

RE-EXAMINING THE MIDDLE-INCOME TRAP HYPOTHESIS: WHAT TO REJECT AND WHAT TO REVIVE?

Xuehui Han and Shang-Jin Wei

NO. 436

July 2015

**ADB ECONOMICS
WORKING PAPER SERIES**



ADB Economics Working Paper Series

Re-examining the Middle-Income Trap Hypothesis: What to Reject and What to Revive?

Xuehui Han and Shang-Jin Wei

No. 436 | July 2015

Xuehui Han (xuehuihan@adb.org), Economist, Asian Development Bank (ADB) and Shang-Jin Wei (swei@adb.org), Chief Economist and Director General, ADB.

The authors are grateful to Orlee P. Velarde, Lea R. Sumulong, and Nedelyn Magtibay-Ramos for excellent research assistance; and to Francesco Caselli, David Cook, Lei Lei Song, and participants of the research seminar at ADB and the ABFER annual conference for useful comments. The authors are responsible for any remaining errors in the paper.

Asian Development Bank
6 ADB Avenue, Mandaluyong City
1550 Metro Manila, Philippines
www.adb.org

© 2015 by Asian Development Bank
July 2015
ISSN 2313-6537 (Print), 2313-6545 (e-ISSN)
Publication Stock No. WPS157472-2

The views expressed in this paper are those of the authors and do not necessarily reflect the views and policies of the Asian Development Bank (ADB) or its Board of Governors or the governments they represent.

ADB does not guarantee the accuracy of the data included in this publication and accepts no responsibility for any consequence of their use.

By making any designation of or reference to a particular territory or geographic area, or by using the term “country” in this document, ADB does not intend to make any judgments as to the legal or other status of any territory or area.

Note: In this publication, “\$” refers to US dollars.

The ADB Economics Working Paper Series is a forum for stimulating discussion and eliciting feedback on ongoing and recently completed research and policy studies undertaken by the Asian Development Bank (ADB) staff, consultants, or resource persons. The series deals with key economic and development problems, particularly those facing the Asia and Pacific region; as well as conceptual, analytical, or methodological issues relating to project/program economic analysis, and statistical data and measurement. The series aims to enhance the knowledge on Asia’s development and policy challenges; strengthen analytical rigor and quality of ADB’s country partnership strategies, and its subregional and country operations; and improve the quality and availability of statistical data and development indicators for monitoring development effectiveness.

The ADB Economics Working Paper Series is a quick-disseminating, informal publication whose titles could subsequently be revised for publication as articles in professional journals or chapters in books. The series is maintained by the Economic Research and Regional Cooperation Department.

CONTENTS

TABLES AND FIGURES	iv
ABSTRACT	v
I. INTRODUCTION	1
II. (UNCONDITIONAL) ECONOMIC TRANSITIONS IN THE LONG RUN	2
A. Transition Matrix and Ergodic Distribution	5
B. Ergodic Distribution Analysis on Convergence in Relative Terms to the United States	6
C. Long-Horizon Analysis with Maddison Data	7
III. EVOLVING CONSTRAINTS TO GROWTH: A PERSPECTIVE FROM REGRESSION TREES AND RANDOM FORESTS	8
A. Variables that Could Alter Growth	8
B. Box Whisker Plot and Pair-Wise Correlation Analysis	11
C. Conditional Regression Tree Analysis	17
D. Robustness Check with Random Forest Analysis	20
IV. CONCLUSION: LINK THE CONDITIONAL AND UNCONDITIONAL ANALYSES	23
APPENDIX	27
REFERENCES	31

TABLES AND FIGURES

TABLES

1	Decade-Average Transition Matrix for 1960–2010	5
2	Decades Needed for X Percent of Economies to Move Up and Out of Their Current Group	6
3	Decade-Average Transition Matrix for 1960–2010 Relative to the United States	7
4	Transition Matrix Based on Maddison Data	7
5	Panel A. Pair-Wise Correlation Matrix for Low-Income Economies	14
5	Panel B. Pair-Wise Correlation Matrix for Middle-Income Economies	15
5	Panel C. Pair-Wise Correlation Matrix for High-Income Economies	16
6	Conditional Inference Forest Results with 1000 Trees in the Forest with $P < 0.05$	22
7	Regression-Tree-Simulated Transition Matrix	24

FIGURES

1	Income Transition from 1960 to 2011	3
2	Income Transition from 1960 to 2011 in Relative Terms to the United States	4
3	Box Whisker Plots for Variables	12
4	Conditional Inference Tree for Low-Income Economies	18
5	Conditional Inference Tree for Middle-Income Economies	19

ABSTRACT

Why do some economies grow faster than others? Do economies in the middle-income range face especially difficult challenges producing consistent growth? Using a transition matrix analysis on decade-level growth rates, we find that the data clearly reject the idea that middle-income economies either have a high absolute probability of being stuck where they are or have a higher relative probability of being stuck than the low- or high-income groups. In this sense, the notion of a “middle-income trap” is not supported by the data. However, economies in a given income range have different fundamentals and policies, and relative growth across economies may depend on these variables. Since development economists and practitioners have proposed a long list of variables that could affect growth, we employ a recently developed nonparametric classification technique (conditional inference tree and random forest) to decipher the relevance and relative importance of various growth determinants. We find that the list of variables that can help distinguish fast- and slow-growing economies is relatively short, and varies by income groups. For low-income economies, favorable demographics, macroeconomic stability, good education system, and good transport infrastructure appear to be the most important separating variables. For middle-income economies, favorable demographics, macroeconomic stability, sound global economic environment, and openness to foreign direct investment (FDI) appear to be the key discriminatory variables. This framework also yields conditions under which economies in the low- and middle-income range are trapped or even move backward.

Keywords: conditional inference regression tree, economic growth, growth determinants, infrastructure, middle-income trap

JEL Classification: C14, C30, F01, F60, O11, O43

I. INTRODUCTION

Are low-income economies likely stuck in a poverty trap? Is a typical economy in the middle-income group likely to be trapped in middle-income status forever, unable to attain a high absolute level of income? Perhaps more importantly, within any given income group, why do some economies grow faster than others? Are there clear and quantifiable indicators that will separate fast-growing economies from slow-growing ones? These are the questions that this paper will investigate.

The notion of middle-income trap has gained attention of policy makers and researchers. While lacking a formal definition, it may be thought of as stating that middle-income economies have a low probability of sustaining sufficiently high growth rates to join the high-income group. For example, Eichengreen, Park, and Shin (2013) documented that economic growth tends to slow down near two modes of \$10,000–\$11,000 and \$15,000–\$16,000 in 2005 purchasing power parity (PPP) terms. Robertson and Ye (2013) showed that a middle-income country's per capita income relative to the reference country tends to lie within a band. Aiyar et al. (2013) reported that a typical country in the middle-income group has a higher frequency of negative deviations from the growth path compared with countries in the other two groups.

If one follows a dictionary definition of a trap as a situation that is difficult or unable to exit once an economy gets in (Merriam-Webster Dictionary, and Oxford Dictionary), we will document in the first part of the paper that the data do not support any notion of an unconditional middle-income trap. That is, a middle-income economy that grows at an average or median rate of the middle-income group will clearly and surely become a high-income economy. In other words, in the data, a typical middle-income economy is not expected to be stuck or trapped in middle-income status. The same thing can be said about a typical low-income economy because the mean or the median growth rate is clearly positive. The only unconditional trap in the data is a “high-income trap.” That is, because the median or the mean growth rate of high-income economies is also positive, once an economy enters the high-income club, it is expected to stay there forever if it is to follow the mean or median growth rate of that group.

If we instead look at the chance that a middle-income economy catches up with the income level of a contemporaneous very rich economy, say the United States (US), we will find that the chance becomes less favorable. We will document that, in the steady state, there is a distribution of relative incomes: some economies will be income leaders, other economies will have lower relative incomes, and there will be no absolute convergence. Because the income level of the income leaders (say, the US) is a moving target, this pattern does not mean that a typical middle-income economy cannot grow beyond the income level that defines the ceiling of the middle-income group. An economy whose income is forever only 75% of the US can nevertheless grow very rich (as long as the US keeps growing as it has been doing in the past).

We are not the first people to discover these patterns in the literature. For example, Im and Rosenblatt (2013) employed transition matrices in the Maddison database over 1950–2008, and found no support for the notion of a middle-income trap in either absolute or relative terms. Felipe, Kumar, and Galope (2014) distinguished economies by fast and slow transitions from middle- to high-income group. They argued that the reason people talk about middle-income trap is because a very small group of countries made these transitions very fast, which are historical outliers rather than the norm. Bulman, Eden, and Nguyen (2014) documented the evidence against the existence of a middle-income trap. After examining the growth rate of countries across different income levels, they did not find stagnations at a particular income level.

Of course, economies in a given income group differ in their fundamentals and policy choices; so also would their growth performances differ. To understand the roles of these fundamentals and policies, in the second part, we introduce a nonparametric classification technique (conditional inference regression tree and random forest) and examine which proposed determinants are most relevant in separating fast- and slow-growing economies, and how these separating variables may differ across income groups. Based on the conditioning variables, we can classify all middle-income economies into three groups: progressive, near-stagnant, and regressive economies. In other words, we can now identify conditions under which middle-income economies can be trapped in their income status or even move backward. We find that demographics, infrastructure, macroeconomic management, and openness to FDI are especially important for growth. However, their relationship to growth is not linear. Different clusters of fundamentals and policy choices produce different growth performance. We can do the same for low-income economies using the same methodology.

Our paper differs from the existing literature in three important dimensions. First, instead of focusing solely on the unconditional income transition or economic growth slowdown (as did, for example, Eichengreen, Park, and Shin [2013]; Felipe, Kumar, and Galope [2014]; and Im and Rosenblatt [2013]), we examine what fundamentals and policies can separate fast- and slow-growing economies in a given income group. Second, as far as we know, this is the first attempt to employ a nonparametric classification scheme—regression tree and random forest—in analyzing economic growth. With this method, we can not only handle more than 20 variables, but we can also tolerate missing data and do not have to make assumptions about the distribution of random shocks. Third, rather than defining a “trap” or “slowdown,” or assuming that any incremental change in a given conditional variable always has the same effect on growth, we examine growth rates directly and let the data speak for itself on whether some of the effects are nonlinear or not.

It is useful to compare our paper with other related ones. Aiyar et al. (2013) examine the impacts of factors such as institutions, demographics, macroeconomic environment, and economic structure on economic slowdown. To deal with the small number of observations and large number of potential right-hand side variables, they use a probit model to include one set of right-hand side variables at a time, which seriously limits the credibility and generalization of their results. Rudengren, Rylander, and Casanova (2014) discuss the roles of governance, education, and other factors in economic growth. However, they only make some qualitative arguments without providing formal tests or analytical evidence. Using income relative to the US,¹ Jones (2015) calculated the long-run stable probability for each income group, which is consistent with our findings. Jones (2015) also touched upon the importance of institutions and governance on economic growth by relying on selected previous literature using a “natural experiment” environment. We will let institutions to be one of the fundamental variables in our regression analysis.

The rest of this paper is organized as follows: Section II discusses the unconditional economic transitions in the long run; Section III presents the evolving constraints analysis by regression tree and random forest; and Section IV concludes.

