CEAMeS Discussion Paper Series

CEAMeS Discussion Paper No. 11/2017

The Deep Determinants at More Subtle Stages of Development – The Example of the Middle-Income Trap Phenomenon

Linda Glawe & Helmut Wagner
(University of Hagen)

Center for East Asia Macro-economic Studies [CEAMeS]
University of Hagen

Universitätsstr. 11, 58084 Hagen, Germany

Also distributed as SSRN Working Paper No. 3085417 (2017).

The Deep Determinants at More Subtle Stages of Economic Development – The Example of the Middle-Income Trap Phenomenon

Linda Glawe[a] and Helmut Wagner[b]

Abstract: The so-called 'deep determinants' of economic growth and development (namely, geography, institutions, and integration) have been found to be decisive for the break out of stagnation and for explaining cross-country income differences by many empirical studies. However, so far, very little has been done to examine to which extent they are also crucial at more subtle stages of economic development. Our paper aims to close this gap by focusing on the phenomenon of the middle-income trap (MIT) which has reached increasing attention in the last 15 years. In particular, we test whether the results of the empirical studies conducted by Acemoglu et al. (2001), Rodrik et al. (2004), and Easterly and Levine (2016) also remain valid when analyzing the MIT. We are the first to analyze the relationship between the deep determinants and the MIT, especially regarding the causal effect of institutional quality on the probability of experiencing a growth slowdown at the middle-income range. Our analysis reveals that while, in general, the deep determinants also seem to play an important role for the middle-income transition (and the question of whether a country falls into an MIT), some differences compared to the results of the standard literature become apparent.

Keywords: deep determinants of growth, economic development, economic growth, middle-income trap, geography, institutions, openness

JEL Classification: O10, O11, O43, O57

[*] We would like to thank all participants and discussants of the Dynamics, Economic Growth and International Trade (DEGIT XXIII) Conference (Moscow, 2018) as well as Daniel Berkowitz for helpful comments.

[a, b] University of Hagen, Faculty of Economics, Chair of Macroeconomics, 58084 Hagen, Germany, phone +4923319872640, fax +492331987391, e-mail linda.glawe@fernuni-hagen.de and helmut.wagner@fernuni-hagen.de

1 Introduction

Since the 1990s, a considerable body of literature has emerged, focusing on the so-called 'deep determinants' as an alternative to the proximate factors of economic growth and development postulated by neoclassical and endogenous growth models (such as physical and human capital, total factor productivity) for explaining current cross-country differences in per capita income.[1] In particular, it is argued that long-term geographic and historical variables such as climate, disease burden, legal origin and colonial heritage are important factors underpinning growth and development – both, directly and indirectly. While the role of the deep determinants is well examined for explaining overall income discrepancy in cross-country comparison, the importance of these factors at more subtle stages of development has not yet been investigated extensively.[2] Our paper aims to close this gap by focusing on the middle-income trap (MIT) phenomenon. In particular, we test whether the results of studies conducted by Acemoglu et al. (2001), Rodrik et al. (2004), and Easterly and Levine (2016) also remain valid when analyzing the MIT. Our analysis reveals that the deep determinants also seem to play an important role for the question of whether a country falls into an MIT. However, several differences compared to the results of the standard literature become apparent, indicating that the deep determinant concept needs to be refined so that it takes into account the special characteristics of the challenging middle-income transition.

The rest of the paper is structured as follows: The next section briefly summarizes the previous literature on the MIT and on the deep determinants of growth. In Sections 3 to 5 we apply the studies of Acemoglu et al. (2001), Rodrik et al. (2004), and Easterly and Levine (2016) to the MIT phenomenon. In each section, we provide some descriptive statistics, present our (modified) estimation strategy as well as our empirical results, and compare our findings with those of the standard literature. In Section 6, we then briefly summarize the main results elaborated in the previous sections. Concluding remarks are provided in Section 7.

2 Literature review

This Section briefly reviews the literature on the MIT as well as on the deep determinants of growth. For extensive surveys on the MIT see Agénor (2016) as well as Glawe and Wagner (2016). Regarding the deep determinants of growth, comprehensive literature overviews are provided by Easterly and Levine (2003) and Spolaore and Wacziarg (2013).

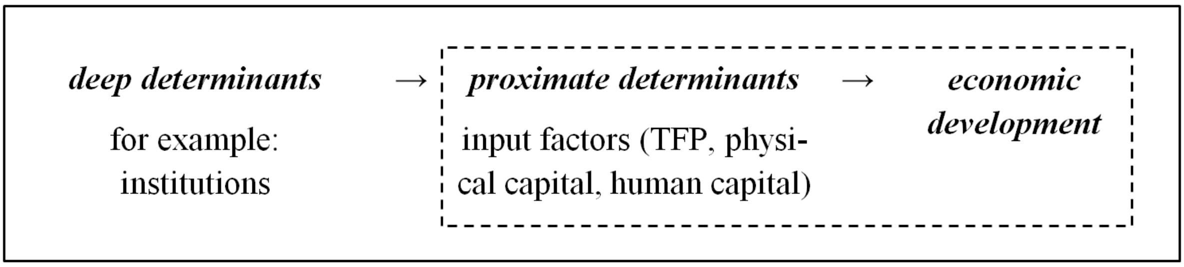
The middle-income trap: The term MIT, introduced by Gill and Kharas in 2007, refers to the often-observed case that a developing country's growth rate decreases significantly when the country reaches the middle-income range (MIR) (Glawe and Wagner, 2016). More precisely, it can be distinguished between absolute and relative empirical definitions of the MIT, the former interpreting it as a prolonged growth slowdown at the MIR, the latter as a failed catching-up process to the advanced economies. According to a widely cited study conducted by the Word Bank (2013), only 13 of 101 countries successfully managed this challenging middle-income transition between 1960 and 2008.

[1] See Easterly and Levine (2003), Rodrik et al. (2004), Owen and Weatherstone (2007), Nunn (2009), as well as Spolaore and Wacziarg (2013).
[2] The article of Lee and Kim (2009) is one of the few exceptions. However, the authors focus especially on the effectiveness of policies at different stages of economic development.

So far, the MIT literature has been mainly empirical and the main triggering factors identified by the empirical studies are the export structure, total factor productivity, and human capital (see Glawe and Wagner, 2017a, for a meta-analysis of the empirical MIT literature).[3] However, the effects of the deep determinants on the MIT phenomenon still remain to be elucidated. A first rather descriptive study by Glawe and Wagner (2017c) explores the relationship between these factors and the MIT probability by using simple hypothesis testing. Their results indicate that it is promising to further investigate this relationship as it does not only seem to be important for the middle-income transition, but their study also reveals several differences to the results of the standard literature.

The deep determinants of growth: The difference between the 'traditional approach' and the 'deep determinant approach' for explaining a country's aggregate output is illustrated by Figure 1. The traditional approach explains economic growth and development solely through the growth of the proximate determinants (that is, the input factors), whereas the deep determinant theory also considers the underlying factors that determine the proximate determinants (North and Thomas, 1973; Acemoglu et al., 2014).

Figure 1. Traditional approach versus the deep determinant approach.



Source: Own representation based on North and Thomas' (1973) theoretical approach.
Notes: The traditional approach only covers the transmission channel in the dashed frame.

In general, the literature agrees that the deep determinants can be broadly classified into three strands, namely (1) geography, (2) institutions, and (3) integration/international trade. More recent analyses that take into account the three determinants simultaneously, postulate the primacy of institutions over the other determinants (examples include Rodrik et al., 2004, and Bhattacharyya, 2004). It has to be noted, however, that very often, geography is found to have an indirect effect on institutions in these studies.

In the following, we present the main results of the studies conducted by Easterly and Levine (2016), Acemoglu et al. (2001), and Rodrik et al. (2004) as we will focus on them in the regression analysis presented in Sections 3 to 5.

Easterly and Levine (2016) (henceforth: EL) find a strong positive correlation between the share of Europeans in colonial population (henceforth: euroshare) and current per capita income. In their analysis, they explore two channels identified in the literature through which

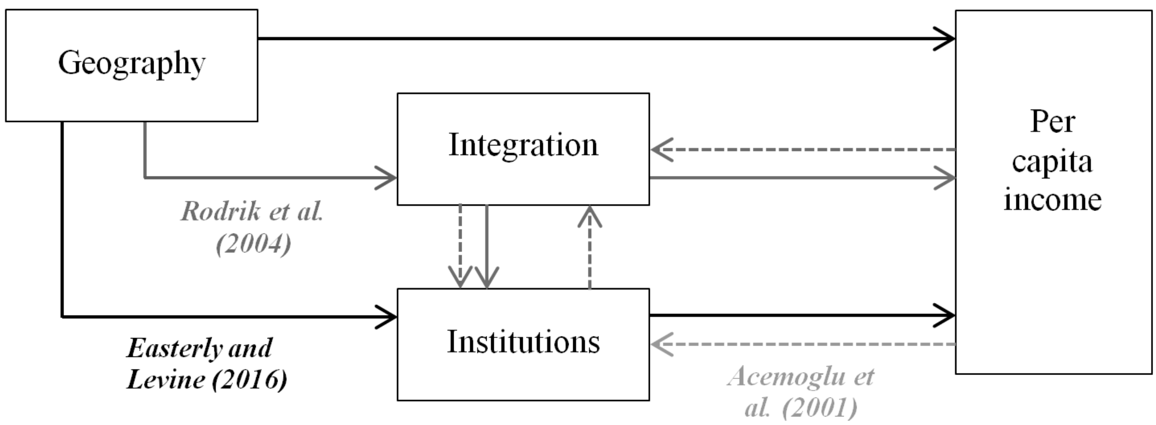
[3] To our knowledge, there are only three growth models, namely a two period overlapping generations model developed by Agénor and Canuto (2015) as well as the country specific models of Dabús et al. (2016) and Glawe and Wagner (2017b).

the euroshare has impacts on the today's differences in cross-countries economic performance, namely the *institutional channel* proposed by Engerman and Sokoloff (1997) as well as by Acemoglu et al. (2001) and the *human capital channel* suggested by Glaeser et al. (2004).

Acemoglu et al. (2001) (henceforth: AJR), suggesting "a 'germs' theory of institutions"[4], argue that European's were more likely to install solid, growth-promoting institutions in areas where they faced a benign disease environment. In areas with a relatively unfavorable disease environment, however, European powers set up extractive colonies and did not introduce much legal protection for private property. AJR use the differences in European mortality rates as an instrument for current institutions to estimate the effects on the GDP per capita.[5]

Rodrik et al. (2004) (henceforth: RST) extend the analysis of AJR by including the integration/trade dimension suggested by Frankel and Romer (1999). RST use the instruments proposed by these two studies, namely the settler mortality as an instrument for institutional quality and the constructed trade share as an instrument for integration. They find that institutional quality "trumps" everything else, whereas geography and integration have no or only weak direct effects on the per capita income (p. 141). In addition, they show that geography and also integration have an indirect effect by influencing institutional quality.

Figure 2. The deep determinants of economic growth and development.



Source: Own representation based on RST (p. 134).

Figure 2 illustrates which direct and indirect deep determinants channels (depicted by the solid lines) and reverse feedbacks (depicted by the dotted lines) are included in the three studies. EL take into account the channels denoted by the black arrows, AJR additionally consider the reverse feedback regarding institutions (light grey arrow) and RST also take into account the integration channel(s) and reverse feedbacks denoted by the dark grey arrows. In contrast to EL using OLS regression, AJR and RST perform two-stage least square regression analyses.

[4] Easterly and Levine (2003).

[5] See Nunn (2009) for a more detailed summary of Acemoglu et al.'s (2001) study.

3 The European origins of the MIT? – Easterly and Levine (2016) revisited

We start with the recent study of EL in which they analyze the impact of the share of Europeans in the colonial population on the current income via human capital and the quality of institutions. The estimation strategy proposed by EL is probably the less complicated among the three studies adopted in this article, however, it nonetheless offers some interesting insights.

3.1 Data and descriptive statistic

Table 1 provides descriptive statistics for the main variables. The definitions and sources for all variables used in this article are summarized in Table A1 in the Appendix A. Our MIT country sample is based on the relative MIT definition of the World Bank (2013).[6] According to this definition (based on Maddison, 2010, data), a country faces an MIT if it stays within the range of 4.5 to 45 percent of the US per capita income (in 1990 international Geary-Khamis dollars) in the period from 1960 to 2008.[7, 8]

Table 1. Descriptive statistics (I).

	Observations	Mean	Standard deviation	Min	Max	Median
MIT dummy	38	0.82	0.39	0.00	1.00	1.00
Euroshare	38	0.16	0.24	0.00	0.901	0.06
British legal origin	38	0.26	0.45	0.00	1.00	0.00
Secondary enrollment	36	68.20	28.76	8.22	152.87	68.39
Secondary total	35	35.90	14.80	2.87	63.31	32.94
Secondary completed	35	19.50	10.18	1.41	54.55	17.37
Independence	38	0.52	0.297	0.00	1.00	0.479
Ethnicity	35	0.34	0.29	0.00	0.83	0.23
Government quality	38	0.13	1.85	-4.27	3.08	0.77
Export share (1960-2000)	37	30.19	16.16	7.69	65.40	25.64
Openness (1960-2000)	37	63.71	31.68	16.09	138.58	56.31

Notes: Variable definitions and sources are provided in Appendix A and in the text below.

