and S. D. Durlauf, 2001, Growth Empirics and Reality, *World Bank Economic Review* 15, 229–272.
- Comin, D., and B. Hobijn, 2004, Cross-Country Technological Adoption: Making the Theories Face the Facts, *Journal of Monetary Economics* 51, 39–83.
- Cox, D. R., 1972, Regression Models and Life Tables, *Journal of the Royal Statistical Society Series B* 34, 187–220.
- Desmet, K., and P. Parente, 2009, The Evolution of Markets and the Revolution of Industry: A Unified Theory of Growth, CEPR Discussion Paper 7290.
- Durlauf, S.N., A. Kourtellos, and C.M. Tan, 2011, Is God in the Details? A Reexamination of the Role of Religion in Economic Growth, *Journal of Applied Econometrics*, forthcoming.
- Durlauf, S. N., A. Kourtellos, and C. M. Tan, 2008, Are Any Growth Theories Robust?, *Economic Journal* 118, 329–346.

- Eicher, T., C. Henn, and C. Papageorgiou, 2012, Trade Creation and Diversion Revisted: Accounting for Model Uncertainty and Natural Trading Partner Effects, *Journal of Applied Econometrics* 27, 296–321.
- Eicher, T., C. Papageorgiou, and A. E. Raftery, 2011, Default Priors And Predictive Performance In Bayesian Model Averaging, With Application To Growth Determinants, *Journal of Applied Econometrics* 26, 30–55.
- Engerman, S., and K. Sokoloff, 2002, Factor Endowments, Inequality and Paths of Development Among New World Economies, NBER Working Papers, 9259.
- Feldkircher, M., and S. Zeugner, 2009, Benchmark priors revisited: On adaptive shrinkage and the supermodel effect in bayesian model averaging, IMF Working Paper.
- Fernandez, C., E. Ley, and M. Steel, 2001, Model Uncertainty in Cross-Country Growth Regressions, *Journal of Applied Econometrics* 16, 563–576.
- Frankema, Ewout, 2009, The Colonial Roots of Land Inequality: Geography, Factor Endowments, or Institutions?, *Economic History Review* 63, 418–451.
- Galor, O., and O. Moav, 2006, Das Human Kapital: A Theory of the Demise of the Class Structure, *Review of Economic Studies* 73, 85–117.
- , and D. Vollrath, 2009, Inequality in Land Ownership, the Emergence of Human Capital Promoting Institutions, and the Great Divergence, *Review of Economic Studies* 76, 143–179.
- Galor, O., and D. Weil, 2000, Population, Technology, and Growth: From Malthusian Stagnation to the Demographic Transition and Beyond, *American Economic Review* 90, 806–828.
- Goldin, C., and L.F. Katz, 1997, Why the United States Led in Education: Lessons from Secondary School Expansion, 1910 to 1940, NBER Working Papers, 6144.
- Hansen, G., and E. C. Prescott, 2001, Malthus to Solow, *American Economic Review* 92, 1205–17.
- Kass, R., and L. Wasserman, 1995, A reference bayesian rest for nested hypotheses and its relationship to the schwarz criterion, *Journal of the American Statistical Association* 90, 928–934.

- Kass, R. E., and A. E. Raftery, 1995, Bayes Factors, *Journal of the American Statistical Association* 90, 773–795.
- Levine, R., and D. Renelt, 1992, A Sensitivity Analysis of Cross-Country Growth Regressions, *American Economic Review* 82, 942–963.
- Liang, F., R. Paulo, G. Molina, M. A. Clyde, and J. O. Berger, 2008, Mixtures of g priors for bayesian variable selection, *Journal of the American Statistical Association* 103, 410–423.
- Maddison, A., 2009, *Statistics on World Population, GDP, and Per Capita GDP, 1-2006 AD*.
- Masanjala, W., and C. Papageorgiou, 2008, Rough and Lonely Road to Prosperity: A Reexamination of the Sources of Growth in Africa Using Bayesian Model Averaging, *Journal of Applied Econometrics* 23, 671–682.
- Miller, M. K., 2011, Democracy by Example? Economic Growth, Policy Diffusion, and Regime Change., Princeton, mimeo.
- Mitchell, B.R, 1998, *International Historical Statistics: The Americas, Africa, Asia, Oceania, and Europe 1750-1993* (MacMillan Press).
- Parente, S., and E. Prescott, 2000, *Barriers to Riches* (MIT Press).
- Raftery, A. E., 1995, Bayesian Model Selection in Social Research (with Discussion), *Sociological Methodology* 25, 111–196.
- Rajan, R. G., and R. Ramcharan, 2010, Land and Credit: A Study of the Political Economy of Banking in the United States in the Early 20th Century, *Journal of Finance* (forthcoming).
- Ramankutty, N., and J. A. Foley, 1999, Estimating historical changes in global land cover: Croplands from 1700 to 1992, *Global Biogeochemical Cycles* 13, 997–1027.
- Sachs, J., J. L. Gallup, and A. Mellinger, 1999, Geography and Economic Development, *International Regional Science Review* 22, 172–232.
- Sala-i Martin, X., G. Doppelhofer, and R. Miller, 2004, Determinants of Long-term Growth: a Bayesian Averaging of Classical Estimates (BACE) Approach, *American Economic Review* 94, 813–835.