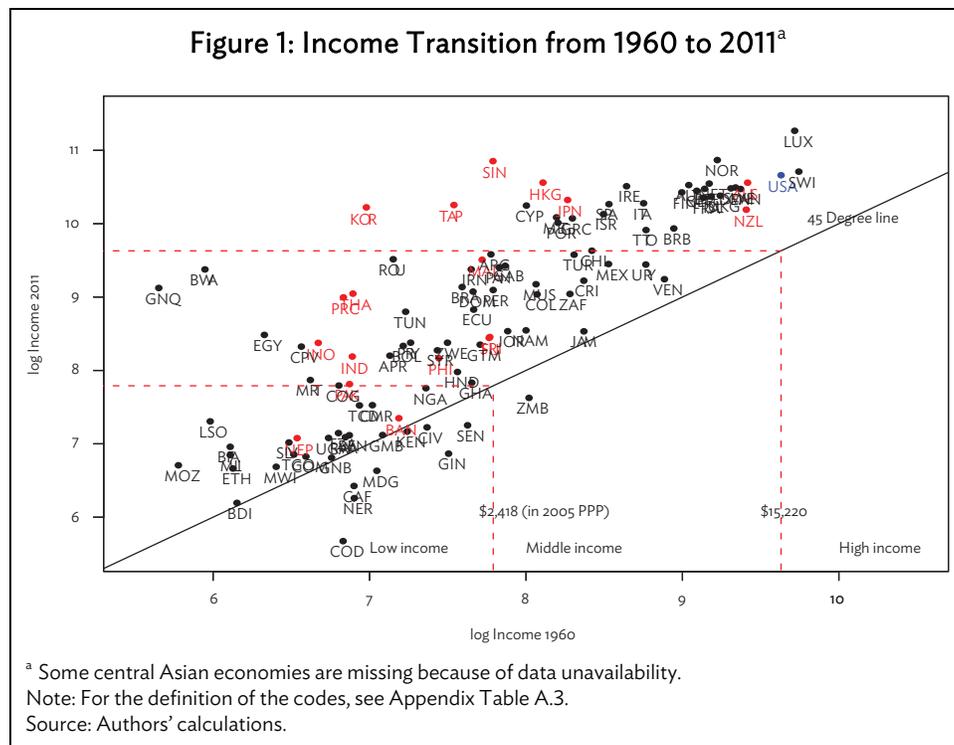
II. (UNCONDITIONAL) ECONOMIC TRANSITIONS IN THE LONG RUN

We measure income levels by real gross domestic product (GDP) per capita from the Penn World Table 8.0. We categorize all economies into five income groups: “Extremely Low Income” (ELI) with

¹ Less than 5%, between 5% and 10%, between 10% and 20% , between 20% and 40%, between 40% and 80%, and more than 80%.

real GDP per capita less than or equal to \$1,096; “Low Income” (LI) with real GDP per capita of \$1,096–\$2,418; “Lower-Middle Income” (LMI) with GDP per capita of \$2,418–\$5,550; “Upper-Middle Income” (UMI) with GDP per capita of \$5,550–\$15,220; and “High Income” (HI) with GDP per capita greater than or equal to \$15,220. The threshold of \$2,418 is equivalent to the World Bank’s cut-off line between low-income and middle-income economies. In addition, another category was added, extremely low-income economies, which comprises economies with per capita income below \$3/day in 2005 PPP or \$1,096/year in 2005 PPP terms. The income of the US in 1960 (\$15,220) was used as the threshold for classifying high-income economies.² Furthermore, the threshold for lower- and upper-middle-income economies was also calibrated so that there are about the same number of economies in the lower- and upper-middle-income categories in 1960, which resulted in a cutoff of \$5,550. Additional details on the mapping between our cut-off lines and the World Bank’s classification can be found in the Appendix.

In Figure 1, we plot log GDP per capita in 2011 against that in 1960. We impose the thresholds that separate middle-income from low-income economies, and high-income from middle-income economies. In terms of overall growth performance from 1960–2011, economies in a given income group fall into one of the following scenarios: (i) those below the 45 degree line, which experienced a negative growth rate; (ii) those above the 45 degree line but still belonging to the same income group; and (iii) those with a positive growth rate and have moved up to a higher income group. All economies that belonged to the middle-income group in 1960, except for Zambia, enjoyed positive growth rate with more than half of them moving up to achieve high-income status (27 out of 41) in 2011. The scenario for the low-income group is much worse: 63 economies in 1960 started as low-income economies, 29 remained as low-income economies in 2011, among which 8 economies experienced negative growth rate.

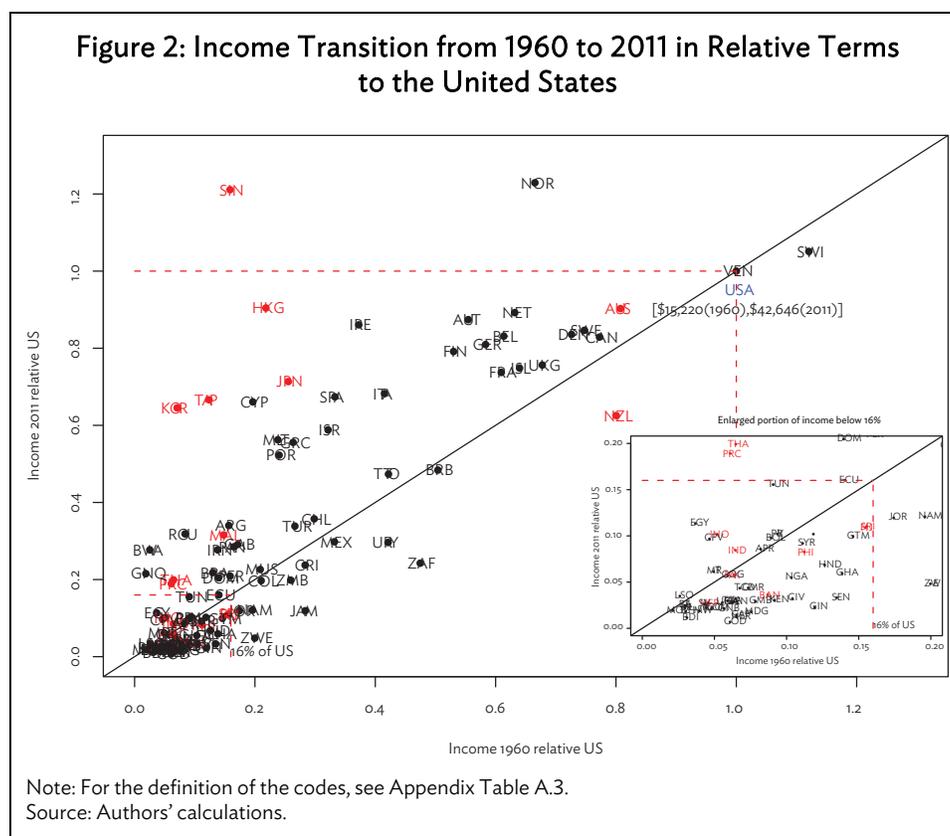


² This is lower than the World Bank’s threshold for high-income countries (\$19,429), since using the World Bank’s threshold would result in the US being classified only as a middle-income country in 1960.

All economies of the Asia and Pacific (in red dots) experienced positive growth rates, with a majority of them managing to move out of low-income status to at least the lower-middle-income group. The Republic of Korea; Singapore; and Taipei, China have burst past middle-income status and attained high-income status.

From Figure 1, the middle-income group does not exhibit any unconditional trap in the sense of nongrowth for a majority of economies. We probe it further by looking at the short-term transition using a shorter time span starting from 1980. A similar pattern is found. All middle-income economies enjoyed positive growth rates, while some of the low-income economies experienced negative growth rates. Since a majority of low-income economies also have positive growth, the unconditional probability of being trapped is also low.

After examining growth patterns in absolute terms, we turn to relative measures. As shown in Figure 2, the threshold for low-income economies is 16% of the 1960 US income level.³ The threshold for separating upper-middle-income and high-income economies is the 1960 US income level (100%).



The economies below the 45 degree line grew slower than the US. Compared to Figure 1, there is less catch up and more economies remain where they are in terms of their income relative to that of the US. As is well known, some Asian economies managed to move up to the higher income group even in relative terms.

³ \$2,418/\$15,220

A. Transition Matrix and Ergodic Distribution

We now investigate transition probabilities of different income groups by introducing the transition matrix and its asymptotic distribution or the Ergodic distribution.

We group economies by their per capita GDP at the beginning of a decade. There are five income groups: extremely low, low, lower middle, upper middle, and high. For each income group, we compute the probabilities that a typical economy moves to each of the possible income groups over a decade. These probabilities are summarized by a transition matrix in Table 1.⁴

Table 1: Decade-Average Transition Matrix for 1960–2010
(%)

	Extremely Low Income	Low Income	Lower- Middle Income	Upper- Middle Income	High Income
ELI	82	18	0	0	0
LI	3	72	25	0	0
LMI	0	3	68	29	0
UMI	0	0	0	70	30
HI	0	0	0	0	100
Ergodic distribution for the average decade transition matrix					
	0	0	0	0	100

ELI = extremely low income, LI = low income, LMI = lower-middle income, UMI = upper-middle income, HI = high income.
Source: Authors' calculations.

The number in a given cell reports the probability that a typical economy with an income status in the row moves to the income status in the corresponding column over a decade. For example, the first cell says that an extremely low-income economy has an 82% probability of remaining in the same income status after a decade, and the second cell says it has an 18% probability of becoming a low-income economy in a decade. The remaining cells in the first row indicate that there is zero probability of moving up any further in a decade. An economy that started as an upper-middle-income economy has a 70% probability of staying in the same income status and 30% probability of moving up as high-income economy at the end of the decade.

Based on the transition matrix, we can see that for all the nonhigh-income groups, the probability of moving up to a higher income level in one decade is greater than 15%. The following question would be, in the long run (allowing enough time to grow), whether all can end up in the high-income group eventually. To address this question, we employ the Ergodic distribution.⁵ As shown in the last row of Table 1, in the long run, regardless of development status from where economies begin, they will all end up in the high-income group (with probability of 1).⁶ In other words, in the long run, there is neither a low-income trap, nor a middle-income trap. The trap we can see in the data is a high-income trap. That is, once an economy reaches high-income status, it is expected to stay there forever.

⁴ The decade average transition matrix is estimated based on the 5-decade transition matrices from 1960 to 2010 by employing a numerical optimization program. Instead of taking the simple average for the five transition matrices (which suffers from Jensen's Inequality), we estimate a transition matrix that can give us an exact 5 decade duration transition matrix (entry in 1960 and exit in 2010) by taking its power 5.

⁵ Ergodic distribution matrix = transition matrix^{tr}. Empirically, we use power 2,000 to approximate the Ergodic distribution matrix.

⁶ We also check the robustness of the results by using a transition matrix with 5 decades as the duration (1960–2010). The result does not change.

The Ergodic distribution tells us the distribution of income status across economies over the very long run. But how long does it take to reach the very long run? From the transition matrix, we estimate that it will take 44 decades for all the extremely low- and low-income economies to move up to the next income level or higher, while it will take 48 decades for all economies to achieve either an upper-middle-income or a high-income status.

We can also compute, based on the transition matrix, the number of decades it takes for a given percentage (e.g., 50% or 90%) of economies in an income group to move out of their current status and into higher income groups.

We summarize the results in Table 2. For extremely low-income economies, it takes 4 decades for half of them to move to higher income groups. Similarly, for low-income, lower-middle-income, and upper-middle-income economies, it takes 3, 3, and 2 decades, respectively, for half of the economies to move to a higher income status.

Table 2: Decades Needed for X Percent of Economies to Move Up and Out of Their Current Group

	Extremely Low Income	Low Income	Lower- Middle Income	Upper- Middle Income
X=0.5	4	3	3	2
X=0.9	14	12	8	7

Source: Authors' calculations.

If we want to see 90% of the economies in a group move to higher incomes instead of 50%, naturally, the required durations would be longer. For the four developing economy groups from the extremely low-income to the upper-middle-income group, it takes 14, 12, 8, and 7 decades respectively, to move into the next income level or higher.

B. Ergodic Distribution Analysis on Convergence in Relative Terms to the United States

So far, we discussed the transitions based on absolute terms. Next, we assess the transition pattern relative to the US income level. We divide the groups into four categories: 16% of US real per capita income as low income, 16%–36% of US real per capita income as lower-middle income, 36%–75% of US income as upper-middle income, and 75% of US income and above as high income indicating catch up with the US.⁷

Table 3 presents the decade average transition matrix relative to US income from 1960 to 2010. For the low-income group, the probability of entering the lower-middle-income category relative to the US is 8%. The probability for an upper-middle-income economy to catch up with the high-income group is 22%. The last row of Table 3 shows the corresponding Ergodic distribution. The last column of the Ergodic distribution shows that 67% of economies cannot exceed 75% of US income in the long run. In relative terms to US income, the “middle-income trap” does exist.

⁷ The reason we have 16% as the cut-off line is to be consistent with the absolute analysis, in which, \$2,418 (the line to differentiate low-income and lower-middle-income in 1960) divided by \$15,220 (US income in 1960) is 0.16. The relative lower-middle-income line is 0.36 (dividing \$5,500, the line differentiating lower-middle- and upper-middle-income, by \$15,220). We choose 75% as the line to indicate a reasonable range with the US.

C. Long-Horizon Analysis with Maddison Data

In Maddison's data, GDP per capita of the US in 1990 international Geary-Khamis dollars for the year 1960 is \$11,328. Aligning with the absolute cut-off lines measured by 2005 PPP international dollars, we use 16%, 36%, and 100%⁸ of the US level as cut-off lines to calculate the cut-off line for income groups in 1990 international Geary-Khamis dollars. These correspond to the following categories: low income (less than \$1,812);⁹ lower-middle income (\$1,813–\$4,078); upper-middle income (\$4,079–\$11,327); and high income (\$11,328 and above).

Table 3: Decade-Average Transition Matrix for 1960–2010 Relative to the United States (%)

	16% and Below (Extremely Low and Low)	16%–36% (Lower-Middle Income)	36%–75% (Upper-Middle Income)	75% and Above (High)
16% and below	92	8	0	0
16%–36%	13	72	15	0
36%–75%	0	4	74	22
75% and above	0	2	19	79
Ergodic distribution associated with transition matrices relative to the United States				
	23	13	31	33

Source: Authors' calculations.