As already described above, *euroshare* equals the share of Europeans in colonial population. As argued by Glawe and Wagner (2017c), simple hypothesis testing reveals that

[6] Other samples, for example based on the definitions of Aiyar et al. (2013) and Felipe et al. (2012) have either a too small sample size to yield interpretable results or include too less non-MIT countries due to limited data availability of the euroshare variable. For example, in case of the Felipe et al. (2012) sample, data on the euroshare is only available for one non-MIT country of the original sample. See also the discussion in Section 4.1.

[7] Due to the fact that the World Bank (2013) study does not provide a comprehensive list of their identified MIT countries, we reproduced their results using their thresholds and the Maddison (2010) database.

[8] In particular, our sample consists of countries that are either caught in an MIT or that have already achieved high-income status. We do not include low-income countries and countries that have not been long enough in the MIR to make a statement as to whether they are already trapped.

the mean euroshare is significantly less (at the 1-percent level) in MIT countries than in non-MIT countries (that is, countries that managed a timely shift from middle- to high-income status) and the result stays robust when using different MIT definitions.[9]

Similar to EL, we use several other variables that capture important characteristics of a country. *British legal origin* is a dummy variable equaling one if the country has a British common law legal tradition and zero otherwise. La Porta et al. (1999) argue that the differences between legal systems implemented by the colonial powers, in particular between the British common law and the French civil law, were important for the development of institutions and thus, for the long-term development. In particular, it is argued that a common law tradition (in comparison to other legal origins) is associated with a stronger emphasis on the protection of property rights and a less interventionist and more efficient government as well as a stronger protection of the individual against the government (see also Finer, 1997 and North, 1990). *Secondary enrollment* presents the average gross rate of secondary school enrollment from 1995 to 2005 (World Bank data) and is used by EL as an indicator for human capital. We construct two additional indicators using the Barro and Lee (2013) dataset, namely the average percentage of the population aged 15 and over with secondary education (total) between 1995 and 2005 (*Secondary total*) as well as the average percentage of the population aged 15 and over with secondary education (completed) between 1995 and 2005 (*Secondary completed*). *Independence* indicates the percentage of years since 1776 that a country has been independent and is used to measure the extent to which a country has had the time to develop and install its own institutions. *Ethnicity* is the average of five different indices of ethnolinguistic fragmentation (ranging from zero to one) and measures the population's heterogeneity. *Government quality* is a measure of government accountability and effectiveness. It is calculated on the base of the Worldwide Governance Indicators (in particular, the first principal component of six individual indicators for the year 2005). *Export share* presents the average export share in GDP between 1960 and 2000 and *Openness* stands for the openness to international markets, measured as the trade share in GDP between 1960 and 2000 (see also Agbor, 2010).

3.2 Probit regression

Estimation strategy

In contrast to general cross-country growth regressions where the dependent variable is a continuous variable such as the log per capita income, we have the dichotomous outcome variable MIT_i which takes the value one if the country i is caught in an MIT and 0 if it succeeded to overcome the MIR without experiencing an MIT. That is, in contrast to EL, we estimate a probit model given by the following equation:

(1) $$P(MIT_i = 1 |\, euroshare_i, z_i) = \Phi(\alpha_1 + \alpha_2\, euroshare_i + z_i'\alpha_3),$$

[9] In particular they use the definitions developed by Aiyar et al. (2013) and the World Bank (2013); the samples based on other definitions are too small to conduct statistic test, however, the descriptive analysis supports their findings.

where MIT_i is the dummy variable for the MIT in country i, $euroshare_i$ denotes the proportion of Europeans in colonial population, z_i is a vector of other covariates (country characteristics), and $\Phi(\cdot)$ is the cumulative distribution function for the standard normal distribution.

Probit regression results

Table 2 presents our probit regression results. Column (1) reveals a strong negative correlation between the MIT dummy and euroshare with a coefficient of -3.99 (s.e. = 1.19) and an average marginal effect of -0.58 (s.e. = 0.15). Similar to EL, this relationship stays statistically significant at the 1-percent level when adding a British legal origin dummy (Column 2) or independence (Column 3) and still at the 5-percent level when adding ethnolinguistic fragmentation (Column 4). However, in contrast to EL who find that ethnic heterogeneity is highly significant for current income, we find no significant relationship between the MIT dummy and ethnicity. It could be hypothesized that ethnolinguistic fragmentation might be important for explaining why some countries manage to break out of stagnation, but not why some countries successfully manage the challenging middle-income transition while others do not.

Table 2. Probit estimates with country characteristics.

	Base sample (1)	Base sample (2)	Base sample (3)	Base sample (4)	Base sample (5)	Base sample (6)
Euroshare	-3.9888*** (1.1870) -0.5784*** (0.1510)	-3.7639*** (1.3784) -0.4793*** (0.1502)	-4.7534*** (1.4744) -0.5660*** (0.1312)	-6.4982** (3.0782) -0.7237** (0.3678)	-5.9322** (2.8123) -0.6638** (0.3302)	-9.7083 (6.1800) -0.7375* (0.4211)
British legal origin		-1.0353 (0.6671) -0.1318 (0.0856)				-0.4311 (0.9167) -0.0327 (0.0680)
Independence			2.5963 (1.7350) 0.3091 (0.2020)			5.6098 (4.9367) 0.4261 (0.3568)
Ethnicity				-1.0537 (1.2894) -0.1174 (0.1492)		0.4130 (1.6953) 0.0314 (0.1297)
Latitude					0.3886 (3.38164) 0.0435 (0.4267)	
McFadden's Pseudo R^2	0.4317	0.4986	0.5298	0.5469	0.5291	0.6807
Correctly classified (%)	92.11	89.47	92.11	94.29	94.44	91.43
Number of observations	38	38	38	35	36	35

Note: Dependent variable: MIT dummy constructed analogously to the World Bank (2013) study, using Maddison (2010) data. We report both, the coefficients and, below, the average marginal effects. The respective standard errors are in parentheses. Significance at the 10, 5, and 1 percent levels are denoted by *, **, and ***, respectively. See Appendix Table Al for detailed variable definitions and sources.

In Column (5), we also add latitude to control for the effect of climate on economic development as proposed by Sachs and co-authors (this variable is actually frequently used in cross-country growth regressions). Again, euroshare stays significant (at the 5-percent level) with a relatively high average marginal effect of -0.66 (s.e. = 0.33). In all Columns, the McFadden's pseudo R-squared is above 0.43, indicating a good model fit.[10] Finally, Column (6) shows that when conditioning for legal origin, independence, and ethnicity (that is, the country characteristics proposed by EL) simultaneously, the euroshare coefficient is not significant, however, the average marginal effect (-0.74, s.e. = 0.42) is significant at the 10-percent level. The decrease in the significance level is probably due to the small sample size (only 35 countries) and the simultaneous use of four regressors.

To learn more about the potential channels (namely the institutional channel and human capital channel) that connect euroshare and – in our case – the MIT dummy, we follow EL by conditioning first on human capital (by including secondary enrollment in our regression equation) and then on institutions (by including government quality in the equation). The argumentation is the following: If euroshare is related to the MIT dummy through institutions and/or human capital, it should not have a significant relationship when conditioning for these variables. This approach is especially useful in our case due to our small sample size that partly restricts the possibility of IV probit regressions that could also be used (see also Section 4.3 and 5.3 for such analyses). However, as EL note, it has to be taken into account that institutions as well as education are both endogenous to our dependent variable and thus, our findings should also be interpreted with caution. Nonetheless, it might deliver us with useful insights. Our results are presented in Table 3.

Regarding the institutional channel, our empirical results are generally in line with EL's findings: Column (7) shows that when adding government quality, the coefficient and average marginal effect of/on euroshare drop considerably and become insignificant whereas the institutional coefficient (-0.69, s.e. = 0.29) and average marginal effect (-0.06, s.e. = 0.02) are significant at the 5- and 1-percent level, respectively.[11]

Regarding the human capital channel, we derive different results than EL: Although in Column (8) *euroshare* turns insignificant when adding *secondary enrollment*, the coefficient of the human capital proxy is not significant (and the average marginal effect of -0.01 (s.e. = 0.003) is only significant at the 10-percent level) indicating that the human capital channel might play a less important role for middle-income transitions. This result is confirmed when using other measures of (secondary) education, compiled on the basis of the widely used Barro and Lee (2013) dataset (see Columns 9 and 10).[12] In particular, when conditioning for the *secondary total* variable, euroshare stays significant at the 5-percent level with an average marginal effect of -0.38 (s.e. = 0.16), whereas the human capital proxy is insignificant. We derive similar results when using the *secondary completed* variable (however, with a slightly reduced significance of the *euroshare* variable compared to Column 9).[13, 14] In all probit mod-

[10] In particular, according to McFadden (1979) a McFadden's pseudo R-squared ranging from 0.2 to 0.4 indicates very good model fit.
[11] Using the first principal components of the six WGIs of various other years does not change our results.
[12] Using data on the average percentage of the population with completed tertiary education leads to similar results.
[13] When only including the human capital proxies, they are highly relevant.
[14] Interestingly, Acemoglu and co-authors derive similar results in a subsequent paper (Acemoglu et al., 2014). In particular, they show that when focusing on historically-determined differences in human capital and control

els, the pseudo R-squared is relatively high (ranging from ca. 0.45 to 0.67), indicating a good model fit.

Table 3. Probit regression with (potential) channels of influence.

	Base sample (7)	Base sample (8)	Base sample (9)	Base sample (10)	Base sample (11)	Base sample (12)
Euroshare	-2.6040 (2.2113) -0.2115 (0.1818)	-1.8757 (1.4354) -0.2051 (0.1413)	-2.8947** (1.2751) -0.3782** (0.1559)	-2.4443* (1.3188) -0.2879** (0.1418)	-4.8464*** (1.5068) -0.6029*** (0.1149)	-4.5454*** (1.3456) -0.5942*** (0.1207)
Government quality	-0.6873** (0.2850) -0.0558*** (0.0201)					
Secondary enrollment		-0.0479 (0.0301) -0.0052* (0.0031)				
Secondary total			-0.0340 (0.0823) -0.0044 (0.0036)			
Secondary completed				-0.0859 (0.0593) -0.0101 (0.0068)		
Export share					-0.0387 (0.0244) -0.0048* (0.0029)	
Openness						-0.0156 (0.0108) -0.0020 (0.0014)
McFadden's Pseudo R^2	0.6743	0.5453	0.4514	0.5021	0.5188	0.4938
Correctly classified (%)	92.11	91.67	88.57	88.57	91.89	91.89
Number of observations	38	36	35	35	37	37

Note: Dependent variable: MIT dummy constructed analogously to the World Bank (2013) study, using Maddison (2010) data. We report both, the coefficients and, below, the average marginal effects. The respective standard errors are in parentheses. Significance at the 10, 5, and 1 percent levels are denoted by *, **, and ***, respectively. See Appendix Table Al for detailed variable definitions and sources.

EL mention that besides institutions and human capital, Europeans brought also other growth-promoting characteristics that had enduring effects on the economic performance, for example connections with international markets. Therefore, we include the international trade channel in our analysis. As proposed by Agbor (2010), we use the average share of exports in GDP as well as the average share of trade in GDP (as a measure of openness) between 1960

for institutions, the estimates on the effect of human capital are reduced significantly while the estimates on the institutional effect stay robust. See also Section 4.6.

and 2000 to test whether *euroshare* and the *MIT dummy* are connected via these variables. As shown in Columns (11) and (12), we find no evidence that support such a trade channel. In both Columns, the institutional coefficients and average marginal effects are relatively high and highly significant.

Our empirical results have shed some light on the importance of the share of European settlers for the middle-income transition and the potential channels that might connect euroshare and the MIT dummy. In line with EL (for general differences in per capita income), we find that the euroshare is strongly negatively related with our MIT dummy and that this result is robust when controlling for various country characteristics such as legal origin, independence or ethnicity. However, our results differ with respect to the importance of the potential channels that might connect euroshare and the MIT dummy. Whereas the results regarding the institutional channel are still valid for the MIT dummy, our findings with respect to the human capital channel suggest that this positive effect may have already faded away when a country has reached the MIR and is not able to explain why some countries become trapped at the MIR. We found no empirical evidence for a potential third channel, namely international trade. Regarding other control variables, our analysis revealed that ethnolinguistic heterogeneity (that is found to be highly significant in EL's regression) is insignificant. That is, it might have played a decisive role for the break out of stagnation, but not for explaining different growth trajectories at more subtle stages of development. As we will see later, although ethnolinguistic fragmentation does not have an direct effect on the MIT dummy, it actually has a strong indirect effect by negatively influencing the institutional quality in some model specifications in Section 4.

4 The colonial origins of the MIT? – Acemoglu et al. (2001) revisited

An econometrically more sophisticated analysis is performed by AJR. They use a two-stage least squares regression analysis to examine the causal effect of the institutional quality on the per capita income. In particular, AJR use the settler mortality rates as instruments for the quality of institutions. In this section, by using a slightly modified version of their econometric model, we examine whether this causal relationship also holds for the MIT phenomenon. In addition, we test various other instruments, for example language variables and the legal origin.