- Van Zanden, J.L., J. Baten, P. Földvari, and B. Van Leeuwen, 2011, The Changing Shape of Global Inequality: Exploring a new dataset, Center for Global Economic History (CGEH), No. 1, CGEH working paper series.
- Volinsky, C., D. Madigan, A. Raftery, and R. Kronmal, 1997, Bayesian Model Averaging in Proportional Hazards Model: Predicting the Risk of a Stroke, *Applied Statistics* 46, 443–448.

Table A1: Data Appendix

Variable	Description
Delay in Schooling	Following the methodology of Comin and Hobijn (2004), we construct historical data for primary schooling enrollments, measured as the number of students in primary school as a fraction of population between 5-14. First, we verify the dataset of Comin and Hobijn (2004), which is limited to 23 industrialized countries and then expand it to 53 countries. Using this new dataset we create the delay in schooling variable, which is the time it takes for each country to first reach a penetration rate in primary schooling enrollment, minus the time it took the first country to pass that penetration rate. Source: Mitchell (1998) for the population data; Banks (1999) for the number of students.
Delays in compulsory schooling	It is the time it takes for each country to introduce compulsory schooling legislation comparing to the first country that introduced such legislation. Source: Various.
Initial Income	Log of GDP per capita, where for the colonies we use the independence date or earliest available, and for the non-colonies the average of earliest possible until 1831 (penetration rate of 50%). Source: Maddison (2009).
Initial Executive Constraints	Institutional variable with the lowest value 0 indicating unlimited executive authority and 1 executive parity or subordination. For the colonies we use the independence date or earliest available, and for the non-colonies the average of earliest possible until 1831 (penetration rate of 50%). Source: Polity IV, http://www.systemicpeace.org/polity/polity4.htm
BFLZ Gini Index	BFLZ Gini Index is based on a large number of observations of within country inequality spanning from 1820-1995. For non-colonies we use earliest available and for colonies we use the Independence date and if not available we use the earliest available after Independence. In particular, the Gini Index is based on direct income Gini estimates; estimates of the net household or expenditure Ginis; Ginis based on income shares; Williamson index, which is the ratio between GDP per capita and real wages of unskilled laborers; and height inequality data. Source: Van Zanden, Baten, Földvari, and Van Leeuwen (2011).

Table A1 continued

Variable	Description
Land Gini	The gini coefficient of the size distribution of land. For all countries we use the earliest observation available. We use data over the period 1880-1999. Source: Frankema (2009).
Arable Land	Log of arable land (absolute) in hectares, in 1700. Source: Ramankutty and Foley (1999)
Total Land	Log of total land in hectares, in 1700. Source: Ramankutty and Foley (1999)
Total Population	Log of total population in 1700. Maddison (2009).
Independence	Independence The time it takes for each country to declare independence relative to the United States who declared independence in 1776. Source: CIA Factbook, https://www.cia.gov/library/publications/the-world-factbook/
Tropics	Percentage of land area classified as tropical and subtropical via the Koeppen-Geiger system. Source: CID at Harvard, http://www.cid.harvard.edu/ciddata/ciddata.html
Ethnic Fractionalization	Variable which combines racial and linguistic characteristics. Source: Alesina, Devleeschauwer, Easterly, Kurlat, and Wacziarg (2003).
Colonial Dummies	Coded zero or one. One indicates that country was colonized by Britain, France, Spain or Portugal. Source: CIA Factbook, https://www.cia.gov/library/publications/the-world-factbook/
Political Instability	Political Instability is measured as the average of the first differences (in absolute values) of the Polity2 variable from Polity IV. The Polity2 variable is a measure of the degree of democracy in a country with a score of +10 representing most democratic and -10 signifying most autocratic. The averages of the first differences are calculated as follows: for colonies we average values of the (absolute) year-to-year changes in the Polity2 variable from the year of independence to the year the colony achieves the schooling penetration rate, while for non-colonies, we take the corresponding average values from the earliest available observation until the year the country achieves the schooling penetration rate. Source: Polity IV, http://www.systemicpeace.org/polity/polity4.htm

Figure A1: BMA Posterior Distributions for θ - Delay in Primary Schooling

Figure A2: BMA Posterior Distributions for θ - Delay in Compulsory Schooling

Figure A3: Cumulative Posterior Model Probabilities for Delay in Primary Schooling for the Top 25 Models

Figure A4: Cumulative Posterior Model Probabilities for Delay in Compulsory Schooling for the Top 25 Models