Table 4 shows the 50-year duration transition matrices for 1850–1900, 1900–1950, and 1950–2000. Compared with 1850–1900 and 1950–2000, in 1900–1950, the low-income group and lower-middle-income group had the highest probability of moving on to the next income level or higher. For the period of 1950–2000, the probability for lower-middle-income economies of moving to high income is 37%, while the probability for the upper-middle-income economies of achieving high-income status is 81%. The Ergodic distribution is consistent with the Ergodic distribution results using the Penn World Table 8.0 data. The probability for all income groups ending up in the high-income group is 100%.

Table 4: Transition Matrix Based on Maddison Data (%)

	Low	Lower-Middle Income	Upper-Middle Income	High
1850–1900 (in 1990 International Geary-Khamis dollars)				
	\$1–\$1,812	\$1,813–\$4,078	\$4,079–\$11,327	\$11,327 and above
Low	52	39	9	0
Lower middle	0	75	25	0
Upper middle	-	-	-	-
High	-	-	-	-
1900–1950				
Low	41	50	9	0
Lower middle	0	21	79	0
Upper middle	0	0	100	0
High	-	-	-	-

continued on next page

⁸ Of the US income used as the high income cut-off line, 100% is consistent with the analysis in absolute terms discussed in section 2.1 using the Penn World Table.

⁹ This corresponds to \$2,418 in the Penn World Table data.

Table 4 continued

	Low	Lower-Middle Income	Upper-Middle Income	High
1950–2000				
Low	59	24	15	2
Lower middle	0	26	37	37
Upper middle	0	0	19	81
High	0	0	0	100
Ergodic Distribution based on transition matrix of 1950–2000				
	0	0	0	100

Source: Authors' calculations.

III. EVOLVING CONSTRAINTS TO GROWTH: A PERSPECTIVE FROM REGRESSION TREES AND RANDOM FORESTS

One implied assumption for the Ergodic distribution is that the transition probability from one income status to another are the same for all economies within a given income group. However, for real growth progress, there is heterogeneity across economies. These dimensions of heterogeneity could be very interesting if they are systematically related to fundamentals or policy choices. In this section, we investigate factors affecting economic growth and their relative importance at different stages of development.

The extant growth literature suggests a long list of factors that have been hypothesized by researchers, policy makers, and practitioners as important factors for growth, especially for low-/middle-income economies. In this camp, there are several papers, such as the abovementioned Aiyar et al. (2013) and Rudengren, Rylander, and Casanova (2014) who made some qualitative arguments about the roles of governance, education, and other factors in economic growth without providing formal tests or analytical evidence. For the general categories of factors, in addition to the well-recognized factors recommended by existing literature, we particularly refer to the Asian Development Bank's Eight Key Actions for Development (Nakao 2014) and the Washington Consensus (Williamson 1990). When we did the variable selection, we also considered the availability of variables. Most of the variables we included can go back to 1960.

A. Variables that Could Alter Growth

We now discuss variables that may separate fast-growing and slow-growing economies. This list is guided by the vast existing literature on determinants of growth.

It is important to note that some plausible determinants of growth are not included in the examination here due to measurement issues. For example, political leaders' vision is identified by Nakao (2014) as one of the key growth determinants. However, we are not aware of a reliable data source that measures the quality of leaders' vision across countries and over time. As a result, we have to leave it out of the analysis here.

Initial income level is commonly accepted as a determinant of the growth rate, and is implied by the Solow growth model and confirmed by a long list of empirical literature (see a summary by Barro and Sala-i-Martin [2004]). The argument is that low-income countries which are farther away from the technology frontier defined by developed countries benefit from adopting the existing technology to improve their productivity and thus enjoy higher growth rates. As they experience productivity

advancements, they move closer to the technology frontier and are compelled to innovate rather than imitate technology, which is harder, and so their growth rates decline. Therefore, in our analysis, we expect economies with higher initial incomes to have lower growth rates, and those with lower initial incomes to have higher growth rates. This expectation is in line with findings in the literature. For example, Pritchett and Summers (2014) argued that there is a strong regression to mean trend in growth rates across countries. Real GDP per capita at the beginning of each decade is used as the initial income.

Demographics are considered basic driving factors of economic growth, as have been explained in growth theory. The contributions of population age structure come from two channels: higher labor supply and higher saving rates as pointed out by Bloom et al. (2007). Empirical evidence has likewise been documented by Bloom, Canning, and Malaney (2000); Bloom, Canning, and Sevilla (2003); and Bloom et al. (2007). We include the share of population 15–64 years old (labor force age population share) and the labor force population growth (difference between the natural logarithm transformed size of population aged 15–64 at the end and at the beginning of the decade) as the demographic variables. Data comes from the World Bank’s World Development Indicators (WDI).

Infrastructure is considered a key input in a country’s investment climate. When Prime Minister Modi of India and President Jokowi of Indonesia came to power in 2014, they both stressed investing in infrastructure as a key to lifting their respective countries’ growth rates. Straub (2008) suggests that infrastructure promotes growth directly through productivity improvements. Indirect channels include: labor productivity improvement by reducing time to commute, health and education improvement, economies of scale and scope, etc. The International Monetary Fund’s World Economic Outlook (2014) also found that increased public infrastructure investment raises output in both the short- and long-term, particularly during periods of economic slack and when investment efficiency is high. Following the recent trend of using direct measures of infrastructure development rather than infrastructure investment (see Egert, Kozluk, and Sutherland [2009] and Calderón, Moral-Benito, and Servén [2014]), we use the indicators developed by Calderón et al. (2014), which include: (i) electricity generating capacity in gigawatts per thousand workers, (ii) total length of paved roads in kilometers per thousand workers, and (iii) total length of rail in kilometers per thousand workers.

We use average years of total schooling from the Barro-Lee database (Barro and Lee 2013) to represent *human capital*. That better education is associated with higher growth is a common assertion, supported by two groups of evidence. The first group uses the estimation-based approach, and includes the following: Barro (1991), who documented a positive correlation between growth rate and initial human capital proxied by initial school-enrollment rates for 98 economies during the period 1960–1985; Mankiw, Romer, and Weil (1992), who used the percentage of working-age population that is in secondary school to approximate the rate of human-capital accumulation rate in an augmented Solow model and found a significant contribution from human capital to economic growth; and Benhabib and Spiegel (1994), who alternatively documented the positive contribution of human capital based on endogenous growth theory by modeling technological progress as a function of human capital. The other group is based on the calibration-based development accounting approach. For example, Caselli and Ciccone (2013) computed the increase in output that can be generated by more schooling which they interpret as an upper bound effect. To measure the attainment of education, we use achieved education years as the indicator. Limited by data availability, we did not consider education quality, although the recent literature pointed out that cognitive skill of the population is related to long-run economic growth (see, for example, Hanushek and Woessmann [2008]).

For *macroeconomic environment and policy*, we include inflation rate, government debt share, and the number of crisis episodes. The inflation rate is consumer price inflation from the World Bank’s World Development Indicators. Both adopted from Reinhart and Rogoff (2009), the government debt

share is the gross central government debt to GDP ratio and the total number of crisis is the sum of currency crises¹⁰ and bank crises¹¹ within the decade. We exclude inflation crises and external and local debt defaults from the Reinhart and Rogoff (2009) crisis data to avoid overlapping with the indicators of inflation and total government debt.

Economic openness is represented by the share of exports and imports to GDP (trade share) and the share of net FDI inflow to GDP (FDI share) from the World Bank's WDI. A vast literature confirms a positive association between trade openness and growth, but causality interpretation is more controversial (see Rodriguez and Rodrik [2001], Frankel and Romer [1999], Feyrer [2009]). There are several channels for FDI to affect growth, including: inducing a more educated workforce (Borensztein, de Gregorio, and Lee [1998]), improving trade openness (Balasubramanyam, Salisu, and Sapsford [1996]), and improving financial markets (Alfaro et al. [2003]). By including the share of net FDI inflow to GDP together with other variables, our framework provides an opportunity to revisit these debates.

The potential importance of *political institutions* in growth is summarized by Glaeser et al. (2004) as follows: with good political institutions (low expropriation risks) in place, there will be greater private sector incentives for investment in human capital and physical capital, which in turn contribute to growth. Well-known papers include Hall and Jones (1999), Acemoglu, Johnson, and Robinson (2001, 2002), Easterly and Levine (2003), Dollar and Kraay (2003), and Rodrik, Subramanian, and Trebbi (2002). Following this line, we adopt the political constraint indices used by Henisz (2000, 2002), which measures constraints on the executives (the president or the prime minister) from legislative, judicial, or other political bodies. The estimate ranges from 0 to 1, where zero means no political constraint (high political discretion) and it moves toward stricter political constraint as its value approaches one.

For *political stability*, Alesina et al. (1996) documented that in economies and time periods with a high propensity of government collapse (political instability), growth is significantly lower than otherwise. In our analysis, we choose the domestic conflicts indicator from Cross-National Time-Series (CNTS) Data Archive to represent political stability. The variable is a weighted conflict measure using the combination of domestic conflicts such as assassinations, strikes, guerrilla warfare, government crises, purges, riots, revolutions, and antigovernment demonstrations. Higher values of the indicator signal more political instability.

Inequality is considered by Nakao (2014) a potential drag on growth. The empirical literature on the relationship between inequality and growth produces mixed results. While Forbes (2000) found that, in a panel regression, a rise in inequality tends to be associated with a pickup in the subsequent growth rate, Deininger and Squire (1998) and Barro (2000) find that higher inequality retards growth in poor countries but encourages growth in rich countries. We include the income share held by the bottom 40% people from the WDI as a measure of inequality. The closer this share to 40%, the less inequality there is. This indicator measures the income share directly, easier to understand than the standard Gini coefficient. The income share held by the bottom 40% people has been chosen by the World Bank as an official measure of the degree of broadly shared prosperity since 2013 (World Bank 2013).

¹⁰ Currency crisis is defined as: currency crashes (an annual depreciation versus the US dollar [or the relevant anchor currency—historically the UK pound, the French franc, or the German DM, and presently the euro] of 15% or more); currency debasement (a reduction in the metallic content of coins in circulation of 5% or more or a currency reform where a new currency replaces a much-depreciated earlier currency in circulation).

¹¹ A banking crisis is defined as bank runs that lead to closure, merger, or takeover by the public sector of one or more financial institutions; and if there are no runs, the closure, merger, takeover, or large-scale government assistance of an important financial institution (or group of institutions), that marks the start of a string of similar outcomes for other financial institutions.

Additionally, we include two more control variables. One is a dummy variable for whether the economy is an “oil exporter.” We define an economy as an oil exporter (with value equal to 1) when its fuel exports exceed 40% of its total exports or its fuel exports exceed 15% of its GDP. Data come from the World Bank’s WDI. Out of 435 economy-decade combinations, 50 observations were labeled as oil exporters.

For *global economic environment*, we construct an indicator of global economic growth using US growth rate before 1980 and the population-weighted average growth rate of Japan, Germany, and the US for and after 1980. We have an average annual growth rate of 3.4% for the 1960s, 2.2% for the 1970s, 2.1% for the 1980s, 2.6% for the 1990s, and 0.7% for the 2000s.