4.1 Data and descriptive statistics

Descriptive statistics of the key variables are provided in Table 4. The definitions and sources for all variables used in this article are summarized in Table A1 in the Appendix A. As before, our MIT country sample is based on the relative MIT definition of the World Bank (2013) (see also the discussion in Section 3.1).

Table 4. Descriptive statistics (II).

	Large sample	Base Sample 1	Base sample 2	Base sample 3
MIT dummy (MIT)	0.68 (0.47)	0.812 (0.39)	0.77 (0.42)	0.76 (0.43)
Average protection against expropriation risk (AVEXPR)	7.45 (2.74)	6.89 (1.43)	7.00 (1.60)	7.02 (1.62)
Log European settler mortality (LOGEM4)	–	4.29 (1.07)	–	–
Distance from equator (Latitude)	0.31 (0.18)	0.21 (0.14)	0.25 (0.17)	0.26 (0.17)
French legal origin dummy (LEGOR_FR)	0.13 (0.33)	0.57 (0.50)	0.54 (0.50)	0.53 (0.50)
Fraction of the population speaking English (ENGFRAC)	0.09 (0.25)	0.13 (0.31)	0.11 (0.28)	0.11 (0.28)
Fraction of the population speaking other European languages (EURFRAC)	0.32 (0.42)	0.42 (0.43)	0.32 (0.42)	0.32 (0.42)
Number of countries	101	44	61	58

Notes: Variable definitions and sources are provided in Appendix A. Standard errors are in parentheses.

Although we cannot econometrically test whether the results also remain valid when using other MIT definitions (due to sample size restrictions), the analysis of Glawe and Wagner (2017c) reveals that other MIT country samples show very similar characteristics regarding the settler mortality, latitude, etc. In addition, we also provide descriptive statistics in form of two graphs for two alternative MIT country definitions (in particular the definitions of Felipe et al., 2012 and Aiyar et al., 2013) in the Appendix which reveal a very similar relationship between our instrument and the institutional measure (see also Section 4.3).

Column (1) of Table 4 presents the statistics for the large sample[15] and Column (2) covers our base sample 1 of 44 countries for which we have data on settler mortality and which are former colonies. The two remaining columns are for two alternative, larger base samples 2 and 3 for which we have data on alternative instruments, namely the French legal origin ($LEGOR\_FR$) (Column 3) and two language variables (Column 4).

$AVEXPR$ denotes the average risk against expropriation between 1985 and 1995, ranging from 0 to 10 where a higher score means less risk. LOGEM4 is the log of the settler mortality rate per thousand. The fraction of the populations speaking English (one of the five major European languages) is denoted by $ENGFRAC$ ($EURFRAC$). $LEGOR\_FR$ is a dummy variable indicating a French civil law tradition.

[15] Consisting of countries that are either caught in an MIT or that have already achieved high-income status.

4.2 Probit regression

Estimation strategy

Our regression model to estimate the effect of institutions on the probability of an MIT is given by the following equation:

(2) $$P(MIT_i = 1 | INS_i, z_i) = \Phi(\beta_1 + \beta_2 INS_i + z_i' \beta_3),$$

where MIT_i is the dummy variable for the MIT in country i, INS_i denotes the institutional measure, in particular the average protection against expropriation ($AVEXPR$), and z_i is a vector of other covariates. $\Phi(\cdot)$ is the cumulative distribution function for the standard normal distribution.

Probit estimation results

Our estimation results for our base sample 1 and our large sample are presented in Table 5. Column (1) reveals that in our large sample, there is a strong correlation between our institutional variable ($AVEXPR$) and the MIT dummy. Column (2) shows that these findings are in general confirmed regarding our base sample 1, however, with a slightly reduced significance (5-percent level) and a slightly lower average marginal effect (-0.14 versus -0.15, both significant at the 1-percent-level). Adding latitude does not change our results significantly, we only report a marginal decrease in the coefficients and average marginal effects for both, the large sample (Column 3) and the base sample 1 (Column 4). In all Columns, the McFadden's pseudo R-squared is above 0.66, indicating a very good model fit. In general, the pseudo R-squared is higher in our base sample 1 than in our large sample. In Column (5)–(8), we add continent dummies, namely Latin America, Asia, and Europe with "Other continent" as the omitted group.[16] Regarding the large sample (Columns 5 and 6), the coefficient and average marginal effect for the institutional variable remain significant at the 1-percent level and, in contrast to the findings of AJR, the latitude variable and the continent dummies are not statistically significant. Regarding our base sample 1, Column (7) reveals that only including continent dummies also does not change our results significantly; we even report a higher institutional coefficient and a higher average marginal effect. (Note that the Europe dummy has been omitted due to perfect separation, which is not unusual regarding the small sample size). When simultaneously including the continent dummies and latitude, we report a much higher p-value for the coefficient (0.074) and for the average marginal effects (0.017) of the institutional measure. However, on account of the fact that our base sample 1 consists of only 44 observations, which is a rather small sample size for a logistic model with four regressors, this drop in significance should not be overrated.

Overall, our regression results in Table 5 indicate a strong correlation between the average expropriation risk as an institutional measure and the MIT dummy. Thus, we derive similar results as AJR regarding the impact on the log GDP per capita. However, in contrast to AJR's findings, latitude and the remaining continent dummies are insignificant.

[16] Note that we chose to include the Latin America dummy instead of the Africa dummy because the literature on MITs agrees that most Latin American countries are affected by the MIT and it seems logical to control for this dummy in our analysis.

Table 5. Probit estimates, institutions, large sample and base sample 1.

	Large sample (1)	Base sample 1 (2)	Large sample (3)	Base sample 1 (4)	Large sample (5)	Large sample (6)	Base sample 1 (7)	Base sample 1 (8)
Institutions (AVEXPR)	-1.3619***	-1.8281**	-1.2853***	-1.7677**	-1.5119***	-1.4107***	-2.1606**	-1.9839*
	(0.2710)	(0.7401)	(0.3054)	(0.7171)	(0.3625)	(0.3688)	(1.0737)	(1.1111)
	-0.1545***	-0.1388***	-0.1460***	-0.1314***	-0.1615***	-0.1492***	-0.1535***	-0.1394**
	(0.0104)	(0.0313)	(0.0211)	(0.0336)	(0.0225)	(0.0261)	(0.0479)	(0.0583)
Latitude			-0.5872	-1.5107		-1.5860		-1.2326
			(1.5096)	(3.3040)		(2.2364)		(3.6590)
			-0.0667	-0.1123		-0.1677		-0.0866
			(0.1713)	(.2418)		(0.2339)		(0.2549)
Latin America dummy					-1.1766	-1.2007	-1.5672	-1.5091
					(1.2256)	(1.3442)	(2.5419)	(2.7729)
					-0.1257	-0.1270	-0.1113	-0.1060
					(0.1291)	(0.1403)	(0.1744)	(0.1900)
Asia dummy					-1.3089	-1.2733	-1.0610	-1.1055
					(1.0700)	(1.2106)	(2.2411)	(2.5202)
					-0.1398	-0.1347	-0.0754	-0.0777
					(0.1110)	(0.1255)	(0.1564)	(0.1745)
Europe dummy					-0.7570	-0.4088	(a)	(a)
					(1.0285)	(1.2677)		
					-0.0809	-0.0432		
					(0.1086)	(0.1338)		
McFadden's Pseudo R^2	0.6688	0.7111	0.6655	0.7161	0.6880	0.6873	0.7263	0.7290
Correctly classified	91.46%	93.18%	92.59%	93.18%	91.46%	91.36%	93.18%	93.18%
Log likelihood	-17.2109	-6.0274	-17.0027	-5.9221	-16.2141	-15.8977	-5.7096	-5.6545
Heteroskedasticity test	0.3387	0.6307	0.6727	0.7219	0.5663	0.6719	0.7580	0.7531
Number of observations	82	44	81	44	82	82	44	44

Note: Dependent variable: MIT dummy constructed analogously to the World Bank (2013) study, using Maddison (2010) data. See Appendix Table Al for detailed variable definitions and sources. We report both, the coefficients and, below, the average marginal effects. The respective standard errors are in parentheses. Significance at the 10, 5, and 1 percent levels are denoted by *, **, and ***, respectively. (a) These variables have been omitted due to perfect separation (due to the small sample size).

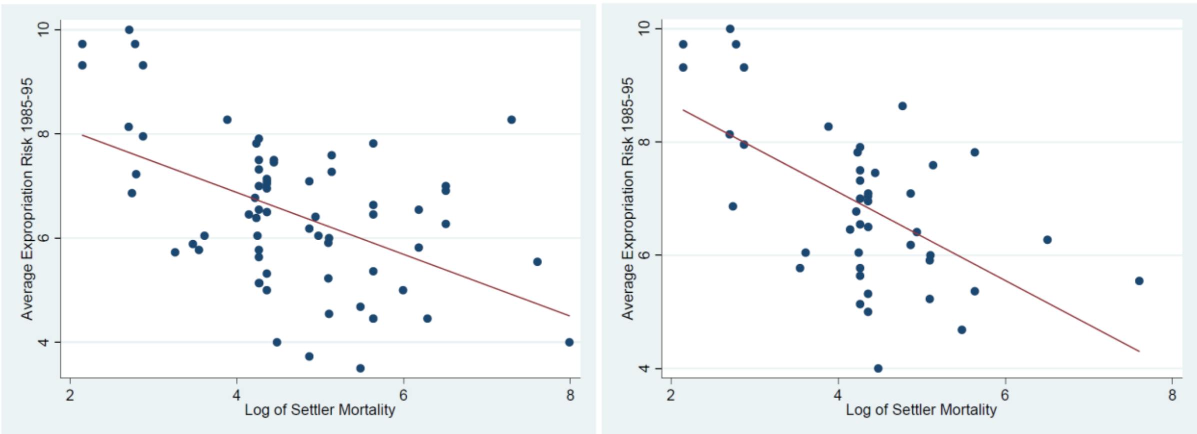
4.3 IV regression model

Estimation strategy

As AJR note, we have to be careful when making statements about the causality of this relationship, inter alia due to the problems of reverse causality (richer economies can afford better institutions) as well as of omitted independent variables correlated with institutions. Therefore, in a next step, we treat INS_i as endogenous and follow AJR by using the settler mortality ($LOGEM4_i$) as an instrument for current institutions INS_i (that is, for $AVEXPR$).[17] Figure 3 illustrates the relationship between settler mortality and current institutions. In Figure (a), we reproduced the scatter plot of AJR's base sample.[18] Figure (b) shows the corresponding scatter plot for our base sample 1 which strongly resembles the AJR scatter plot (actually, there is even a slightly stronger negative relationship in our base sample 1). Scatter plots of two other samples based on other MIT definitions are provided in the Appendix B, Figure B1 (a) and (b). They both confirm the strong negative correlation between settler mortality and the institutional measure.

Figure 3. First-stage relationship between settler mortality and expropriation risk.

(a) AJR base sample **(b)** World Bank (2013) base sample 1



Formally, our probit model with an endogenous explanatory variable is

(3) $MIT_i^* = \beta_1 + \beta_2 INS_i + z_i'\beta_3 + \upsilon_i$

(4) $INS_i = \gamma_1 + \gamma_2\, LOGEM4_i + z_i'\gamma_3 + \nu_i,$

(5) $MIT_i = 1[MIT_i^* > 0].$

where (υ_i, ν_i) has a zero mean and a bivariate normal distribution, that is $(\upsilon_i, \nu_i) \sim N(0, \Sigma)$. Moreover, (υ_i, ν_i) is independent of $LOGEM4_i$ and z_i. Equation (3), along with equation (5), is the structural equation. Equation (4) is a reduced form for INS_i which is endogenous if u_i and ν_i are correlated.[19]

---

[17] For a detailed/in-depth discussion of the choice of this instrument see AJR (p. 1383).

[18] As ARJ suggest, we use the log because this prevents that the extreme African mortality rates do not play a disproportionate role (see ARJ, p. 1383).

[19] See Wooldridge (2002), Chapter 15.7.2.

IV estimation results

The IV estimation results are reported in Table 6. Note that due to the fact that we have an IV probit model, we additionally report the average marginal effects below the respective coefficients. Columns (1) and (2) of Panel B in Table 6 reveals a strong first-stage relationship between (log) settler mortality and current institutions in our base sample 1. In contrast, the relationship between the quality of institutions and latitude is insignificant (see Column 2). The corresponding two-stage probit estimates are reported in Panel A: In Column (1), the institutional coefficient is highly significant (-1.89, s.e. = 0.71). Regarding the average marginal effect we find that for an infinitesimal rise in the institutional measure, the probability of a country to experience an MIT is reduced by around 16 percent. Column (2) shows that similar as by AJR, adding latitude does not change this relationship; we report almost identical coefficients and average marginal effects for our institutional variable. However, in contrast to AJR's finding, the latitude variable does not have the "wrong" sign.

Table 6. IV regression, with and without geographical controls (ARJ).