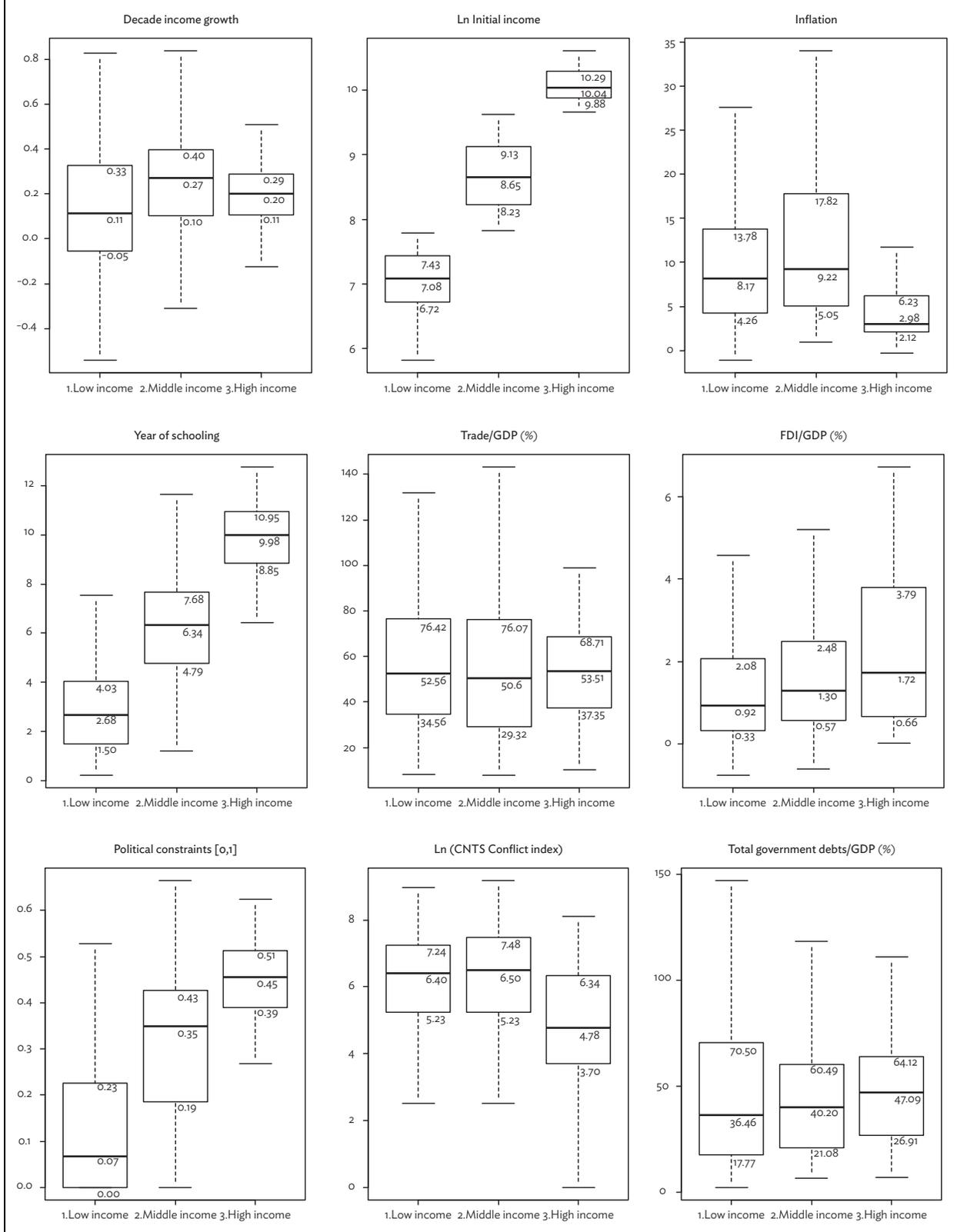
We use real GDP per capita from the Penn World Table 8.0 to measure economic growth. Most data are converted to decade average values, unless otherwise specified. The first decade is from 1960–1969. The annual growth rate is the compounded rate calculated based on the decade growth rate. We include economy-decades with at least 15 variables available (out of 17 potential predictors), which resulted in a total of 453 observations in the dataset. The dataset includes 94 economies, with 5 decades for some economies and less than 5 decades for others.

B. Box Whisker Plot and Pair-Wise Correlation Analysis

For each income group, we draw the Box Whisker Plots for each variable and present them in Figure 3. The middle-income economies have the highest decade median growth rate at 27%, whereas the low-income economies have the lowest decade median growth at 11%, but with the largest variation. There are clear strong associations between income levels (low/middle/high) and years of schooling, political constraints, electricity generating capacity, paved road, and railway, that is, higher levels of each factor are associated with higher income groups. For inflation, the median levels of the low-income and middle-income groups are close to each other, with a higher degree of variation among the middle-income group. For trade share, all three income groups share similar median levels, with the middle-income group having the largest variation. For FDI share, the high-income group has the highest median and largest variation. For domestic conflicts, the low-income and middle-income groups face relatively worse situations than the high-income group. For government debt share, the high-income group has the highest median level at around 47%, while the low-income group has the largest variation. For the total number of crises, the high-income group is in better situation. For labor force population growth, the low-income group has the highest decade growth rate at 25%, while the high-income group has the lowest decade growth at 7%.

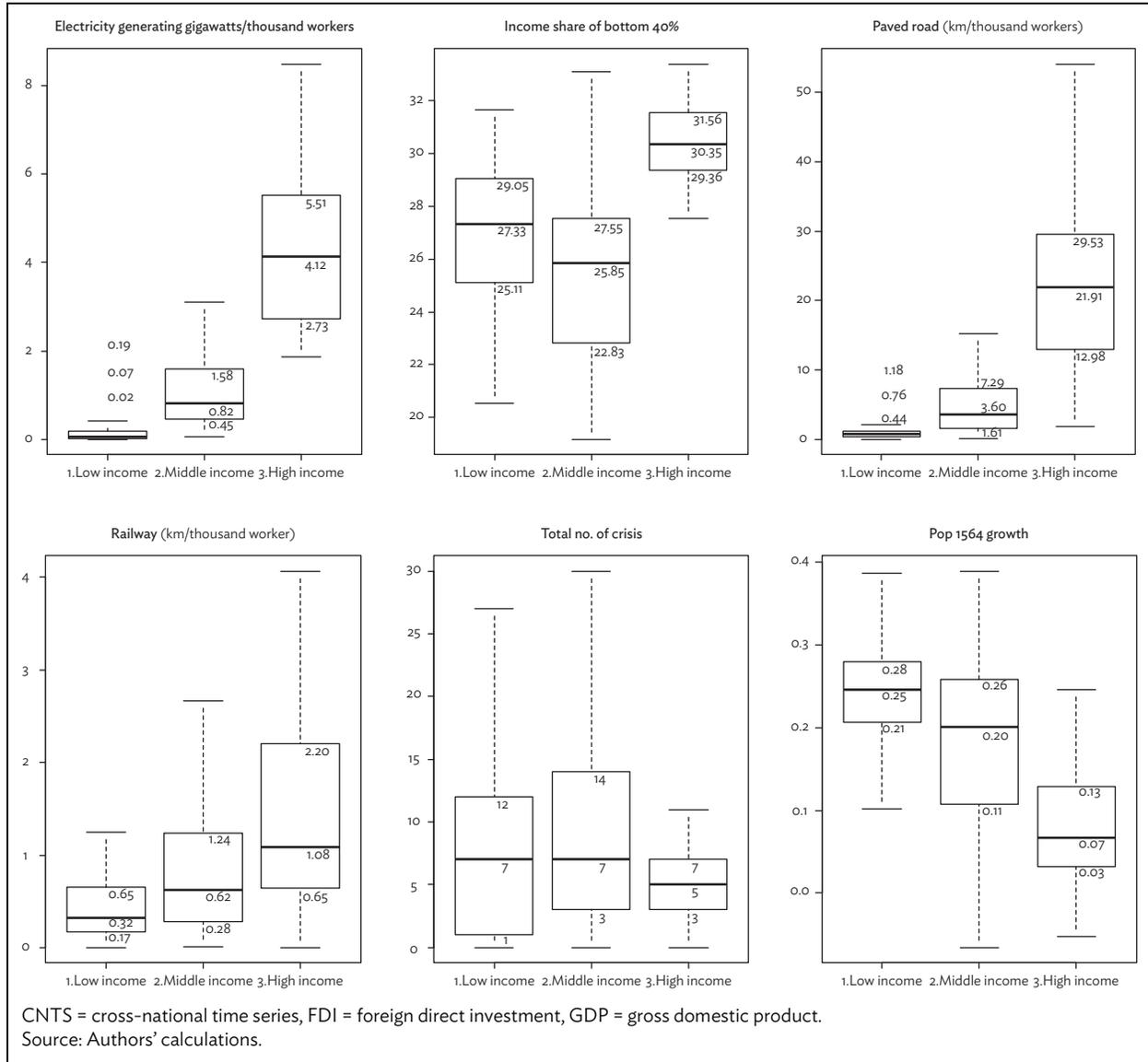
The Box Whisker Plots show heterogeneity in variables among different income groups, indirectly supporting our hypothesis that different subsets of factors matter more for growth among economies in different income groups. Based on this hypothesis, we construct the pair-wise correlation matrices for low-income, middle-income, and high-income groups. In Table 5, the red color highlights correlations between the growth rate and the factors (the first row), which are either higher than 0.15 or lower than -0.15 . The second through the last rows show the correlations between factors. The green color highlights correlations that are either higher than 0.4 or lower than -0.4 . As shown in red highlight, different variables are correlated with the growth rate for different income groups: (i) for the low-income group, years of schooling, political constraints, share of population 15–64 years old, government debt share, and paved roads have higher correlations with growth than other factors; (ii) for the middle-income group, FDI share, share of population 15–64 years old, government debt share, crisis indicator, and growth of working-age population have relatively higher correlations; and (iii) for the high-income group, initial income, trade share, share of population 15–64 years old, and crisis indicator correlate more with growth.

Figure 3: Box Whisker Plots for Variables



continued on next page

Figure 3 continued



The correlation observations show further evidence supporting our hypothesis that conditional on the stage of development, driving factors of economic growth vary among income groups. We also see that correlations between factors have different patterns across economies with different income levels (as shown in green color). The relatively high correlations between factors raise particular challenges for estimation, especially for linear regression analysis. In the following examination, we employ Conditional Inference Regression Tree and Random Forest approaches, which have the advantage of assessing the contribution of each factor conditional on the correlated predictors.

Table 5: Pair-Wise Correlation Matrix

Panel A. Low-Income Economies

	Initial Income	Inflation	Schooling	Trade	FDI	Political Constr.	Conflicts	Pop 15-64 Share	Debt	Power Gen.	Paved Road	Railway	Crisis	Pop 15-64 growth	Oil Exporter	Global Growth
Growth rate	(0.06)	(0.05)	0.23	(0.03)	0.03	0.20	0.02	0.28	(0.22)	0.06	0.17	0.00	(0.15)	0.06	0.14	(0.06)
Initial income		0.06	0.33	0.07	(0.05)	(0.03)	(0.01)	0.15	(0.15)	0.26	0.44	0.21	0.05	0.19	0.14	0.18
Inflation			0.14	(0.09)	(0.05)	0.08	0.08	0.04	0.18	0.01	(0.04)	0.15	0.16	(0.02)	0.21	0.01
Schooling				(0.01)	0.21	0.26	0.11	0.33	0.25	0.34	0.17	(0.05)	0.39	0.10	0.24	(0.34)
Trade					0.35	(0.09)	(0.33)	(0.07)	(0.13)	0.12	0.08	0.02	(0.36)	0.02	0.11	(0.09)
FDI						(0.02)	(0.09)	0.13	0.43	0.21	(0.02)	0.04	0.03	(0.08)	0.12	(0.32)
Political constr.							0.11	0.18	(0.03)	0.03	0.08	(0.13)	0.17	0.12	(0.07)	(0.26)
Conflicts								0.13	(0.17)	(0.05)	(0.02)	(0.02)	0.14	(0.10)	0.03	0.04
Pop 15-64 share									(0.04)	0.12	0.10	0.05	0.14	(0.39)	0.10	(0.06)
Debt										0.30	0.10	(0.16)	0.29	0.03	0.22	(0.30)
Power gen.											0.17	0.14	0.20	0.09	0.53	(0.06)
Paved road												0.16	(0.03)	0.02	0.01	0.03
Railway													(0.08)	(0.08)	(0.02)	0.25
Crisis														0.16	0.12	(0.06)
Pop 15-64 growth															0.08	(0.17)
Oil exporter																(0.09)

continued on next page

Table 5 continued

Panel B. Middle-Income Economies

	Initial							Pop						Pop		
	Income	Inflation	Schooling	Trade	FDI	Political	Conflicts	15-64	Debt	Power	Paved	Railway	Crisis	15-64	Oil	Global
						Constr.		Share		Gen.	Road			growth	Exporter	Growth
Growth rate	0.03	(0.11)	0.14	(0.06)	0.26	0.07	0.01	0.36	(0.34)	0.09	0.09	0.06	(0.22)	(0.24)	0.03	0.09
Initial income		(0.08)	0.48	(0.16)	0.14	0.32	(0.26)	0.49	(0.11)	0.61	0.46	0.35	0.06	(0.46)	0.05	0.01
Inflation			0.04	(0.06)	(0.07)	(0.03)	0.04	(0.07)	0.39	(0.07)	(0.06)	(0.03)	0.37	(0.10)	(0.06)	(0.10)
Schooling				0.06	0.4	0.30	(0.13)	0.58	0.28	0.46	0.34	0.23	0.05	(0.37)	(0.14)	(0.31)
Trade					0.41	(0.16)	(0.31)	(0.04)	0.35	(0.10)	(0.16)	(0.25)	(0.31)	0.3	0.16	(0.33)
FDI						0.08	(0.21)	0.37	0.17	0.03	0.01	(0.01)	(0.22)	(0.17)	0.09	(0.38)
Political constr.							(0.07)	0.24	(0.08)	0.30	0.19	0.08	(0.04)	(0.21)	(0.24)	0.11
Conflicts								(0.05)	(0.16)	(0.19)	(0.17)	(0.11)	0.26	0.16	(0.08)	0.00
Pop 15-64 share									(0.02)	0.32	0.15	0.10	0.06	(0.64)	(0.13)	(0.26)
Debt										(0.12)	0.07	(0.04)	0.00	0.04	0.05	(0.23)
Power gen.											0.42	0.41	0.00	(0.31)	(0.02)	0.06
Paved road												0.44	(0.14)	(0.34)	(0.14)	0.20
Railway													(0.05)	(0.23)	(0.12)	0.29
Crisis														(0.10)	0.00	0.03
Pop 15-64 growth															0.24	(0.01)
Oil exporter																(0.11)