	Base sample 1 (1)	Base sample 1 (2)	Base sample 1 (3)	Base sample 1 (4)
Panel A: Two-Stage Probit				
Institutions (AVEXPR)	-1.8875*** (0.7106) -0.1569*** (0.0514)	-1.8602*** (0.6717) -0.1581** (0.0708)	-2.2599** (1.1006) -0.1807* (0.0940)	-2.0846** (1.0491) -0.1903 (0.1295)
Latitude		-0.7517 (3.5554) -0.0639 (0.2874)		-0.2125 (3.9014) -0.0194 (0.3474)
Latin America dummy			-1.5283 (2.8198) -0.1222 (0.2073)	-1.3588 (3.0742) -0.1066 (0.2502)
Asia dummy			-0.7512 (2.6964) -0.0600 (0.1996)	-0.5513 (3.1273) -0.0503 (0.2641)
Europe dummy			(a)	(a)
Wald test of exogeneity	0.6670	0.6704	0.7255	0.6694
F-stat	21.90	12.27	8.02	6.84
Log likelihood	-74.2764	-73.0871	-72.8562	-71.4519
Number of observations	44	44	44	44
Panel B: First-Stage for Average Risk Against Expropriation				
Settler mortality (LOGEM4)	-0.7814*** (0.1631)	-0.6558*** (0.1800)	-0.7042*** (0.1758)	-0.5229*** (0.2028)
Latitude		2.0366 (1.3639)		2.3605* (1.4278)
Latin America dummy			-0.3314 (0.3995)	-0.1098 (0.4101)
Asia dummy			0.3284 (0.4809)	0.6074 (0.4962)
Europe dummy			(a)	(a)

Table 6 continued.

	Base sample 1 (1)	Base sample 1 (2)	Base sample 1 (3)	Base sample 1 (4)
Panel C: Probit Regression				
Institutions (AVEXPR)	-1.8281**	-1.7677**	-2.1606**	-1.9839*
	(0.7401)	(0.7171)	(1.0737)	(1.1111)
	-0.1388***	-0.1314***	-0.1535***	-0.1394**
	(0.0313)	(0.0336)	(0.0479)	(0.0583)
Latitude		-1.5107		-1.2326
		(3.3040)		(3.6590)
		-0.1123		-0.0866
		(0.2418)		(0.2549)
Latin America dummy			-1.5672	-1.5091
			(2.5419)	(2.7729)
			-0.1113	-0.1060
			(0.1744)	(0.1900)
Asia dummy			-1.0610	-1.1055
			-0.0754	(2.5202)
			(0.1564)	-0.0777
			(2.2411)	(0.1745)
Europe dummy			(a)	(a)
Log likelihood	-73.0871	-5.9221	-5.7096	-5.6545
McFadden's Pseudo R^2	0.7111	0.7161	0.7263	0.7290
Correctly classified	93.18%	93.18%	93.18%	81.82%
Heteroscedasticity test	0.6307	0.7219	0.7580	0.7531

Note: Dependent variable: MIT dummy constructed analogously to the World Bank (2013) study, using Maddison (2010) data. See Appendix Table Al for detailed variable definitions and sources. 'F-stat' is the F-statistic for the first-stage regression. The row 'Wald test of exogeneity' presents the p-value of the Wald test of exogeneity for the suspected endogenous variable and the row 'Heteroscedasticity test' presents the p-value for the test for heteroscedasticity. In panel A and C, the average marginal effects are reported below the respective coefficients. Standard errors are in parentheses. Significance at the 10, 5, and 1 percent levels are denoted by *, **, and ***, respectively. (a) These variables have been omitted due to perfect separation (due to the small sample size).

We briefly turn to some diagnostic statistics. The F-stat for our first-stage regression for both models is larger than the Staiger and Stock's rule of thumb of 10 and thus, our results do not suffer from the presence of weak instruments. The Wald test of exogeneity results in p-values above 0.65 for both models and thus, it does not support the theoretical claim of endogeneity of institutions. Therefore, a regular probit regression may be appropriate (indeed, the coefficients and average marginal effects do not vary very much between the two-stage probit model and the ordinary probit model). The test for heteroscedasticity yields $\chi^2(1) = 0.63$ and $\chi^2(1) = 0.72$ when adding latitude, suggesting that our probit estimates are homoscedastic.

When adding continent dummies in Columns (3) and (4) of Panel A, both, the institutional coefficients and the average marginal effects are slightly higher than in Column (1) and (2). The institutional coefficients remain significant at the 5-percent level; however, the significance of the average marginal effects is sharply reduced (with p-values of 0.055 and 0.141 when adding latitude). The latitude variable and the continent dummies are insignificant. The corresponding first-stage relationships between the settler mortality and the average expropriation risk (reported in Panel B) remain highly significant. With respect to the diagnostic statistics, it seems noteworthy that the F-stat drops below the threshold of 10. However, the fact that our dataset has only 44 observations (which is very small for a probit model with 3 to 4 regressors – not to speak of one instrumental variable) might (at least to some extent) contribute to the lower F-stats and the higher p-values of the institutional measure. Thus, these results have to be treated with considerable caution. Our ordinary probit regression presented in Panel C reveals very similar coefficients for the institutional measure compared to the two-stage probit model; however, our coefficients are slightly less significant (at the 5-percent level for Columns 1–3 and at the 5-percent level for Column 4). In contrast, the average marginal effects are significant at the 1-percent level for Columns (1)–(3) and at the 5-percent level for Column (4). In all probit models, the estimates are found to be homoscedastic and the Pseudo-R-squared is relatively high (ranging from ca. 0.50 to 0.76), indicating a good model fit.

Due to the small sample size, it might be more sensible to construct a relatively simple model without IV estimation. Therefore, inspired by EL's approach presented in the previous section, we set up the following equation:

(6) $$P(MIT_i = 1 | LOGEM4_i, z_i) = \Phi(\delta_1 + \delta_2 LOGEM4_i + z_i' \delta_3)$$

Our results are presented in Table 7. Column (1) reveals a strong positive correlation between the MIT dummy and settler mortality. The average marginal effect is around 0.18 (s.e. = 0.04), that is, an infinitesimal rise in the settler mortality increases the probability of experiencing an MIT by ca. 18 percent. The relationship stays statistically significant at the 1-percent level (with average marginal effects about 0.16) when adding latitude (Column 2) or continent dummies (Column 3) or both (Column 4). To learn more about the potential institutional channel that connects settler mortality and the MIT dummy, we condition on institutions by including $AVEXPR$ in the equation: If settler mortality is related to the MIT dummy through institutions, it should not have a significant relationship when conditioning for these variables. However, as already mentioned in Section 3, it has to be taken into account that institutions are endogenous to our dependent variable and thus, our findings should also be interpreted with caution. Nonetheless, it might be a useful alternative with respect to our very small sample size. Columns (5) and (6) show that – as hypothesized – the settler mortality coefficient turns insignificant when adding $AVEXPR$. In addition, the average marginal effects of the settler mortality variable are reduced sharply (to around 0.02, s.e. = 0.04) and are also insignificant. In contrast, the coefficient and average marginal effect of our institutional measure are highly significant. Column (6) reveals that adding latitude does not change our results.

Up to now, our empirical results have shed some light on the importance of the institutional quality for the MIT phenomenon. We conducted IV probit regressions (using the average settler mortality as instrument) and the corresponding ordinary probit regressions as well as a "simplified" probit regression (analogous to the probit version of EL's model described in Section 3.2). Our main results can be summarized as follows: Analogous to AJR's findings for general cross-country comparisons, we find that institutions also seem to be decisive for the question of whether a country experiences an MIT. Our geographical control variables (latitude and continent dummies) are insignificant – not only in our two-stage probit estimates but also in our ordinary probit estimates.

In general, our findings hold for all our models, however, due to the very small sample size, more complex models such as the IV probit model are afflicted with several problems. Probably most important, the Wald test of exogeneity indicates no need for correcting for endogeneity regarding our institutional measure (which is in contrast to what theory suggests). Nonetheless, as all models, including the corresponding probit model presented in Panel C of Table 6 as well as the simplified probit model presented in Table 7 confirm our findings, we can conclude that good institutions, in particular the protection of private property rights, do play a decisive role regarding the MIT phenomenon. In the next subsections, we test whether our findings also hold when controlling for various other variables. Moreover, we test whether the legal origin or language variables are even more appropriate instruments for $AVEXPR$ than the settler mortality.

Table 7. Probit regression, directly controlling for the settler mortality (ARJ).

	Base sample 1 (1)	Base sample 1 (2)	Base sample 1 (3)	Base sample 1 (4)	Base sample 1 (5)	Base sample 1 (6)
Settler mortality (LOGEM4)	1.0814***	0.9984***	0.9961***	1.0213**	0.2116	0.2159
	(0.3238)	(0.3460)	(0.3816)	(0.4790)	(0.5044)	(0.5252)
	0.1827***	0.1559***	0.1595***	0.1462***	0.0159	0.0159
	(0.0353)	(0.0413)	(0.0499)	(0.0569)	(0.0373)	(0.0386)
Institutions (AVEXPR)					-1.7052**	-1.6525**
					(0.7840)	(0.7531)
					-0.1278***	-0.1218***
					(0.0414)	(0.0388)
Latitude		-2.3996		-3.9097		-1.4714
		(1.9928)		(2.4324)		(3.2785)
		-0.3747		-0.5596*		-0.1085
		(0.2954)		(0.3215)		(0.2393)
Latin America dummy			-0.0300	-0.9433		
			(0.8308)	(1.1671)		
			-0.0048	-0.1350		
			(0.1330)	(0.1629)		
Asian dummy			-0.5315	-1.2728		
			(0.7059)	(0.9590)		
			-0.0851	-0.1822		
			(0.1095)	(0.1298)		
Europe dummy			(a)	(a)		
Log likelihood	-13.1019	-12.3180	-12.7080	-11.2575	-5.9390	-5.8373
McFadden's Pseudo R^2	0.372	0.4102	0.3909	0.4604	0.7153	0.7202
Correctly classified	90.91%	90.91%	90.91%	90.91%	95.45%	95.45%
Heteroscedasticity test	0.1289	0.4299	0.2582	-	0.8901	0.6282
Number of observations	44	44	44	44	44	44

Note: Dependent variable: MIT dummy constructed analogously to the World Bank (2013) study, using Maddison (2010) data. See Appendix Table Al for detailed variable definitions and sources. The row 'Heteroscedasticity test' presents the p-value for the test for heteroscedasticity. We report both, the coefficients and, below, the average marginal effects. The respective standard errors are in parentheses. Significance at the 10, 5, and 1 percent levels are denoted by *, **, and ***, respectively. (a) These variables have been omitted due to perfect separation (due to the small sample size).

4.4 Alternative instruments

In this subsection we test alternative instruments for the average risk against expropriation and compare our results with those obtained using the settler mortality instrument.

First, we use the French legal origin of a country as instrument. One main advantage of this indicator is the better data availability, leading to an increased sample size. But also from a theoretical perspective, this choice of instrument seems reasonable: For example, La Porta et al. (1999) show that the differences between legal systems implemented by the colonial powers were important for the development of institutions (especially the protection of investor rights) and thus, on the long-term development of a country. In particular, they find that countries with a British common law origin are less interventionist, have a more efficient government and more political freedom as well as a better public good provision than countries with a French civil law origin (see also Section 3.1). Glawe and Wagner (2017c) analyze this relationship for the phenomenon of the MIT. Using simple hypothesis testing, they show that regarding the question of whether a country faces an MIT, the negative influence of a French legal origin seems to persist, whereas the positive effect of a former British legal origin seems to fade out. Nonetheless, we also checked the suitability of the British legal origin as instrument. We obtain similar results compared to our $LEGOR\_FR$ instrument, however, the F-stat is around 7 which indicates weak instruments.[20] In addition, we also used the British and French colonial origin as instruments. With respect to the British dummy, we obtain similar results regarding the institutional effect, however, a very low F-stat (below 2 for the British colonial origin dummy). Regarding the French colonial origin dummy, we have a perfect separation problem.

We decided to report our estimates for the French legal origin instrument due to the stronger instrument fit (that is, the higher F-stat); however, our results regarding the British legal origin instrument confirm these findings. For example, we obtain very similar results regarding the significance-levels as well as the average marginal effects. Our results are presented in Table 8.

In general, our previous findings (when using settler mortality as instrument) are confirmed, that is we find a strong relationship between the institutional measure and the MIT dummy. Nonetheless, some differences become apparent. First of all, although the institutional coefficients remain highly significant, they are slightly lower than those reported in Table 6. Moreover, regarding the first-stage regressions reported in Panel B, latitude has a significant impact on the institutional quality (at the 5-percent level).[21] Regarding the diagnostic statistics, two aspects seem noteworthy: First, we have extremely lower p-values regarding the Wald test of exogeneity (0.0889 and 0.1066 when adding latitude). Second, the F-stats are slightly lower, however, (as before) they are still above the threshold of 10 for Columns (1) and (2). When adding continent dummies in Columns (3) and (4), our institutional coefficients remain significant at the 1-percent level (remind that their significance dropped to the 5-percent level when using the settler mortality instrument). As in Table 6, the F-stats drop below the threshold of 10 when adding continent dummies and the p-value of the Wald test of exogeneity increases significantly (Note that although our sample size is slightly increased to

[20] Note that the p-value of the Wald test of exogeneity is slightly lower compared to the p-value we obtain when using the $LEGOR\_FR$ instrument.

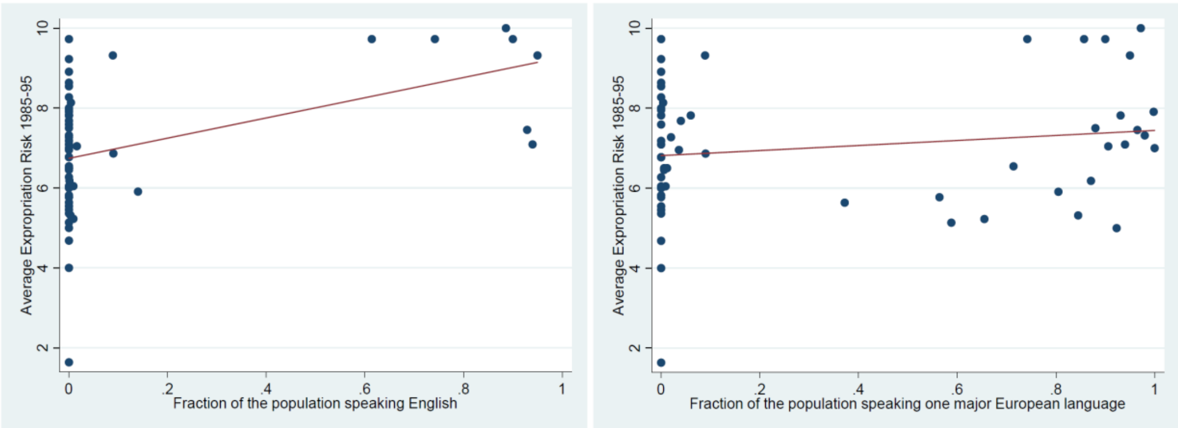
[21] This positive relationship becomes insignificant when adding continent dummies.

60 when using the French legal origin IV, this is still relatively small for such an econometric model). In all probit models reported in Panel C, the Pseudo-R-squared is relatively high (ranging from 0.56 to 0.61), indicating a good model fit (however, in Table 6, we report even higher values) and the estimates are found to be homoscedastic. Again, our coefficients are lower than when using the settler mortality instrument.

Next, as suggested by Hall and Jones (1999), we use the fraction of the population speaking English ($ENGFRAC$) and the fraction of the population speaking one of the five major European languages ($EURFRAC$) as instruments. Figure 4 illustrates the relationship between the language variables and current institutions for base sample 3.

Figure 4. First-stage relationship between settler mortality and language variables.

(a) ENGFRAC **(b)** EURFRAC



It is striking that while there is a clear positive relationship between $ENGFRAC$ and $AVEXPR$, there is only a very weak relationship between $EURFRAC$ and $AVEXPR$. Thus, it is not surprising that Columns (5) and (6) of Panel B reveal an insignificant first-stage relationship between $EURFRAC$ and $AVEXPR$ (the sign is even negative). In contrast, there is a strong first-stage relationship between $ENGFRAC$ and $AVEXPR$. Regarding the corresponding two-stage probit estimates reported in Panel A, the institutional coefficient is highly significant, so are the average marginal effects. The Wald test of exogeneity yields a p-value of 0.0550 and 0.0668 when adding latitude, indicating that we need to correct for endogeneity. The F-stat is below 10, and thus, our results suffer from the presence of weak instruments. The overidentification test reveals that (with p-values of 0.69 and 0.70) we fail to reject the null hypothesis that all instruments are exogenous.[22, 23]

[22] We also tried to focus solely on the $ENGFRAC$ variable as instrument; however, when adding continent dummies the institutional measure turns insignificant (the continent dummies are insignificant) while when additionally adding latitude, it stays significant at the 1-percent level. In both cases, the average marginal effects for the institutional measure are as well insignificant. Therefore, $ENGFRAC$ does not appear to be the best choice as instrument.

[23] We also used the ethnolinguistic fragmentation as an instrument. In general, our results are confirmed, however the F-stat is relatively low, indicating weak instruments.

Table 8. IV regression – alternative instruments (ARJ).

	Base sample 2 (1)	Base sample 2 (2)	Base sample 2 (3)	Base sample 2 (4)	Base sample 3 (5)	Base sample 3 (6)
Panel A: Two-Stage Probit						
Institutions (AVEXPR)	-1.2827*** (0.2756)	-1.2475*** (0.2739)	-1.3696*** (0.3253)	-1.3110*** (0.3162)	-1.4042*** (0.3123)	-1.3819*** (0.3109)
	-0.1823*** (0.0161)	-0.1907** (0.0207)	-0.1719*** (0.0353)	-0.1637*** (0.0521)	-0.1975*** (0.0175)	-0.2053*** (0.0187)
Latitude		0.9883 (1.5889)		-0.8677 (2.7688)		1.1523 (1.5499)
		0.1511 (0.2605)		-0.1084 (0.3229)		0.1712 (0.2422)
Continent dummies	No	No	Yes	Yes	No	No
Wald test of exogeneity	0.0889	0.1066	0.2920	0.4013	0.0550	0.0668
F-stat	16.96	10.94	5.09	4.39	6.94	8.30
Log likelihood	-120.03111	-115.4266	-117.79283	-114.21515	-114.7586	-107.9777
Overidentification test	-	-	-	-	0.70	0.69
Panel B: First-Stage for Average Risk Against Expropriation						
LEGOR_FR	-1.5066*** (0.3598)	-1.1704*** (0.3680)	-1.356*** (0.4251)	-1.2261*** (0.4204)		
ENGFRAC					2.9125*** (0.7772)	2.2511*** (0.7263)
EURFRAC					-0.5807 (0.5051)	-0.1968 (0.4752)
Latitude		2.5913** (1.1026)		2.1018 (1.3452)		3.1563*** (1.0938)

Table 8 continued.

	Base sample 2 (1)	Base sample 2 (2)	Base sample 2 (3)	Base sample 2 (4)	Base sample 3 (5)	Base sample 3 (6)
			Panel C: Probit Regression			
Institutions (AVEXPR)	-1.2527***	-1.1639***	-1.2866***	-1.1676***	-1.2953***	-1.2325***
	(.3269)	(.3377)	(0.3841)	(0.3730)	(0.3444)	(0.3615)
	-0.1587***	-0.1467***	-0.1471***	-0.1292***	-0.1623***	-0.1554***
	(0.0152)	(0.0237)	(0.0239)	(0.0259)	(0.0165)	(0.0247)
Latitude		-1.0524		-2.3086		-0.4528
		(1.6143)		(2.3203)		(1.7358)
		-0.13261		-0.2554		-0.0571
		(0.2017)		(0.2499)		(0.2187)
McFadden‘s Pseudo R^2	0.5691	0.5599	0.6065	0.6066	0.5845	0.5697
Correctly classified	90.16%	91.67%	90.16%	77.05%	89.66%	75.86%
Log likelihood	-14.1577	-13.8010	-12.9289	-12.3378	-13.3184	-13.1702
Heteroscedasticity test	0.6616	0.9394	0.7284	0.8011	0.7596	0.9527
Number of observations	61	60	61	60	58	57

Note: Dependent variable: MIT dummy constructed analogously to the World Bank (2013) study, using Maddison (2010) data. See Appendix Table Al for detailed variable definitions and sources. ‘F-stat’ is the F-statistic for the first-stage regression. The row ‘Wald test of exogeneity’ presents the p-value of the Wald test of exogeneity for the suspected endogenous variable and the row ‘Heteroscedasticity test’ presents the p-value for the test for heteroscedasticity. In panels A and C, the average marginal effects are reported below the respective coefficients. Standard errors are in parentheses. Significance at the 10, 5, and 1 percent levels are denoted by *, **, and ***, respectively. In Panel B and C, we only report the results regarding the average risk against expropriation and latitude to save space (the other regressors are insignificant).

4.5 Robustness checks – additional controls

In this section, we test the validity of our results obtained in Sections 4.3 and 4.4 by controlling for various other variables that could be correlated with the settler mortality or the MIT dummy. As AJR, we find that adding these controls in most cases has only little effect on the impact of institutions. Regarding our base sample 1 (using the settler mortality as instrument), the coefficients and average marginal effects of the institutional measure stay significant at the 1- to 5-percent level, in our larger base sample 2 (using the legal origin as instrument) even at the 1-percent level.

As AJR suggest, we control for various geographical characteristics that could be correlated with the settler mortality. In Columns (1) to (4) of Table B1 (for base sample 1) and Table B3 (for base sample 2) in the Appendix B, we include variables regarding the temperature[24], and the oil reserves. These variables are all insignificant, whereas the effect of institutions on the MIT dummy stays significant (at the 1- and 5-percent level, depending on the base sample used).[25]

We then control for the disease environment. Including a yellow fever dummy (indicating whether yellow fever epidemics occurred before 1900) has only little effect on institutions for both base samples (Column 5 and 6 in Tables B1 and B3, respectively). Regarding the malaria variable, due to perfect separation, we can only report the results for base sample 2 (Columns 7 and 8 of Table B3). We find that while the coefficients of institutions and the corresponding significance levels remain nearly unchanged, the average marginal effect (still significant at the 1-percent level) drops to -0.08 (s.e. = 0.01) and -0.10 (s.e. = 0.02) when adding latitude; however the malaria variable itself is insignificant.[26]

Also controlling for cultural variables such as the ethnolinguistic fragmentation (Columns (1) and (2) of Tables B2 and B4, respectively) or religion (Columns (3) to (4) of Tables B2 and B4, respectively) has only little effect on our results. In Columns (5) and (6) of Tables B2 and B4, we control for the "fraction of the population of European descent" as Europeans might have brought along a 'European culture' (AJR, p. 1390). However, this variable is also insignificant, whereas the coefficient and average marginal effects of the institutional measure remains significant at the 5-percent level for base sample 1 and at the 1-percent level for base sample 2.

In Section 4.4, we have already discussed the importance of the colonial origin as well as the legal origin for institutional development. Among others, we have checked whether these variables might be also appropriate instruments. Indeed, the French legal origin has proven to be an adequate instrument in our case. With respect to the British colonial origin

---

[24] Temperature is the first principal component of the five temperature variables compiled by Parker (1997).

[25] For our base sample 2, we additionally test humidity variables (namely, morning minimum and maximum humidity, as well as afternoon minimum and maximum humidity) and soil quality variables (steppe, low latitude; desert, low latitude; steppe, middle latitude; desert, middle latitude; dry steppe wasteland, and desert dry winter), both compiled by Parker (1997). (Note that our base sample 1 is too small for including these variables.) The p-value for the joint significance of humidity variables is 0.6187 and 0.1828 when adding latitude, whereas the coefficients and average marginal effects of the institutional measure stay significant at the 1-percent level. The p-value for the joint significance of soil quality variables is 0.3396 and 0.5227 when adding latitude, whereas the coefficients and average marginal effects of the institutional measure stay significant at the 1-percent level.

[26] Note that some researchers (among them ARJ) argue that malaria (in particular, the percentage of the population living where falciparum malaria is endemic in 1994) is endogenous because poor countries with weak institutions have not managed to eradicate malaria.

dummy, we obtained very low F-stats and regarding the French legal origin dummy, we have a perfect separation problem. Analogous to AJR, we also control for these variables in Tables B2 and B4. Unfortunately, we can only include the British colonial dummy in our equation, as regarding the other variables, we (again) have a perfect separation problem in both samples (and of course, we cannot include the French legal origin dummy in our base sample 2 as this variable is our instrument). With respect to the base sample 1, the institutional coefficients turn insignificant, while the average marginal effects stay significant at the 1-percent level (and at the 10-percent level when including latitude). The British colonial origin dummy is insignificant in Column (7), however, when adding latitude in Column (8), it becomes significant at the 10-percent level. Strikingly, although AJR also found a marginally positive effect of the British dummy, their institutional coefficients stayed relatively unchanged. Regarding the base sample 2, including the British dummy does not affect our results regarding the institutional measure. As before, the coefficient of the British colonial dummy is slightly significant when adding latitude (in addition, the average marginal effect of the dummy becomes significant at the 5–percent level).

Finally, some findings regarding the first-stage relationships seem to be noteworthy: the fraction of the population of European descent has a significant impact in determining the quality of institutions (at the 1-percent level for the base sample 2 and at the 5-percent level for the base sample 1), supporting the findings of Section 3. In addition, we have additional significant first-stage relationships regarding the base sample 2: For example, the oil reserves and temperature variables are negatively correlated with the institutional measure. The same applies to the ethnolinguistic fragmentation variable, indicating that we might have as well a kind of "cultural channel". However, as Alesina and Giuliano (2014) note, culture and institutions actually interact, and thus, we have to be careful when making statements on the causality/direction of this relationship. Nonetheless, it appears to be an interesting observation because this potential channel is not identified by AJR. Last but not least, there is a significant negative (!) relationship between a former British origin and the average expropriation risk. Compared with the standard literature, the British dummy has the "wrong" sign, however, as Glawe and Wagner (2017c) argue, the positive effect of being a former British colony fades out and is not relevant for the question of whether a country falls into an MIT (or not).

Overall, our results stay relatively robust also when adding various control variables. In general, our findings of this and the previous sections suggest that colonial and legal origin variables might play a more important role for analyzing the MIT phenomenon than for general cross-country growth regressions. Moreover, although the findings of AJR seem in general also to apply for the MIT, we identify various significant first-stage relationships that are not identified by AJR (especially regarding our base sample 2).