continued on next page

Table 5 continued

Panel C. High-Income Economies

	Initial Income	Inflation	Schooling	Trade	FDI	Political Constr.	Conflicts	Pop 15-64 Share	Debt	Power Gen.	Paved Road	Railway	Crisis	Pop 15-64 growth	Oil Exporter
Growth rate	(0.35)	(0.17)	(0.04)	0.23	0.10	(0.12)	0.01	0.35	0.06	0.03	0.11	0.01	(0.28)	0.11	(0.15)
Initial income		(0.27)	0.51	0.25	0.40	0.06	(0.12)	0.19	0.09	0.32	0.11	(0.16)	0.24	(0.03)	0.16
Inflation			(0.29)	(0.17)	(0.27)	0.07	(0.06)	(0.40)	(0.23)	0.10	(0.16)	0.12	0.17	0.20	0.03
Schooling				(0.09)	0.11	0.16	0.04	(0.04)	(0.09)	0.28	0.28	0.18	0.08	0.21	(0.03)
Trade					0.83	(0.48)	(0.30)	0.45	0.21	(0.04)	(0.19)	(0.26)	(0.08)	0.39	0.33
FDI						(0.36)	(0.11)	0.36	0.09	(0.08)	0.09	(0.16)	0.03	0.43	0.17
Political constr.							(0.29)	(0.26)	(0.26)	0.14	0.06	0.01	(0.02)	(0.24)	(0.14)
Conflicts								(0.02)	0.39	(0.42)	(0.14)	(0.03)	0.25	0.05	(0.19)
Pop 15-64 share									0.30	(0.18)	(0.19)	(0.13)	(0.08)	(0.09)	(0.12)
Debt										(0.25)	(0.25)	(0.26)	0.14	(0.27)	(0.06)
Power gen.											0.25	0.35	0.01	0.01	0.48
Paved Road												0.36	0.02	0.03	(0.09)
Railway													(0.23)	0.26	(0.07)
Crisis														(0.22)	0.08
Pop 15-64 growth															0.12

() = negative, FDI = foreign direct investment.

Source: Authors' calculations.

C. Conditional Regression Tree Analysis

The regression tree analysis is a data driven machine learning method pioneered by Breiman et al. (1984) and refined in the subsequent literature. The general idea of a regression tree analysis (in the context of growth prediction) is this: the algorithm searches for all possible binary splitting points for each predictor (i.e., the independent variables we consider to affect economic growth), and chooses the one split point of the predictor that yields the highest gains in predicting growth and uses that particular predictor and splitting point to grow two children branches from the parent node. Following the same procedures, the algorithm searches and splits the children nodes until any further splitting does not yield any gain in improving predictability. In the final tree structure, the observations will end up in one of the ending nodes. The prediction of growth of each end node is simply the average of the growth of economy-decades falling into that node. Therefore, for prediction purposes, we can predict the economy with given predictors having the same expected growth as the average growth of the ending node where the economy belongs to.

Durlauf and Johnson (1995) were the first in applying a regression tree approach to economic growth. They consider only two conditioning variables: initial income and literacy rate. Their key point is that multiple growth regimes are a better characterization of the data than a single regime. The Conditional Inference Regression Tree, as suggested by Hothorn, Hornik, and Zeileis (2006), is a refinement of the Regression Tree Analysis that introduces hypothesis testing in deciding on each split—a split is made if one can reject the null hypothesis that the proposed split does not improve the predictive power. Because it makes a split of one predictor conditional on other correlated predictors, it overcomes criticisms of the traditional regression tree analysis that favors the choice of correlated predictors to do the splitting method. In the Conditional Inference Regression Tree, searching for the best predictor to make the split and searching for the optimal cut-off split value are conducted separately. First, based on linear statistics proposed by Strasser and Weber (1999), the relation of a variable to the response assessed by permutation tests follows a χ^2 distribution. The null hypothesis is that there is no association between a predictor and the response. With a smaller p-value, the probability of incorrectly rejecting the null hypothesis is lower. Therefore, in the first step, the variable with the smallest p-value is chosen to do the split. In the second step, the best cut-off point for the most significant variable chosen in step one is determined. For each of the two branches associated with the first split, another variable with the strongest association to the response is searched for. The remaining branches of the tree will grow in the same fashion. To grow the conditional inference tree, we require that all splits have p-values of 0.05 or smaller,¹² a minimum number of 7 observations for each ending node, and a minimum size of 20 in a branch before any split.

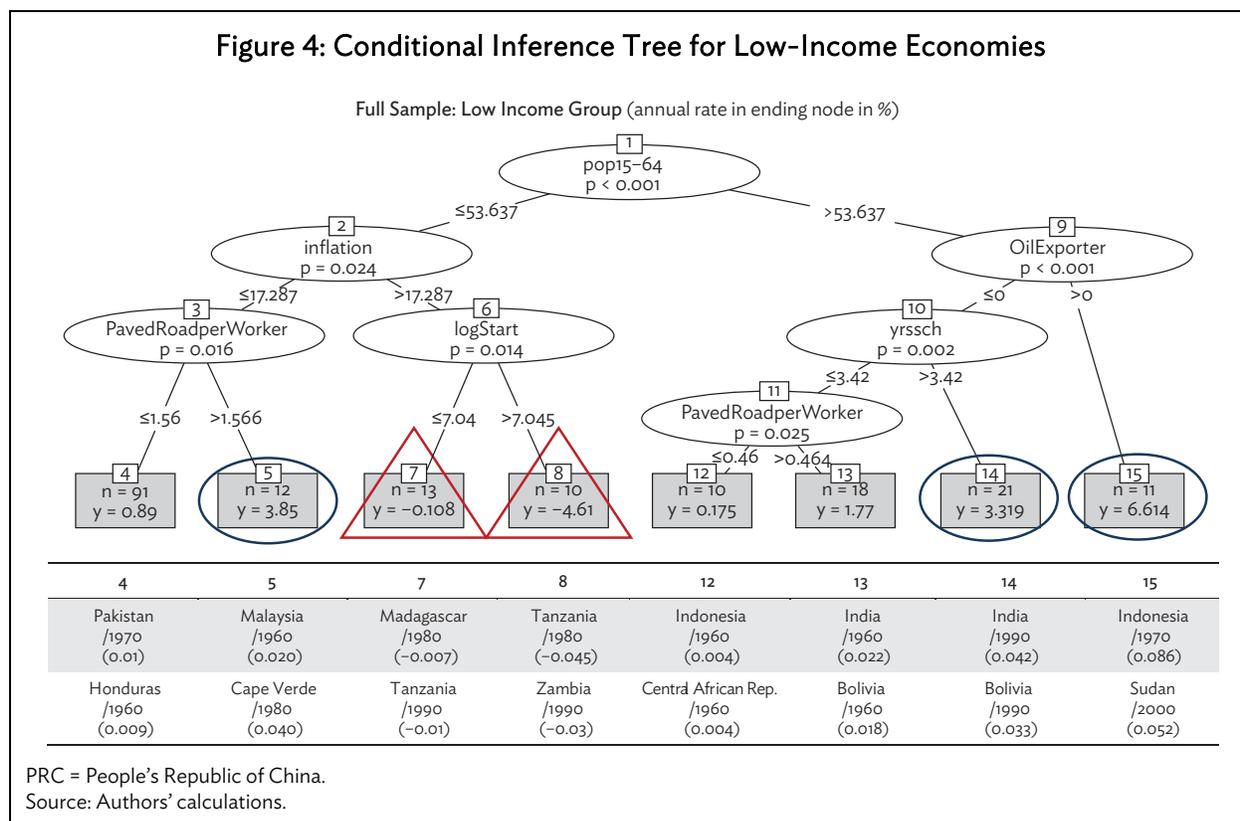
Figures 4 and 5 show the conditional regression trees for low-income and middle-income economies, respectively.¹³ In the tree, the variables used for each split and the associated p-values are labeled in each splitting node. For each split, the right branch indicates the branch with values higher than the splitting value of the parent node, while the left indicates the branch with values lower than the splitting value. The ending nodes are shaded with gray color. The number of observations and the predicted growth rates (average of growth rates) are listed. The predicted growth rates are average annual growth rates of economy-decades falling in each ending node.

¹² The estimate is carried out with the package party in R.

¹³ We skip the investigation of high-income countries so as to focus the analysis on low-income and middle-income countries.

1. For Low-Income Economies

We pool extremely low-income and low-income economies together and label them as one low-income group. As shown in the Conditional Inference Tree in Figure 4, among all the variables we included in the analysis, the important variables for categorizing their growth performance include: demographics (share of population 15–64 years old), macroeconomic environment (inflation), infrastructure (paved road), education (years of schooling), initial income level, and whether the economy is an oil exporter or not.



If there is a threshold effect, a virtue of the conditional tree approach is that it estimates the value of the threshold by hypothesis testing. In comparison, most papers in the existing literature such as Reinhart and Rogoff (2009) would preimpose a value of the threshold, which necessarily involve elements of arbitrariness.

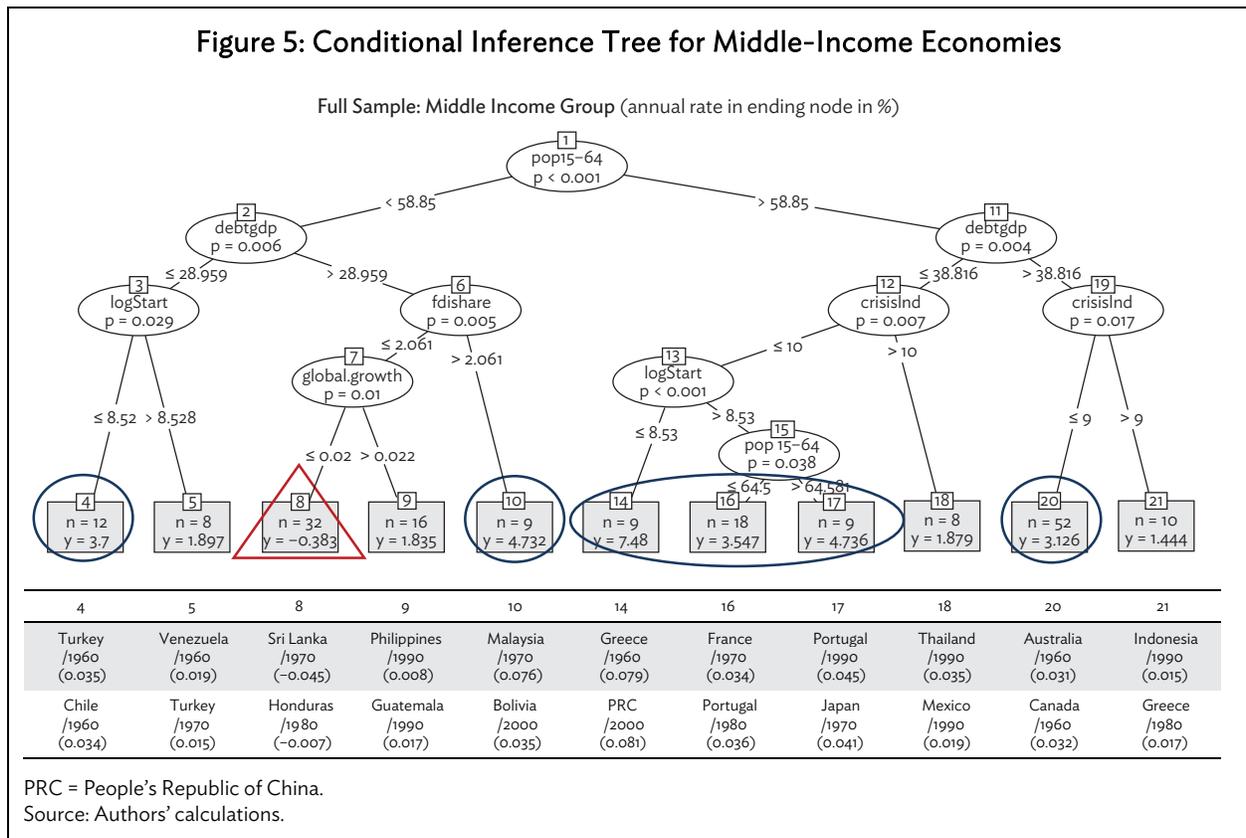
Based on the ending node results, we further categorize economies into three groups: progressive (with expected annual growth rate higher than 3%), near-stagnant (with expected growth rate between 0 and 3%), and regressive (with expected negative growth rate) economies. For progressive economies, three combinations of variables produce relatively high growth (labeled with blue circles): Conditional on favorable demographics (share of population 15–64 years old higher than 53.6%), if the economies are oil exporters, their expected annual growth rate is 6.6%; if not an oil exporter, but with relatively good education (years of schooling higher than 3.42), an annual growth of 3.3% can be expected. Another group of good performers with expected annual growth rate of 3.9% are economies with better macroeconomic environment (inflation lower than 17.3%) and sound infrastructure (paved road higher than 1.566 km per thousand workers) when facing unfavorable demographics (share of population 15–64 years old lower than 53.6%).

There are two groups with alarmingly negative expected growth rates (labeled with red triangles). Both are featured with unfavorable demographics (share of population 15–64 years old lower than 53.6%) and an unfavorable macroeconomic environment (inflation higher than 17.3%). When the logarithm transformed initial income is higher than 7.045 (\$1,147), the expected growth rate is more negative at -4.6%, than otherwise at -0.1%.

All other groups have growth rates between 0 and 3%. For all the ending nodes, we listed two sample economies with the decade and the actual growth rates in parentheses.

2. For Middle-Income Economies

For middle-income economies, we pooled the lower-middle- and upper-middle-income economies. As shown in Figure 5, the important variables for middle-income economies in explaining growth performance include: demographics (share of population 15–64 years old), macroeconomic environment (government debt to GDP ratio and the number of crises in the decade), openness (net FDI inflow as a share of GDP), global economic growth, and initial income level.



Based on growth performance, similar to the low-income group analysis, we categorize economies into three groups: progressive, near-stagnant, and regressive.

In the progressive group, economies with favorable demographics (share of population 15–64 years old higher than 58.9%), sound macroeconomic situation (government debt to GDP ratio lower than 38.8% and decade number of crises lower than 10), and lower initial income (lower than \$5,064) can expect an annual growth rate as high as 7.5%; economies under similar circumstances but with higher initial income (higher than \$5,064) can expect lower but still solid performance of either 3.5%

(if the share of population 15–64 years old is lower than 64.5% but higher than 58.9%) or 4.7% (if the share of population 15–64 years old is higher than 64.5%). A third group featuring favorable demographics (share of population 15–64 years old higher than 58.9%) but unfavorable government debt ratio (higher than 38.8%), as long as the decade number of crises is lower than 9, can still expect an annual growth rate of 3.1%.

For economies with unfavorable demographics (i.e., with a share of the 15–64 age cohort in the population lower than 58.9%), there are still two groups of economies that have reached a reasonable growth rate. They are on the left half of the graph. One of these groups, with a relatively low level of government debt (29% of GDP or less) and an initial income level of \$5,064 or less, has an annual growth rate of 3.7%. The other group, with their high debt-to-GDP ratios (in excess of 29% of GDP) offset by an open policy toward FDI (with FDI inflows at 2.06% of GDP or more), produces a growth rate of 4.7%. Since all the progressive groups grow faster than the average of the high-income group, they have hope of catching up with the existing high-income economies in due course.

Middle-income economies can also produce their regressive group. In particular, for a combination of unfavorable demographics (share of population 15–64 years lower than 58.85%), relatively high government debt (greater than 29% of GDP), low FDI openness (with a share of net FDI inflow to GDP lower than 2.061%), and unfavorable global economic environment (global annual growth lower than 2%), growth becomes $-0.4%$ a year. Of course, since these economies become poorer over time, they are doing worse than being trapped in a middle-income trap. If their policy choices and fundamentals do not change, in principle, they can slip out of the middle-income group and become low-income economies again.

Economies with other characteristics can have growth rates between 0 and 3%. While these economies are not formally trapped in a particular income status in terms of their absolute income, their anemic growth rate would leave them behind the existing high-income economies as a group in relative terms.

To summarize, for economies in the middle-income group to attain a strong growth rate (i.e., a growth rate higher than the high-income group), a favorable demographic pattern (a high share of working-age population) and prudent macro debt management are helpful. Without a favorable demographic pattern, a combination of prudent macro debt management and openness to FDI can still deliver strong growth. In contrast, macroeconomic instability in the form of frequent crises and inadequate openness to FDI are likely to lead to anemic or even negative growth rates.

D. Robustness Check with Random Forest Analysis

As a nonparametric technique, relative to linear regression analysis, regression tree analysis enjoys several advantages: no required transformation of variables, robustness to outliers, and greater tolerance of missing data without having to impute values. However, results of the regression tree are potentially sensitive to changes in the sample (Shmueli, Patel, and Bruce [2007], page 132). To obtain a sense of the results in different subsamples, a random forest technique is proposed and used by Breiman (2001) and Hapfelmeier (2012). A random forest is a combination of many trees, with each tree constructed on the basis of an independently and randomly drawn subsample and subject to random errors. Therefore, as the number of trees in the forest increases, the random errors are averaged out by taking the average of the trees in the forest, helping to yield more robust results compared with a single tree based on the whole sample. Since the size of the subsample for each tree in the forest is smaller than the whole sample, the forest does not include the particular tree that was constructed based on the whole sample and presented earlier.

For each income group, we will grow a forest with 1,000 trees¹⁴ (based on 1,000 randomly drawn subsamples). In defining the parameters to grow the trees, we choose to use the unbiased random forest as suggested by Strobl et al. (2007).¹⁵ For each tree, we require the maximum p-value for a split to be 5%, the minimum size for a split to be 20, the minimum size for the ending node to be 7, and the resample size of 90% as the whole sample.

Unlike a regression tree, the results of a random forest are harder to visualize and are summarized in Table 6. The first column ranks the importance of factors based on the frequencies listed in column 3. The frequency pertains to the total number of appearances of each variable in all trees in the forest. The fourth column is the average split value of the corresponding variables. For example, the share of population 15–64 years old appears 1,277 times in the forest and the average of its split value across all its 1,277 appearances is 53.77%. As illustrated in the regression tree analysis, for each split, the right branch includes observations with values higher than the split value, while those with lower values are on the left branch. Column 5 lists the average difference of the decade growth rates between observations on the right branch and those on the left branch when the corresponding variable is used for the split. Therefore, if the difference is a positive number, the variable used for the split has a positive association with the growth rate. Using the share of population 15–64 years old as an example, we can say that on average, economies with a share higher than 53.77% has an annual growth rate around 2.54% higher than that of economies with a share lower than 53.77%. The last column is a statistic we constructed to indicate the significance of the results in column 5. They are the frequencies of positive differences against the frequencies of negative differences. With larger differences in the positive–negative votes, we have higher confidence in the results listed in column 5.

We highlight all variables with a total frequency of 600 (out of 1000) or higher. For low-income economies, the share of population 15–64 years old, paved roads, share of net FDI inflow to GDP, power generating capacity, initial income, population growth, whether an economy is an oil exporter or not, years of schooling, inflation, and government debt share are the important variables. The variables shown in the regression tree for the whole sample such as favorable demographics and openness to FDI are all picked up as important variables by the forest, which suggests robustness of these variables.

In addition, several measures of infrastructure, especially roads and power generation, are often important in subsamples.

One notable difference of the forest results from the regression tree results is the high ranking of FDI share, power generating capacity, and population growth. The difference between the right and left branches of each split in the regression tree is conditional on the unique tree structure that was constructed based on the whole sample. By contrast, the contribution of each variable (Table 6, column 5) in the forest is the average of the contributions of all splits using that variable conditional on the tree structures across the forest. Conditional on the tree structures in the forest, we consider their contribution in the forest as the “marginal” effect of that variable on growth. The regression tree and the random forest differ in serving policy purposes. For “diagnostic” types of purposes, such as determining what institutional/fundamental combination can help one economy improve its growth (moving from the regressive group to the progressive group), the regression tree is better. For prioritization purposes, the forest is better since it ranks the importance (marginal effect) of each variable.

¹⁴ The function of `cforest` in the package of `party` in R is employed.

¹⁵ With `teststat="quad"`, `testtype="Univariate"` defined in the `cforest_control`.

Table 6: Conditional Inference Forest Results with 1,000 Trees in the Forest with $P < 0.05$

Rank (1)	Factors (2)	Total Frequency (3)	Average Split Value (4)	Difference of Growth between Nodes with Higher / Lower Value than the Splitting Value (decade growth rate) (5)	Frequencies of Positive Contributions / Negative Contributions (%) (6)
Low Income					
1	Population share (15–64)	1,277	53.77%	2.54	93.89/6.11
2	Paved road (km/1,000 workers)	1,181	1.39	1.21	69.18/30.82
3	FDI share	955	1.83%	0.79	65.03/34.97
4	Power generating capacity (gigawatts/1,000 workers)	839	0.12	1.54	78.67/21.33
5	Log initial income	832	6.86	(2.47)	9.62/90.38
6	Population growth (15–64)	828	0.20	2.37	85.75/14.25
7	Oil exporter or not	740	0.00	2.36	78.65/21.35
8	Years of schooling	736	3.21	2.02	89.95/10.05
9	Inflation	617	12.83%	(2.21)	15.24/84.76
10	Government debt share	603	41.82%	(1.72)	5.14/94.86
11	Global growth rate	571	2.00%	0.84	71.98/28.02
12	Log conflicts index	507	5.50	(1.12)	30.97/69.03
13	Political constraints	480	0.12	1.62	78.96/21.04
14	Railway (km/1,000 workers)	294	0.39	0.29	53.06/46.94
15	Trade share	282	60.44%	(0.92)	20.21/79.79
16	No. of crises	278	3.00	(1.39)	24.82/75.18
17	Income share held by bottom 40% income	124	27.00%	0.77	63.71/36.29
Middle Income					
1	Population share (15–64)	2,093	59.03%	2.38	96.42/3.58
2	Government debt share	1,690	38.29%	(2.01)	2.9/97.1
3	No. of crisis	1,414	8.00	(2.02)	4.03/95.97
4	Log initial income	1,152	8.51	(2.37)	7.12/92.88
5	FDI share	1,036	1.84%	2.25	91.89/8.11
6	Global growth rate	974	2.00%	2.04	97.33/2.67
7	Political constraints	721	0.28	(0.15)	51.87/48.13
8	Years of schooling	606	6.41	0.22	45.87/54.13
9	Inflation	603	13.98%	(0.64)	32.01/67.99
10	Population growth (15–64)	505	0.18	(0.98)	22.57/77.43
11	Log conflicts index	488	5.75	0.58	59.84/40.16
12	Oil exporter or not	458	0.00	1.92	89.96/10.04
13	Trade share	422	62.75%	(1.22)	27.49/72.51
14	Power generating capacity (gigawatts/1,000 workers)	286	0.80	(0.46)	41.96/58.04
15	Paved road (km/1,000 workers)	256	4.32	1.09	75.39/24.61
16	Railway (km/1,000 workers)	246	0.69	1.19	68.7/31.3
17	Income share held by bottom 40% income	98	25.35%	0.47	54.08/45.92

() = negative.

Source: Authors' calculations.

In our analysis, although the structure of the regression tree is potentially sensitive to changes in the sample as noted by Shmueli, Patel, and Bruce (2007), the resulting economy groupings in the end nodes are quite stable. Our intuition is that when the factors are close competitors in explaining the growth differentiation between groups (making the splits), the structure of the tree is more sensitive to data changes. For example, when factor A and factor B are equally good in doing the split, either one can be chosen to do the split. While the structure of the tree would be different (since one tree would have factor A and the other tree would have factor B), the economy groupings in the end

nodes would be exactly the same. This may imply a change in perspective in the application of the regression tree. That is, as long as factor A can represent factor B (or vice versa), it does not matter much which one is on the tree to do the split.

For middle-income economies, variables with frequencies higher than 600 include the share of population 15–64 years old, government debt share, the number of crises in a decade, initial income, share of net FDI inflow to GDP, global growth, political constraints, years of schooling, and inflation. However, for political constraints and years of schooling, the entries in the last column of frequencies of positive contributions/negative contributions are 51.87%/48.13% and 45.87%/54.13%, respectively, which indicate no dominating votes, so we dropped them from the list. Again, the variables picked up by the forest cover all variables shown in the regression tree, which suggests that our regression tree results are robust.