4.6 Further extensions

A subsequent article of Acemoglu and co-authors (Acemolgu et al., 2014) examines the relationship between institutions, human capital and development. This study is an interesting extension of AJR and also investigates issues mentioned in Section 3. Acemoglu et al. (2014) find that when focusing on the historically-determined differences in human capital, once controlling for institutions, the human capital estimates are reduced significantly while the effect of institutional quality is robust. The authors use a two-stage least squares model with two endogenous variables (human capital and institutions). For each endogenous variable, they use two instruments. In addition, Acemoglu et al. (2014) also perform semi-structural models and simple OLS estimation. We also tried to adopt this analysis to the MIT phenomenon, however, the necessary data is only available for 37 countries of our sample (even the sample of Acemoglu et al., 2014, with 62 observations, is rather small). Thus, a two-stage least squares model and also the semi-structural models unfortunately are far too complicated for our data. However, in the previous sections we have found that our probit and two-stage probit estimates are very similar (especially when using settler mortality as instrument), and the results of the Wald test of exogeneity (often) do not support the theoretical claim of endogeneity of institutions. Thus, we estimate an ordinary probit model. Our results are presented in Table B5 in the Appendix B. Column (1) shows the bivariate relationship between average years of schooling and the MIT dummy. The relationship is highly significant with a coefficient of -0.71 (s.e. = 0.26) and an average marginal effect of -0.07 (s.e. = 0.02). Moreover, as reported in Column (2), there is also a very strong bilateral correlation between the institutional measure and the MIT dummy with a coefficient of -1.57 (s.e. = 0.45) and an average marginal effect of -0.15 (s.e. = 0.03). That is, an infinitesimal rise in the institutional measure (average years of schooling) reduces the probability of a country to experience an MIT by around 15 percent (7 percent). Controlling for latitude and continent dummies has only little effect on our results (see Columns 3-8). In Column (9), we simultaneously include human capital and institutions in the regression. In contrast to Acemoglu et al. (2014), the average years of schooling variable turns insignificant and the average marginal effect is reduced to -0.03 (s.e. = 0.02). The average marginal effect of the institutional measure (-0.08, s.e. = 0.03) is still significant at the 5-percent level (and the coefficient at the 10-percent level). Again, our results are robust for adding latitude and continent dummies.[26] Indeed, the institutional and human capital coefficients and average marginal effects are slightly higher (and the corresponding p-values are slightly reduced) when adding latitude (eventually reflecting the indirect effect of geography). As Acemoglu et al. (2014) note, such a model (without instruments) is afflicted with several problems, leading to, among others, an upward bias of the human capital estimate and downward bias of the institutional estimate. However, even our ordinary probit models suggest that human capital has a minor role compared to the institutional quality (especially regarding the significance of the coefficients and the average marginal effects), thus confirming Acemoglu et al.'s (2014) two-stage least squares estimation results.

5 The primacy of institutions over geography and integration for the MIT? – Rodrik et al. (2004) revisited

To our knowledge, RST are the first who take into account the three deep determinants simultaneously, thus extending AJR's analysis by adding the integration/international trade dimension proposed by Frankel and Romer (1999). They find that institutional quality "trumps everything else" (p. 135). We test whether this also applies to the MIT phenomenon. In particular, we examine whether the inclusion of international trade as a third potential channel

[26] When using an extended sample which is not restricted to former colonies, the institutional measure is significant at the 1-percent level (the average years of schooling are significant at the 10-percent level). Including latitude or continent dummies does not change these results.

changes the results of the previous section, especially regarding the importance of institutions for the MIT.

5.1 Data and descriptive statistics

In Table 9, we extend the descriptive statistics provided in Table 4 in Section 4.1 with the integration measure (namely the actual trade share in GDP, $ATRADE$) and the respective instrument (the trade share in GDP constructed on the basis of a gravity equation for bilateral trade flows, $CTRADE$). Column (1) presents the statistics for the large sample and Column (2) covers our base sample 1 of 44 countries for which we have data on settler mortality and which are former colonies. The remaining column is for the larger base sample 2 for which we have data on alternative instrument, namely the French legal origin. We do not use the base sample 3 with the two language variables as instruments because they turned out to be inadequate instruments in our case (see also discussion in Section 4.4).

Table 9. Descriptive statistics (III).

	Large sample	Base sample 1	Base sample 2
MIT dummy	0.68 (0.47)	0.812 (0.39)	0.77 (0.42)
Average protection against expropriation risk (AVEXPR)	7.45 (2.74)	6.89 (1.43)	7.00 (1.60)
Distance from the equator (Latitude)	0.31 (0.18)	0.21 (0.14)	0.25 (0.17)
Settler Mortality (LOGEM4)	–	4.29 (1.07)	–
French legal origin dummy (LEGOR_FR)	0.13 (0.33)	0.57 (0.50)	0.54 (0.50)
Actual trade share (ATRADE)	66.14 (45.04)	58.06 (49.48)	62.43 (48.17)
Constructed trade share (CTRADE)	21.15 (15.96)	14.82 (9.66)	19.21 (14.81)
Number of countries	76	44	61

Notes: Variable definitions and sources are provided in Appendix A. Standard errors are in parentheses.

5.2 Probit regression

Estimation strategy

Our regression model presented in Section 4.3 is slightly modified to include a measure of integration (analogous to RST):

(7) $$P(MIT_i = 1 | INS_i, INT_i, GEO_i) = \Phi(\lambda_1 + \lambda_2 INS_i + \lambda_3 INT_i + \lambda_4 GEO_i),$$

where INT_i is the measure of integration, in particular the actual trade share ($ATRADE$) proposed by Frankel and Romer (1999), INS_i is the institutional measure ($AVEXPR$) and GEO_i is the measure of geography, in particular latitude (the latter replaces X_i, that is our vector of other covariates used in Section 4). In the following, we follow RST and use standardized measures of our three regressors, which enables us to directly compare the estimated coefficients.[27]

Table 10. Probit regression, large sample and base sample 2 (RST).

	Large sample (1)	Large sample (2)	Large sample (3)	Base Sample 2 (4)	Base Sample 2 (5)	Base Sample 2 (6)
Geography (Latitude)	-1.0646*** (-4.92) -0.2433*** (-10.78)	-0.1747 (-0.53) -0.0159 (-0.53)	-0.2544 (-0.68) -0.0226 (-0.68)	-0.7779*** (-3.18) -0.1711*** (-4.04)	-0.2716 (-0.77) -0.0269 (-0.78)	-0.3832 (-0.94) -0.0364 (-0.97)
Institutions (AVEXPR)		-2.4190*** (-4.06) -0.2198*** (-5.52)	-2.3073*** (-3.73) -0.2047*** (-4.48)		-2.2103*** (-3.27) -0.2189*** (-5.53)	-2.0517*** (-2.95) -0.1949*** (-3.99)
Integration (ATRADE)			-0.1494 (-0.56) -0.0133 (-0.57)			-0.1884 (-0.69) -0.0179 (-0.71)
McFadden's Pseudo R^2	0.3632	0.7311	0.7347	0.2221	0.6409	0.6502
Correctly classified	84.21%	94.74%	94.74%	81.82%	94.55%	94.55%
Log likelihood	-30.1806	-12.7453	-12.5759	-21.4090	-9.8823	-9.6267
Heteroskedasticity test	0.1191	0.3869	0.4613	0.2903	0.5581	0.6351
Number of observations	76	76	76	55	55	55

Note: Dependent variable: MIT dummy constructed analogously to the World Bank (2013) study, using Maddison (2010) data. See Appendix Table Al for detailed variable definitions and sources. We report both, the coefficients and, below, the average marginal effects. Z-statistics are in parentheses. Significance at the 10, 5, and 1 percent levels are denoted by *, **, and ***, respectively.

Probit estimation results

Our estimation results for our base sample 2 and our large sample are presented in Table 10. Unfortunately, we have a perfect separation problem when using our base sample 1 ($AVEXPR$ predicts the MIT perfectly).

Column (1) reveals that in our large sample, there is a strong correlation between our geographical measure (latitude) and the MIT dummy; however, latitude turns insignificant if we additionally include our institutional measure (see Column 2). In addition, Column (2) shows a highly significant relationship between institutions and the MIT dummy with a coefficient of -2.42 and an average marginal effect of -0.22. Also adding the measure of integra-

[27] The standardized variable x^* is obtained by using the following formula: $x^* = \frac{x-\mu}{sd}$, where x denotes the original variable and μ (sd) is the mean (standard deviation) of x.

tion ($ATRADE$) in Column (3) does not change our results; the institutional effect is still highly significant with a coefficient of -2.30 and an average marginal effect of -0.20. We obtain very similar results for our base sample 2 (Columns 4–6). With regard to the strong correlation between institutions and the dependent variable, our findings are in line with RST. However, RST find that the signs of geography and integration are also statistically significant or close to (p. 139). Columns (3) and (6) of Table 10 reveal that in our case, these variables are insignificant (we actually get relatively high p-values).

5.3 IV regression model

Estimation strategy

As outlined in the previous sections, there are various reasons why the above relationship between institutions/trade and the MIT dummy cannot be interpreted as causal (e.g., the reversed causality problem). Therefore, we employ a two-stage probit model. As already mentioned, we follow Frankel and Romer (1999) by using the constructed trade share as an instrument for integration. Regarding the instrument for institutions, due to our perfect separation problem, we cannot use the instrument proposed by AJR (namely the settler mortality rate), but we have to rely on the French legal origin which we identified as an adequate instrument in the previous section. One advantage of this instrument is that it increases our sample size from 41 to 55. The first-stage regressions are given by equations (9) and (10). Our full model is given by equations (8)–(11):

(8) $MIT_i^* = \lambda_1 + \lambda_2 INS_i + \lambda_3 INT_i + \lambda_4 GEO_i + \sigma_i,$

(9) $INS_i = \mu_1 + \mu_2 LEGOR\_FR_i + \mu_3 CTRADE_i + \mu_4 GEO_i + \varepsilon_{INS_i},$

(10) $INT_i = \varphi_1 + \varphi_2 CTRADE_i + \varphi_3 LEGOR\_FR_i + \varphi_4 GEO_i + \varepsilon_{INT_i},$

(11) $MIT_i = 1[MIT_i^* > 0].$

where $\varepsilon_i = (\varepsilon_{INS_i}, \varepsilon_{INT_i})$ and $(\sigma_i, \varepsilon_i) \sim N(0, \sum)$. $LEGOR\_FR_i$ denotes the French legal origin dummy and $CTRADE_i$ is the constructed trade share. $(\sigma_i, \varepsilon_i)$ is independent of $LEGOR\_FR_i$, $CTRADE_i$, and GEO_i. Equation (8), along with equation (11), is the structural equation. Equations (9) and (10) are a reduced form for INS_i and INT_i, respectively.

IV estimation results

Our IV estimation results are presented in Table 11. Column (1) of Panel B reveals a strong first-stage relationship between the French legal origin dummy and current institutions. In addition, latitude (i.e., our geography variable), has a significant impact in determining the quality of institutions. The corresponding two-stage probit estimates are reported in Panel A:[28] In Column (1), the institutional coefficient (-2.40) is highly significant, confirming the results presented in Section 4.3. In contrast, our geography measure is insignificant. The F-stat for our first-stage regression is larger than the Staiger and Stock’s rule of thumb of 10 and thus, our results do not suffer from the presence of weak instruments. The Wald test of exogeneity yields a p-value of 0.64, thus not supporting the theoretical claim of endogeneity of institu-

[28] As before, we report the average marginal effects below the respective coefficients.

tions. Therefore, the results of our probit regression reported in Panel C may be better suited (indeed, the coefficients and average marginal effects do not vary very much between Panel A and C, indicating that the reverse causality bias is not very large). The test for heteroscedasticity yields $\chi^2(1) = 0.56$, suggesting that our probit estimates are homoscedastic. In addition, the Pseudo-R-squared is relatively high (0.64), indicating a good model fit. Before we test all three deep determinant regressors simultaneously in Column (3), we first only include the institutional and integration measure in Column (2). Our results regarding the institutional measure remain relatively unchanged and are still highly significant, while the coefficient and average marginal effect of the integrational measure are insignificant. Regarding the first-stage relationship, the integrational variable has no impact on the institutional measure. The F-stat for the trade share is relatively high (24.25), however, the F-stat for the institutional measure drops below the threshold of 10. We perform an AR-Test for testing the robustness of the weak instrument which yields a p-value of 0.02 (see also the two graphs of the 95 percent confidence set and the rejection surface, respectively, in the Appendix C, Figure C1).

Table 11. IV regression – including the integration channel (RST).