We also check the robustness of the random forest results by carrying out estimations with subsamples, excluding decades of 1970–1980, 1980–1990, 1990–2000, and 2000–2010, one decade at a time. We compare the important variables in each subsample with the top 10 important variables for low-income economies: among the top 10 variables (with frequencies higher than 600) listed in Table 6, eight appear in the top 10 list for the subsample excluding 1970–1980 (years of schooling and government debt share are excluded), nine appear in the top 10 list for the subsample excluding 1980–1990 (net FDI inflow is excluded), eight appear in the top 10 list for the subsample excluding 1990–2000 (population growth and government debt share are excluded), and nine appear in the top 10 list for the subsample excluding 2000–2010 (the oil exporter dummy is excluded).

Among the top nine variables (with frequencies higher than 600) for middle-income economies listed in Table 6, eight appear in the top nine list for the subsample excluding 1970–1980 (years of schooling is excluded), eight appear in the top nine list for the subsample excluding 1980–1990 (global growth is excluded), eight appear in the top nine list for the subsample excluding 1990–2000 (years of schooling is excluded), and eight appear in the top nine list for the subsample excluding 2000–2010 (political constraints is excluded).

Although the results are not based on randomly drawn subsamples (such as by excluding economy decades randomly), they still to some extent lend confidence to the robustness of our forest results. The small variations in the subsamples may be reflective of the decade-specific features of the growth patterns.

Another robustness check is the use of initial values for all variables at the beginning of each decade, rather than their decade average values to help us address the “endogeneity” challenge. We conducted the exercise and obtained forest results that are similar to the results listed in Table 6.

IV. CONCLUSION: LINK THE CONDITIONAL AND UNCONDITIONAL ANALYSES

In this paper, we examine the growth performance of economies in different income status. In the first half of the paper, we reject the unconditional notion of a “middle-income trap,” or a “low-income trap.” That is, an average economy in either the low- or middle-income group has more than 50% chance of having a positive growth rate. Therefore, given enough time, an average economy is always expected to move to a higher income status. The only trap in the data is a high-income trap in the sense that once an economy enters the high-income club, it is always expected to stay there. In the second half of the paper, we find that a relatively succinct list of variables can separate fast-growing and slow-growing economies in any given income group.

We now link the conditional results based on the regression trees to the unconditional results based on transition matrices. We divide the economies into five groups: extremely low-income, low-income, lower-middle-income, upper-middle-income, and high-income economies using the same criteria as in Section II. In each group, we have three types of economies: progressive (with an expected annual growth rate higher than 3% based on the regression tree), near-stagnant (with an annual growth rate between 0 and 3%), and regressive (with a negative annual growth rate). The results are presented in Table 7.

Table 7: Regression-Tree-Simulated Transition Matrix
(%)

	Obs. in the Group / Total Obs.	Obs. Percentage	Extremely Low Income	Low Income	Lower- Middle Income	Upper- Middle Income	High Income	# Years for 50% Moving Up	# Years for 90% Moving Up
Extremely Low-Income Economies									
Progressive	12/82	14.63	17	83	0	0	0	4	13
Near-stagnant	60/82	73.17	85	15	0	0	0	43	142
Regressive	10/82	12.20	100	0	0	0	0	never	never
Low-Income Economies									
Progressive	32/104	30.77	0	38	62	0	0	8	24
Near-stagnant	59/104	56.73	0	92	8	0	0	79	261
Regressive	13/104	12.50	54	46	0	0	0	never	never
Lower-Middle-Income Economies									
Progressive	43/86	50.00	0	0	44	56	0	9	29
Near-stagnant	19/86	22.09	0	0	95	5	0	129	426
Regressive	24/86	27.91	0	0	100	0	0	never	never
Upper-Middle-Income Economies									
Progressive	66/97	68.04	0	0	0	65	35	17	54
Near-stagnant	23/97	23.71	0	0	0	96	4	156	518
Regressive	8/97	8.25	0	0	12	88	0	never	never

Source: Authors' calculations.

Conditional on the sample and the regression tree results, we show that for extremely low-income groups, 12 out of 82 economy-decades belong to progressive economies, and they have an 83% probability of moving up to the next higher income group—i.e., the low-income group—within 1 decade. It only takes 4 years for half of them to move up to the next higher income group or 13 years for 90% of them to move up. For these economies, there is clearly no low-income trap. For the near-stagnant economies (60 out of 82), the scenario is much worse; the upward decade transition probability is only 15%. It will take 43 years (142 years) for 50% (90%) of them to move up to higher income groups. For the regressive group (10 out of 82), i.e., those with negative growth rates, they will never move up to higher income groups if nothing else changes. With policy choices and fundamentals that characterize the regressive group of low-income economies (i.e., high inflation and unfavorable demographics), these economies are likely trapped in poverty.

We perform a similar exercise for the other three income groups. In general, there is no trap for economies in a progressive group. They are expected to move to the next income group within a relatively short period of time. For economies in a regressive group, the negative expected growth rate implies that they may do worse than being simply trapped in their current income status. For

economies in a near-stagnant group, because growth is low, they may look like they are being trapped in their current income status for a long time. One interesting observation is that even for the progressive economies, it takes longer for the upper-middle-income economies to move up compared with the other income groups since the income interval covered by the upper-middle-income group is much wider than the other groups. For example, it takes 54 years for 90% of the upper-middle-income economies to join the high-income club, but only 29 years for 90% of the lower-middle-income economies to move up, and 24 years for 90% of the low-income economies to move up. (In other words, part of the differences in the time it takes to move up are due to the income thresholds one chooses for the income groups.)

Based on what characterizes a progressive group in a given income group, one can also infer the types of changes in policies (and fundamentals) that can help hasten the pace of progress toward high-income status. The regression tree results therefore provide plausible drivers of growth for economies in a given income group. For a given economy, comparing its own policy regimes and fundamentals to these growth drivers provide hints for plausible priority reform items.

APPENDIX

Linkage between our income group classifications and the World Bank's classifications

The World Bank classifies economies according to the following thresholds in 2013 US dollars (Atlas method):

- Low-income economies (L): gross national income (GNI) per capita (Atlas method) \leq \$1,045
- Lower-middle-income economies (LM): $\$1,045 < \text{GNI per capita (Atlas method)} \leq \$4,125$
- Upper-middle-income economies (UM): $\$4,125 < \text{GNI per capita (Atlas method)} < \$12,746$
- High-income economies (H): GNI per capita (Atlas method) \geq \$12,746

We use data on GDP per capita in 2005 purchasing power parity (PPP) terms from the Penn World Tables 8.0. To make the World Bank thresholds, which are in GNI per capita (Atlas method) terms, compatible with our data in 2005 PPP, we use the ratios of the average GNI in Atlas method for 2013 to that in 2005 PPP per economy group (i.e., L, LM, UM, H) and apply them to the thresholds in GNI Atlas method to get the equivalent thresholds in 2005 PPP.

Table A.1: Group Averages in Different Units

	Atlas Method (2013 US dollars)	PPP (2005 international dollars)
L	664	1,536
LM	2,068	5,152
UM	7,540	11,494
H	39,312	34,800

H = high, L = low, LM = lower middle, PPP = purchasing power parity, UM = upper middle, US = United States.
 Sources: World Bank country classification; World Bank, World Development Indicators. <http://data.worldbank.org/data-catalog/world-development-indicators> (accessed 29 August 2014).

The resulting thresholds in 2005 PPP are as follows:

- L: GNI per capita (2005 PPP) \leq \$2,418
- LM: $\$2,418 < \text{GNI per capita (2005 PPP)} \leq \$10,276$
- UM: $\$10,276 < \text{GNI per capita (2005 PPP)} < \$19,429$
- H: GNI per capita (2005 PPP) \geq \$19,429

However, using the revised thresholds, the US would be classified only as a middle-income economy in 1960. As this appears to be too strict, an adjustment was effected to make US income as the threshold for classifying high-income economies. In addition, another category was included, extremely low-income economies, which comprise economies with per capita income below \$3/day in 2005 PPP or \$1,096/year in 2005 PPP terms. Furthermore, the threshold for lower- and upper-middle-income economies was also calibrated so that there are about the same number of economies in the lower- and upper-middle-income categories in 1960.

The final thresholds used are as follows:

- Extremely low-income economies: GNI per capita (2005 PPP) \leq \$1,096
- Low-income economies: $\$1,096 < \text{GNI per capita (2005 PPP)} \leq \$2,418$
- Lower-middle-income economies: $\$2,418 < \text{GNI per capita (2005 PPP)} \leq \$5,500$
- Upper-middle-income economies: $\$5,500 < \text{GNI per capita (2005 PPP)} < \$15,220$
- High-income economies: GNI per capita (2005 PPP) \geq \$15,220

Table A.2: Factors Included in the Regression Tree Analysis

Indicators	Included in ADB's Eight Key Actions	Included in Washington Consensus	Other Important Factors for Development
1. Political stability: Domestic conflicts indicator from Cross-National Time-Series database;	✓	✓	
2. Macroeconomic stability: inflation, government debt share, and the number of crisis episodes (currency, banking, and stock market crises);	✓	✓	
3. Investment in infrastructure: paved road, railway, and power generating capacity;	✓	✓	
4. Investment in human capital: average years of schooling from Barro-Lee database;	✓	✓	
5. An open trade and investment regime: trade share and share of net FDI inflow in GDP;	✓	✓	✓
6. Good governance: political constraints indices by Henisz (2002);			✓
7. Inequality: share of population with incomes belonging to the bottom 40%			✓
8. Initial income level;			✓
9. Demography: share of population aged 15–64 and changes of log (population aged 15–64);			
10. Global economic environment: growth rate of the leading economies; and			
11. Oil exporter indicator.			

FDI = foreign direct investment, GDP = gross domestic product.

Note: In the eight key actions of ADB, one action is not covered in our analysis—a clear vision for the future since it is hard to find an appropriate indicator.

Sources: Nakao (2014); Williamson (1990).

Table A.3: Economy Codes

Code	Economy	Code	Economy	Code	Economy
AGO	Angola	BGR	Bulgaria	CAF	Central African Republic
ALB	Albania	BHR	Bahrain	CAN	Canada
ARG	Argentina	BHS	Bahamas, The	SWI	Switzerland
ARM	Armenia	BIH	Bosnia and Herzegovina	CHL	Chile
ATG	Antigua and Barbuda	BLR	Belarus	PRC	China, People's Rep. of
AUS	Australia	BLZ	Belize	CIV	Cote d'Ivoire
AUT	Austria	BMU	Bermuda	CMR	Cameroon
AZE	Azerbaijan	BOL	Bolivia	COD	Congo, Dem. Rep.
BDI	Burundi	BRA	Brazil	COG	Congo, Rep.
BEL	Belgium	BRB	Barbados	COL	Colombia
BEN	Benin	BRU	Brunei Darussalam	COM	Comoros
BFA	Burkina Faso	BHU	Bhutan	CPV	Cape Verde
BAN	Bangladesh	BWA	Botswana	CRI	Costa Rica

continued on next page

Table A.3 continued

Code	Economy	Code	Economy	Code	Economy
CYP	Cyprus	KGZ	Kyrgyz Republic	POR	Portugal
CZE	Czech Republic	CAM	Cambodia	PRY	Paraguay
GER	Germany	KNA	St. Kitts and Nevis	QAT	Qatar
DJI	Djibouti	KOR	Korea, Rep. of	ROU	Romania
DMA	Dominica	KWT	Kuwait	RUS	Russian Federation
DEN	Denmark	LAO	Lao People's Democratic Republic	RWA	Rwanda
DOM	Dominican Republic	LBN	Lebanon	SAU	Saudi Arabia
ECU	Ecuador	LBR	Liberia	SDN	Sudan
EGY	Egypt, Arab Rep.	LCA	St. Lucia	SEN	Senegal
SPA	Spain	SRI	Sri Lanka	SIN	Singapore
EST	Estonia	LSO	Lesotho	SLE	Sierra Leone
ETH	Ethiopia	LTU	Lithuania	SLV	El Salvador
FIN	Finland	LUX	Luxembourg	SRB	Serbia
FIJ	Fiji	LVA	Latvia	STP	Sao Tome and Principe
FRA	France	MAC	Macau, China	SUR	Suriname
GAB	Gabon	APR	Morocco	SVK	Slovak Republic
UKG	United Kingdom	MDA	Moldova	SVN	Slovenia
GEO	Georgia	MDG	Madagascar	SWE	Sweden
GHA	Ghana	MLD	Maldives	SWZ	Swaziland
GIN	Guinea	MEX	Mexico	SYR	Syrian Arab Republic
GMB	Gambia, The	MKD	Macedonia, FYR	TCD	Chad
GNB	Guinea-Bissau	MLI	Mali	TGO	Togo
GNQ	Equatorial Guinea	MLT	Malta	THA	Thailand
GRC	Greece	MNE	Montenegro	TAJ	Tajikistan
GRD	Grenada	MON	Mongolia	TKM	Turkmenistan
GTM	Guatemala	MOZ	Mozambique	TTO	Trinidad and Tobago
HKG	Hong Kong, China	MRT	Mauritania	TUN	Tunisia
HND	Honduras	MUS	Mauritius	TUR	Turkey
HRV	Croatia	MWI	Malawi	TAP	Taipei, China
HUN	Hungary	MAL	Malaysia	TZA	Tanzania
INO	Indonesia	NAM	Namibia	UGA	Uganda
IND	India	NER	Niger	UKR	Ukraine
IRE	Ireland	NGA	Nigeria	URY	Uruguay
IRN	Iran, Islamic Rep.	NET	Netherlands	USA	United States
IRQ	Iraq	NOR	Norway	UZB	Uzbekistan
ISL	Iceland	NEP	Nepal	VCT	St. Vincent and the Grenadines
ISR	Israel	NZL	New Zealand	VEN	Venezuela, RB
ITA	Italy	OMN	Oman	VIE	Viet Nam
JAM	Jamaica	PAK	Pakistan	YEM	Yemen, Rep.
JOR	Jordan	PAN	Panama	ZAF	South Africa
JPN	Japan	PER	Peru	ZMB	Zambia
KAZ	Kazakhstan	PHI	Philippines	ZWE	Zimbabwe
KEN	Kenya	POL	Poland		

Source: Authors.