	Base sample 2 (1)	Base sample 2 (2)	Base sample 2 (3)
Panel A: Two-Stage Probit			
Institutions (AVEXPR)	-2.3982*** (-3.78) -0.2555*** (-2.75)	-2.3787*** (-3.40) -0.2373*** (-4.53)	-2.1370** (-2.16) -0.1999* (-1.81)
Integration (ATRADE)		-0.2367 (-0.69) -0.0236 (-0.68)	-0.2586 (-0.73) -0.0242 (-0.78)
Geography (Latitude)	-0.0806 (-0.15) -0.0086 (-0.16)		-0.3108 (-0.47) -0.0291 (-0.51)
Wald test of exogeneity	0.6409	0.6533	0.8723
Log likelihood	-103.6549	-384.2990	-369.0907
AR Test (p-value)	0.0425	0.0229	0.1299

Panel B: First-Stage for AVEXPR/ATRADE					
Column	(1)	(2)	(2)	(3)	(3)
Dependent variable	AVEXPR	AVEXPR	ATRADE	AVEXPR	ATRADE
LEGOR_FR	-0.3417*** (-3.04)	-0.4619*** (-4.20)	-0.1915* (-1.85)	-0.3412*** (-3.03)	-0.2687** (-2.51)
CTRADE		0.0210 (0.18)	0.7605*** (6.84)	0.0121 (0.11)	0.7717*** (7.24)
LATITUDE	0.3150*** (2.61)			0.3145*** (2.60)	-0.2690** (-2.34)
F-Stat	11.7	8.40	24.45	7.70	19.00

Table 11 continued.

	Base sample 2 (1)	Base sample 2 (2)	Base sample 2 (3)
Panel C: Probit Regression			
Institutions (AVEXPR)	-2.2103***	-2.3700***	-2.0517***
	(-3.27)	(-3.59)	(-2.95)
	-0.2189***	-0.2375***	-0.1949***
	(-5.53)	(-7.44)	(-3.99)
Integration (ATRADE)		-0.1128	-0.1884
		(-0.40)	(-0.69)
		-0.0113	-0.0179
		(-0.40)	(-0.71)
Geography (Latitude)	-0.2716		-0.3832
	(-0.77)		(-0.94)
	-0.0269		-0.0364
	(-0.78)		(-0.97)
McFadden's Pseudo R^2	0.6409	0.6482	0.6502
Correctly classified	94.55%	94.64%	94.55%
Log likelihood	-9.8823	-10.2365	-9.6267
Heteroskedasticity test	0.5581	0.6956	0.6351
Number of observations	55	56	55

Note: Dependent variable: MIT dummy constructed analogously to the World Bank (2013) study, using Maddison (2010) data. See Appendix Table Al for detailed variable definitions and sources. 'F-stat' is the F-statistic for the first-stage regression. The row 'Wald test of exogeneity' presents the p-value of the Wald test of exogeneity for the suspected endogenous variable and the row 'Heteroscedasticity test' presents the p-value for the test for heteroscedasticity. In panels A and C, the average marginal effects are reported below the respective coefficients. Z-statistics are in parentheses. Significance at the 10, 5, and 1 percent levels are denoted by *, **, and ***, respectively.

In Column (3), we finally include all three variables simultaneous. The institutional coefficient in the two-stage estimation (Panel A) decreases slightly and is significant at the 5-percent level. In contrast to RST, who find that the coefficients of geography and integration have the wrong sign when simultaneously including all three deep determinants, our coefficients all have the expected negative sign. The first-stage relationships reported in Panel B confirm our previous results of Columns (1) and (2): Geography has a significant impact in determining the institutional quality, while integration has not. The latter finding is in contrast to the results of RST who found a significant positive impact of integration on the quality of institutions (at the 5-percent level). Moreover, Column (3) of Panel B reveals that geography has a significant negative impact on trade. In RST, this relationship is not significant. Another finding of Panel B seems to be noteworthy as well: In Column (3), the French legal origin dummy has a negative impact on integration (at the 5-percent level), whereas in Column (2), without including latitude, there is no such effect.

Regarding the diagnostic statistics, as before, the F-stat of the institutional measure drops below the threshold of 10 and this time, the AR-Test results in a p-value of 0.13 (see also the corresponding graphs of the AR Test in the Appendix C, Figure C2). However, as the Wald test of exogeneity suggest that there is no need for an endogeneity correction (the p-value is above 0.87), the probit estimates should deliver us more insightful results. Indeed, the institutional coefficient reported in Panel C is highly significant, but slightly slower than that reported in Panel A. Thus, (in combination with the diagnostic statistic results described above) the two-stage probit estimation is probably slightly upward biased.

Due to the fact that our results regarding the institutional measure stay relatively unchanged when including integration, we refer to the Section 4.5 for a detailed discussion of robustness checks.[29]

We briefly discuss the inter-relationships between institutions and integration, by (a) regressing trade and geography on institutions and (b) regressing institutions and geography on trade. Our results are presented in Table 12. Panel A presents the OLS regressions and Panel B the 2SLS regressions (the first-stages are provided in Panel C). We find that institutional quality and integration have no significant positive effect on each other at the 5-percent level (RST find a positive effect of institutional quality on trade at the 5-percent level). As RST, we find a highly significant positive effect of geography on institutional quality and a slightly less significant negative impact of geography on trade (at the 5-percent level).

Overall, our results are in line with the findings of RST in the sense that institutional quality "trumps" everything else. In addition, our *first-stage relationships* reveal that there is a significant positive effect of latitude (i.e., the geographical variable) on institutional quality. However, in contrast to RST, our geography variable also has a significant impact in determining integration while we do not find a significant positive effect of integration on institutional quality. Regarding the *inter-relationships*, our findings suggest that there is an indirect positive effect of geography on the probability of experiencing an MIT via institutional quality (at the 1-percent level) and integration (at the 5-percent level). Moreover, in contrast to RST, we find no significant positive effect of institutional quality on integration.

[29] We also repeated the analysis with the Rule of Law index of the Worldwide Governance Indicators dataset initiated by Daniel Kaufmann and Aart Kraay that is also used by RST to increase the sample size. However, regarding our sample, using this index instead of the average expropriation risk does not increase our sample size when using the settler mortality as instrument (the sample size is actually slightly reduced). When using the Rule of Law index and instrument it with the French legal origin, the results of Column (1) and (2) of Table 11 are confirmed, however, regarding Column (3), convergence is not achieved.

Table 12. Inter-relationships between institutional quality and integration (RST).

	(1)	(2)
Panel A: OLS Regression		
Dependent variable	Institutions (AVEXPR)	Integration (ATRADE)
Geography (Latitude)	0.4807*** (4.08)	-0.3258* (-1.96)
Institutions (AVEXPR)	–	0.4560*** (2.76)
Integration (ATRADE)	0.2804*** (2.76)	–
R-square	0.2989	0.1376
Panel B: 2SLS Regression		
Dependent variable	Institutions (AVEXPR)	Integration (ATRADE)
Geography (Latitude)	0.4503*** (3.70)	-0.5120** (-2.03)
Institutions (AVEXPR)		0.8730* (1.95)
Integration (ATRADE)	0.0321 (0.21)	
R-square	0.2182	
Number of observations	55	55
Panel C: First-Stage for Institutions and Trade		
Dependent variable	Integration (ATRADE)	Institutions (AVEXPR)
LEGOR_FR		-0.3417*** (-2.96)
CTRADE	0.7819*** (6.76)	
Geography (Latitude)	-0.1663 (-1.42)	0.3150** (2.54)
F-statistic	45.723	8.73
R-square	0.4738	0.3116
Test for endogeneity (p-value)	0.0184	0.2873

Note: Dependent variable: Log per capita GDP. The independent variables are all scaled in the sense that they present deviations from the mean divided by the standard deviation. T-statistics are in parentheses. Significance at the 10, 5, and 1 percent levels are denoted by *, **, and ***, respectively.

6 Summary

In Section 3 to 5, we have examined the relationship between the deep determinants and the MIT phenomenon. Our main findings can be summarized as follows:

Our first analysis based on the study of EL revealed that the fraction of Europeans in colonial population is strongly negatively correlated with our MIT dummy. Controlling for various country characteristics such as legal origin, independence or ethnicity does also not change our results significantly. Ethnolinguistic heterogeneity (that is found to be highly significant in EL's regression) has actually no direct effect on the MIT dummy (however, we later identified an indirect effect via institutional quality). Our findings also differ regarding the potential channels that might connect euroshare and the MIT dummy: Most importantly, our findings do not support a human capital channel, suggesting that this positive effect may have already faded away when a country has reached the MIR. In contrast, our results confirm the existence of an institutional channel. We additionally tested for a third channel (integration/trade); however we found no empirical evidence for its existence. As this analysis is based on relatively simple estimation methods (EL use OLS estimation and we, due to our binary outcome variable, a probit model), we have to be careful regarding statements about the causality of the relationships.

In the next two sections, focusing on the studies conducted by AJR and RST, we estimated more sophisticated econometric models that provided us with additional insights regarding the reverse feedback channels. Moreover, these models take into account the criticism of Acemolgu et al. (2014) who argue that "empirical models that treat institutions […] as exogenous are misspecified".

Our empirical results based on AJR's study show that good institutions, in particular the protection of private property rights, play a decisive role regarding the question of whether a country faces an MIT. Furthermore, our robustness check revealed that adding geographical and cultural control variables has only little effect on our results. Regarding these aspects, AJR's and our study derive very similar results. However, (in contrast to AJR) the first-stage regressions show that that the fraction of the population of European descent has a significant impact in determining the quality of institutions for both base samples (supporting our findings of Section 3). Regarding the base sample 2, the oil reserves and temperature variables variable are negatively correlated with the institutional measure. The same applies to the ethnolinguistic fragmentation variable, indicating that we might have as well a kind of "cultural channel". Moreover, we find a significant positive effect of latitude on institutional quality.

Our final analysis adapted the study of RST, thus extending the previous analysis by the integration/trade channel. In general, our results are in line with the findings of RST in the sense that institutional quality "trumps" everything else. We also find that latitude (i.e., the geographical variable) has a significant positive impact in determining institutional quality and a significant negative impact on integration. However, we do not find a significant positive indirect effect of institutions on income via integration.

7 Conclusion

So far, very little has been done to examine the relationship between the deep determinants and economic performance at more subtle stages of economic development. Our paper has contributed to close this gap by focusing on the MIT phenomenon. In particular, we have applied and modified/extended various empirical analyses of the general literature. We especially focused on the studies conducted by Acemoglu et al. (2001), Rodrik et al. (2004), and Easterly and Levine (2016) which have all been published in highly ranked journals.

According to our analysis, the deep determinants, especially institutional quality, appear to play an important role for overcoming the MIR without falling into an MIT. In contrast, the integration channel does not seem to play a major role for the middle-income transition (when conditioning on institutions). From this perspective, the findings of the general literature seem also to hold for the MIT phenomenon. However, our analysis has also revealed some differences, for example regarding the transmission channels, the first-stage relationships, and the inter-relationships between integration and institutions. For instance, our findings do not support the existence of a human capital channel proposed by Easterly and Levine (2016) (Section 3.2). Moreover, in contrast to Acemoglu et al. (2001), we find statistically significant (negative) first-stage relationships between some cultural as well as geographical variables and institutions (Section 4.5). Finally, unlike Rodrik et al. (2004), we do not find a statistically significant positive effect of institutional quality on integration regarding the inter-relationship between institutions and trade (Section 5.2). This may raise the question whether we need a modified version of the deep determinants regarding the MIT phenomenon, taking into account the specific challenges and characteristics associated with the middle-income transition. Subsequent research should more intensively investigate the MIT-deep determinants relationship, especially regarding the indirect channels and reverse feedbacks. However, further research does not necessarily have to be restricted only to the MIT phenomenon, but could also be extended to other subtle stages of economic development.

References

Acemoglu, D., Johnson, S., and Robinson, J.A. (2001) The Colonial Origins of Comparative Development: An Empirical Investigation. *American Economic Review*, *91*(5), 1369-401.

Acemoglu, D., Gallego, F., and Robinson, J.A. (2014) Institutions, Human Capital and Development. *Annual Review of Economics, 6*, 875-912.

Agbor, J.A. (2011) How Does Colonial Origin Matter for Economic Performance in Sub-Saharan Africa? UNU WIDER Working Paper No. 2011/27. Helsinki: World Institute for Development Economics Research.

Agénor, P.-R. (2016) Caught in the Middle? The Economics of Middle-Income Traps. *Journal of Economic Surveys*, *31*(3): 771-91.

Agénor, P.-R., and Canuto, O. (2015) Middle-Income Growth Traps. *Research in Economics,* 69(4): 641–60.

Aiyar, S., Duval, R., Puy, D. Wu, Y., and Zhang, L. (2013) Growth Slowdowns and the Middle-Income Trap. IMF Working Paper 13/71. Washington, DC: International Monetary Fund.

Alesina, A., and Giuliano, P. (2015) Culture and Institutions. *Journal of Economic Literature, 53*(4): 898-944.

Amin, M., and Mattoo, A. (2006) Do Institutions Matter More for Services? World Bank Policy Research Working Paper 4023. Washington, DC: The World Bank.

Barro, R.J., and Lee, J.-W. (2013) A New Data Set of Educational Attainment in the World, 1950-2010. *Journal of Development Economics*, *104*, 184–98.

Bhattacharyya, S. (2004) Deep determinants of economic growth. *Applied Economics Letters* 11(9): 587-90.

Blanchard, O., and Kremer, M. (1997) Disorganization. *Quarterly Journal of Economics, 112*(4): 1091-126.

Dabús, C., Tohmé, F., and Carabello, M.A. (2016) A Middle Income Trap in a Small Open Economy: Modeling the Argentinian Case. *Economic Modeling*, *53*, 436-44.