REFERENCES

- Acemoglu, D., S. Johnson, and J. A. Robinson. 2001. The Colonial Origins of Comparative Development: An Empirical Investigation. *American Economic Review*. 91 (5). pp. 1369–1401.
- . 2002. Reversal of Fortune: Geography and Development in the Making of the Modern World Income Distribution. *Quarterly Journal of Economics*. 117 (4). pp. 1231–94.
- Aiyar, S., R. Duval, D. Puy, Y. Wu, and L. Zhang. 2013. Growth Slowdowns and the Middle-Income Trap. *IMF Working Paper*. No. 13/71. International Monetary Fund. Washington, DC.
- Alesina, A., S. Ozler, N. Roubini, and P. Swagel. 1996. Political Instability and Economic Growth. *Journal of Economic Growth*. 1 (2). pp. 189–211.
- Alfaro, L., C. Areendam, S. Kalemli-Ozcan, and S. Selin. 2003. FDI and Economic Growth: The Role of Local Financial Markets. *Journal of International Economics*. 61 (1). pp. 512–33.
- Balasubramanyam, V. N., M. Salisu, and D. Sapsford. 1996. Foreign Direct Investment and Growth in EP and IS Countries. *Economic Journal*. 106 (434). pp. 92–105.
- Barro, R. 1991. Economic Growth in a Cross-Section of Countries. *Quarterly Journal of Economics*. 106 (2). pp. 407–43.
- . 2000. Inequality and Growth in a Panel of Countries. *Journal of Economic Growth*. 5 (1). pp. 5–32.
- Barro, R. J. and J.-W. Lee. 2013. A New Data Set of Educational Attainment in the World, 1950–2010. *Journal of Development Economics*. 104 (C). pp. 184–98.
- Barro, R. J. and X. Sala-i-Martin. 2004. *Economic Growth*. Cambridge, MA: The MIT Press.
- Benhabib, Jess and M. Spiegel, 1994. The Role of Human Capital in Economic Development: Evidence from Aggregate Cross-Country Data. *Journal of Monetary Economics*. 34 (2). pp. 143–73.
- Bloom, D. E., D. Canning, and P. Malaney. 2000. Demographic Change and Economic Growth in Asia. *Population and Development Review* 26(Supplement). pp. 257–90.
- Bloom, D. E., D. Canning, and J. Sevilla. 2003. The Demographic Dividend: A New Perspective on the Economic Consequences of Population Change. Population Matters Monograph MR-1274. RAND. Santa Monica.
- Bloom, D. E., D. Canning, G. Fink, and J. E. Finlay. 2007. Does Age Structure Forecast Economic Growth? *NBER Working Paper*. No. 13221. National Bureau of Economic Research. Cambridge, MA.
- Borensztein, E., J. De Gregorio, and J.-W. Lee. 1998. How Does Foreign Investment Affect Growth? *Journal of International Economics*. 45 (1). pp. 115–72.
- Breiman, L. 2001. Random Forests. *Machine Learning*. 45 (1). pp. 5–32.

- Breiman, L., J. Friedman, R. Ohlsen, and C. Stone. 1984. *Classification and Regression Trees*. Belmont, CA: Wadsworth.
- Bulman, D., M. Eden, and H. Nguyen. 2014. Transitioning from Low-Income Growth to High-Income Growth: Is there a Middle Income Trap? *World Bank Policy Research Working Paper*. No. 7104.
- Calderón, C., E. Moral-Benito, and L. Servén. 2014. Is Infrastructure Capital Productive? A Dynamic Heterogeneous Approach. *Journal of Applied Econometrics*. 30 (2). pp. 177–98.
- Caselli, F. and A. Ciccone. 2013. The Contribution of Schooling in Development Accounting: Results from a Nonparametric Upper Bound. *Journal of Development Economics*. 104. pp. 199–211.
- Deininger, K. and L. Squire. 1998. New Ways of Looking at Old Issues: Inequality and Growth. *Journal of Development Economics*. 57. pp. 259–87.
- Dollar, D. and A. Kraay. 2003. Institutions, Trade and Growth. *Journal of Monetary Economics*. 50 (1). pp. 133–62.
- Durlauf, S. N. and P. A. Johnson. 1995. Multiple Regimes and Cross-Country Growth Behavior. *Journal of Applied Econometrics*. 10 (4). pp. 365–84.
- Easterly, W. and R. Levine. 2003. Tropics, Germs, and Crops: How Endowments Influence Economic Development. *Journal of Monetary Economics*. 50 (1). pp. 3–39.
- Egert, B., T. Kozluk, and D. Sutherland. 2009. Infrastructure and Growth: Empirical Evidence. *William Davidson Institute Working Paper*. No. 957. The William Davidson Institute at the University of Michigan. Ann Arbor.
- Eichengreen, B., D. Park, and K. Shin. 2013. Growth Slowdowns Redux: New Evidence on the Middle-Income Trap. *NBER Working Paper*. No. 18673. National Bureau of Economic Research. Cambridge, MA.
- Felipe, J., U. Kumar, and R. Galope. 2014. Middle-Income Transition: Trap or Myth? *ADB Economics Working Paper*. No. 421. Asian Development Bank. Manila.
- Feyrer, J. 2009. Trade and Income – Exploiting Time Series in Geography. *NBER Working Paper*. No. 14910. National Bureau of Economic Research. Cambridge, MA.
- Forbes, K. 2000. A Reassessment of the Relationship Between Inequality and Growth. *American Economic Review*. 90 (4). pp. 869–87.
- Frankel, J. and D. Romer. 1999. Does Trade Cause Growth? *American Economic Review*. 89 (3). pp. 379–99.
- Glaeser, E. L., R. La Porta, F. Lopez-de-Silanes, and A. Shleifer. 2004. Do Institutions Cause Growth? *NBER Working Paper*. No. 10568. National Bureau of Economic Research. Cambridge, MA.
- Hall, R. E. and C. I. Jones. 1999. Why Do Some Countries Produce so Much More Output per Worker than Others? *The Quarterly Journal of Economics*. 114 (1). pp. 83–116.

- Hanushek, E. A. and L. Woessmann. 2008. The Role of Cognitive Skills in Economic Development. *Journal of Economic Literature*. 46 (3). pp. 607–68.
- Hapfelmeier, A. 2012. Analysis of Missing Data with Random Forests. PhD. Dissertation. Ludwig-Maximilians-Universität. München.
- Henisz, W. J. 2000. The Institutional Environment for Economic Growth. *Economics and Politics*. 12 (1). pp. 1–31.
- . 2002. The Institutional Environment for Infrastructure Investment. *Industrial and Corporate Change*. 11 (2). pp. 355–89.
- Hothorn, T., K. Hornik, and A. Zeileis. 2006. Unbiased Recursive Partitioning: A Conditional Inference Framework. *Journal of Computational and Graphical Statistics*. 15 (3). pp. 651–74.
- Im, F. G. and D. Rosenblatt. 2013. Middle-Income Traps: A Conceptual and Empirical Survey. *Poly Research Working Paper*. No. 6594. The World Bank. Washington, DC.
- International Monetary Fund (IMF). 2014. Is It Time for An Infrastructure Push? The Macroeconomic Effects of Public Investment. Chapter 3. In *World Economic Outlook 2014*. Washington, DC.
- Jones, C. 2015. The Facts of Economic Growth. *NBER Working Paper* No. 21142.
- Mankiw, G., D. Romer, and D. Weil. 1992. A Contribution to the Empirics of Growth. *Quarterly Journal of Economics*. 107 (2). pp. 407–37.
- Nakao, Takehiko. 2014. An Eight-Point Development Agenda for Sustainable Growth in Asia. Keynote address at the 5th Asian Development Forum. Ha Noi, Viet Nam. 19 September. <http://www.adb.org/news/speeches/eight-point-development-agenda-sustained-growth-asia-takehiko-nakao>
- Pritchett, L. and L. H. Summers. 2014. Asiaphoria Meets Regression to the Mean. *NBER Working Paper*. No. 20573. National Bureau of Economic Research. Cambridge, MA.
- Reinhart, C. M. and K. S. Rogoff. 2009. *This Time is Different: Eight Centuries of Financial Folly*. Princeton: Princeton University Press.
- Robertson, P. E. and L. Ye. 2013. On the Existence of a Middle Income Trap. *Economics Discussion Paper*. No. 13.12. University of Western Australia. Perth.
- Rodriguez, F. and D. Rodrik. 2001. Trade Policy and Economic Growth: a Skeptic’s Guide to the Cross-National Evidence. *NBER Macroeconomics Annual 2000*. 15. pp. 261–338.
- Rodrik, D., A. Subramanian, and F. Trebbi. 2002. Institutions Rule: The Primacy of Institutions over Geography and Integration in Economic Development. *NBER Working Paper*. No. 9305. National Bureau of Economic Research. Cambridge, MA.
- Rudengren, T., L. Rylander, and C. R. Casanova. 2014. It’s Democracy, Stupid: Reappraising the Middle-Income Trap. Stockholm Paper. Institute for Security and Development Policy. Stockholm-Nacka.

- Shmueli, G., N. R. Patel, and P. C. Bruce. 2007. *Data Mining for Business Intelligence: Concepts, Techniques, and Applications in Microsoft Office Excel with XLMiner*. New Jersey: Wiley-Interscience.
- Strasser, H. and C. Weber. 1999. On the Asymptotic Theory of Permutation Statistics. *Mathematical Methods of Statistics*. 8 pp. 220–50.
- Straub, S. 2008. Infrastructure and Growth in Developing Countries: Recent Advances and Research Challenges. *World Bank Working Paper*. No. 4460. World Bank. Washington, DC.
- Strobl, C., A. L. Boulesteix, A. Zelielis, and T. Hothorn. 2007. Bias in Random Forest Variable Importance Measures: Illustrations, Sources and a Solution. *BMC Bioinformatics*. 8 (1). p. 25.
- Williamson, J. 1990. “What Washington Means by Policy Reform.” In J. Williamson (ed.), *Latin American Adjustment: How Much Has Happened?* Washington, DC: Institute for International Economics.
- World Bank. 2013. *The World Bank Group Goals: End Extreme Poverty and Promote Shared Prosperity*. Washington, DC.
- . World Development Indicators. <http://data.worldbank.org/data-catalog/world-development-indicators>

Re-examining the Middle-Income Trap Hypothesis: What to Reject and What to Revive?

Contrary to a popular belief, the data reject a simple-minded hypothesis of a middle-income trap; most middle-income countries will attain the living standard of today's United States. However, policy choices and fundamentals matter. Macroeconomic stability, openness to foreign direct investment, and good infrastructure, besides favorable demographics help to produce faster growth. There are also conditions under which countries can be trapped or even move backward.

About the Asian Development Bank

ADB's vision is an Asia and Pacific region free of poverty. Its mission is to help its developing member countries reduce poverty and improve the quality of life of their people. Despite the region's many successes, it remains home to the majority of the world's poor. ADB is committed to reducing poverty through inclusive economic growth, environmentally sustainable growth, and regional integration.

Based in Manila, ADB is owned by 67 members, including 48 from the region. Its main instruments for helping its developing member countries are policy dialogue, loans, equity investments, guarantees, grants, and technical assistance.