Easterly, W., and Levine, R. (1997) Africa's Growth Tragedy: Policies and Ethnic Divisions. *Quarterly Journal of Economics*, *112*(4): 1203-50.

Easterly, W., and Levine, R. (2003) Tropics, germs, and crops: how endowments influence economic development. *Journal of Monetary Economics*, *50*(1): 3-39.

Easterly, W., and Levine, R. (2016) The European Origins of Economic Development. *Journal of Economic Growth*, *21*, 225-57.

Engerman, S., and Sokoloff, K. (1997) Factor Endowments, Institutions, and Differential Paths of Growth Among New World Economies. In: S. Haber (ed.), How Latin America Fell Behind: Essays on the Economic Histories of Brazil and Mexico, 1800-1914 (pp. 260-304). Stanford: Stanford University Press.

Felipe, J., Abdon, A., and Kumar, U. (2012) Tracking the Middle-Income Trap: What is it, Who is in it, and Why? Levy Economics Institute of Bard College, Working Paper No. 715. Annandale-on-Hudson, NY: Levy Economics Institute.

Finer, S. (1997) *The History of Government, Vol I-III*. Cambridge, UK: Cambridge University Press.

Frankel, J., and Romer, D. (1999) Does trade cause growth? *American Economic Review*, *89*, 379-99.

Gallup, J.L., and Sachs, J.D. (2001) The Economic Burden of Malaria. Supplement to: *The American Journal of Tropical Medicine & HygieneK, 64*(1/2): 85-96.

Gill, I., and Kharas, H. (2007) *An East Asian Renaissance – Ideas for Economic Growth*. Washington, DC: World Bank.

Glawe, L., and Wagner, H. (2016) The Middle-Income Trap: Definitions, Theories and Countries Concerned—A Literature Survey. *Comparative Economic Studies*, *58*(4): 507-38.

Glawe, L., and Wagner, H. (2017a) The People's Republic of China in the Middle-Income Trap? Asian Development Bank Institute (ADBI) Working Paper No. 749.

Glawe, L., and Wagner, H. (2017b) A Stylized Model of China's Growth Since 1978. CEAMeS Discussion Paper No. 5. Hagen: Center for East Asia Macroeconomic Studies. Also published as SSRN Working Paper No. 2917433.

Glawe, L., and Wagner, H. (2017c) The Deep Determinants of the Middle-Income Trap. CEAMeS Discussion Paper No. 10. Hagen: Center for East Asia Macroeconomic Studies.

Glawe, L., and Wagner, H. (2017d) The Deep Determinants Revisited – The Role of the Cultural Channel and Implications for the Middle-Income Trap. Mimeo.

Hall, R.E., and Jones, C.I. (1999) Why do some countries produce so much more output per worker than others? *Quarterly Journal of Economics*, *114* (1): 83-116.

McFadden, D. (1979) Quantitative Methods for Analyzing Travel Behaviour of Individuals: Some Recent Developments. In: D.A. Hensher, and P.R. Stopher (Eds.), Behavioural Travel Modelling (pp. 279-318). London: Croom Helm.

Holmstrom, B. (1985) The Provisions of Services in a Market Economy. In: R.P. Inman (Ed.), Managing the Service Economy: Prospects and Problems (pp. 138-213). Cambridge: Cambridge University Press.

Klerman, D., Mahoney, P., Spamann, H., and Weinstein, M. (2011) Legal Origin or Colonial History. Harvard Discussion Paper No. 704.

La Porta, R., Lopez-de-Silanes, F., Shleifer, A., and Vishny, R. W. (1999) The Quality of Government. *Journal of Law, Economics, and Organization*, *15*(1): 222-79.

Mayer, T., and Zignago, S. (2011) Notes on CEPII's distances measures: the GeoDist Database. *CEPII Working Paper 2011-25.*

North, D.C. (1990) *Institutions, Institutional Change, and Economic Performance*. Cambridge and London: Cambridge University Press.

North, D.C., and Thomas, R.P. (1973) *The Rise of the Western World: A New Economic History*. Cambridge, UK: Cambridge University Press.

Nunn, N. (2009) The Importance of History for Economic Development. Annual Review of Economics. *Annual Reviews*, *1*(1): 65-92.

Owen, P.D., and Weatherston, C.R. (2007) The Deep Determinants of Long-run Growth and Economic Development. In: W. R. Garside (Ed.), Institutions and Market Economies: The Political Economy of Growth and Development pp. 137-66). Basingstoke: Palgrave Macmillan.

Parker, P.M. (1997) *National cultures of the world: A statistical reference, Cross-Cultural Statistical Encyclopedia of the World*. Vol. 4. Westport, CT: Greenwood Press.

Rodrik, D., Subramanian, A., and Trebbi, F. (2004) Institutions Rule: The Primacy of Institutions Over Geography and Integration in Economic Development. *Journal of Economic Growth*, *9*, 131-65.

Spolaore, E., and Wacziarg, R. (2013) How Deep Are the Roots of Economic Development? *Journal of Economic Literature*, *51*(2): 325-69.

Wooldridge, J.M. (2002) *Econometric Analysis of Cross Section and Panel Data*. Cambridge, MA and London: MIT Press.

World Bank (2013) *China 2030: Building a Modern, Harmonious, and Creative High-Income Society*. Washington, DC: World Bank.

World Bank (2016) *Worldwide Governance Indicators (WGI).* Washington, DC: World Bank. Available online at http://info.worldbank.org/governance/wgi/index.aspx#doc, 2015 Update.

World Bank (2016) *World Development Indicators (WDI)*. Washington, DC: World Bank. Available online at http://data.worldbank.org/data-catalog/world-development-indicators.

Appendix A. Data description

Table A1. Description of the variables and data sources.

Variable Name	Description	Source
MIT	Dummy indicating whether a country is in an MIT (1) or succeeded in achieving high-income status without falling into an MIT (0).	Own calculations based on the World Bank definition.
Euroshare	Proportion of Europeans during colonization.	Easterly and Levine (2016)
Latitude	A measure of distance from the equator, in particular the absolute value of the latitude of a country scaled to take values between 0 and 1.	La Porta et al. (1999)
British legal origin	Legal dummy indicating whether a country has a British Common Law legal origin.	La Porta et al. (1999)
Ethnicity	Average of five different indices of ethnolinguistic fragmentation ranging from 1 to 0.	*Easterly and Levine (1997) as used in La Porta et al. (1999)*
Independence	Percentage of years since 1776 that a country has been independent	CIA Factbook, own calculations.
Government quality	The first principal component of the six governance indicators.	World Bank (2016).
Secondary enrollment	Average gross rate of secondary school enrollment from 1995 to 2005.	World Bank (2017), own calculations.
Secondary total	Average percentage of the population aged 15 and over with secondary education (total) between 1995 and 2005	Barro and Lee (2013), own calculations.
Secondary completed	Average percentage of the population aged 15 and over with secondary education (completed) between 1995 and 2005	Barro and Lee (2013), own calculations.
Export share	Average export share over the period between 1960 and 2000.	World Bank (2017)
Openness	Average trade share in GDP between 1960 and 2000.	World Bank (2017)
AVEXPR	Average Risk of expropriation of private foreign investment by government between 1985 and 1995, ranging from 0 to 10 where a higher score means less risk.	Acemoglu et al. (2001)
Continent dummies	Continent dummy indicating whether a country is located in Africa, Asia, Europe, or Latin America.	World Bank (2017), Mayer and Zignago (2011)
LOGEM4	Logarithm of the baseline settler mortality measured in terms of deaths per annum per 1,000 "mean strength".	Acemoglu et al. (2001)
LEGOR_FR	Legal dummy indicating whether a country has a French Commercial Code legal origin.	La Porta et al. (1999)
ENGFRAC	the fraction of a country's population speaking English as a mother tongue	Hunter (1992) and Gunnemark (1991) as used in Hall and Jones (1999).
EURFRAC	the fraction of a country's population speaking one of the five primary Western European languages (including English)	Hunter (1992) and Gunnemark (1991) as used in Hall and Jones (1999).

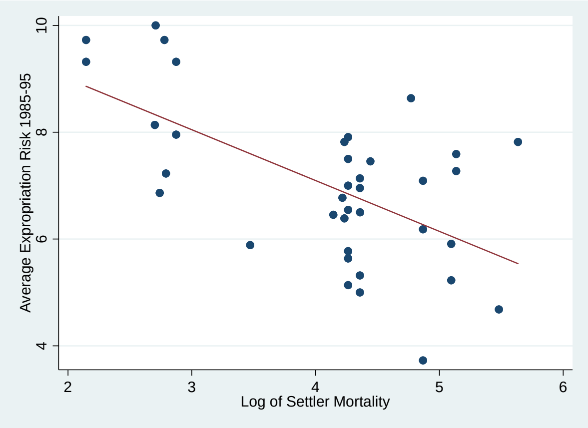
Table A1 continued.

Variable Name	Description	Source
Temperature	The first principal component of 5 temperature variables compiled by Parker (1997).	Parker (1997), own calculations.
Oil reserves	Oil resources of the country in thousands of barrels per capita.	Parker (1997)
Soil quality	Steppe, low latitude; desert, low latitude; steppe, middle latitude; desert, middle latitude; dry steppe wasteland, and desert dry winter.	Parker (1997)
Humidity	Morning minimum and maximum humidity, as well as afternoon minimum and maximum humidity (all in percentage).	Parker (1997)
Malaria dummy	Dummy equal to 1 if the percentage of the population living where falciparum malaria is endemic is greater than *zero*.	Gallup and Sachs (2001), own calculations.
Yellow fever	Dummy equal to 1 if yellow fever epidemics before 1900 and 0 otherwise.	Acemoglu et al. (2001) using data from Oldstone (1998: 69) and Curtin (1989, 1998)
Religion variables (fraction)	Percentage of the population that belonged to the three most widely spread religions in the world (in 1980 and in 1990-95 for more recently formed countries), namely Roman Catholic, Protestant, Muslim. The group "other religions" is the residual.	La Porta et al. (1999)
F_BRIT	Colonial dummy indicating whether a country was a British colony.	La Porta et al. (1999), Klerman et al. (2011), Price (2003), Treisman et al. (2014), Mayer and Zignago (2011)
European descent	Percent of population of European descent in 1975.	Acemoglu et al. (2001)
Rule of law	Rule of law index ranging from -2.5 to 2.5 for 2005 from the Worldwide Governance Indicators.	World Bank (2016)
ATRADE	Actual trade share, that is ratio of imports plus exports to GDP, 1985	Frankel and Romer (1999)
CTRADE	Constructed trade share, that is share—aggregated fitted values of bilateral trade equation with geographic variables.	Frankel and Romer (1999)

Appendix B. AJR

Figure B1. Scatter plots of other MIT country samples.

(a) Aiyar et al. (2013) base sample



(b) Felipe et al. (2012) base sample

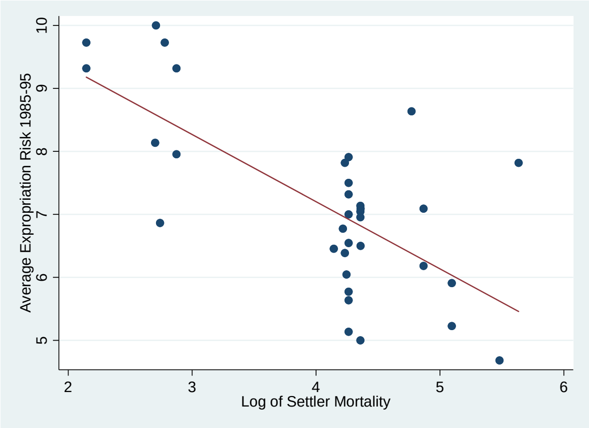


Table B1. IV regression with additional controls (I), base sample 1 (ARJ).

	Base sample 1 (1)	Base sample 1 (2)	Base sample 1 (3)	Base sample 1 (4)	Base sample 1 (5)	Base sample 1 (6)
Panel A: Two-Stage Probit						
Institutions (AVEXPR)	-1.8527*** (0.6931) -0.1577*** (0.0581)	-1.8600*** (0.6766) -0.1579** (0.0727)	-2.24511** (1.085508) -0.1761*** (0.0464)	-2.2295** (0.9209) -0.1823*** (0.0350)	-1.9197** (0.7669) -0.1634** (0.0659)	-2.0474** (0.8262) -0.1964*** (0.0658)
Latitude		-0.7476 (5.6016) -0.0635 (0.4590)		-1.8080 (3.4401) -0.1478 (0.2553)		-3.9348 (5.1594) -0.3775 (0.4547)
Temperature	0.0757 (0.2817) 0.0064 (0.0239)	0.0019 (0.4728) 0.0002 (0.4590)				
Oil reserves			-1.12e-06 (2.14e-06) -8.81e-08 (1.48e-07)	-1.46e-06 (2.24e-06) -1.20e-07 (1.40e-07)		
Yellow					-0.1972 (0.8991) -0.0168 (0.0792)	-1.1930 (1.2562) -0.1144 (0.1264)
Wald test of exogeneity	0.6385	0.6781	0.4484	0.3111	0.6576	0.4682
F-stat	10.93	7.98	12.80	9.45	11.48	8.04
Log likelihood	-74.0485	-73.0848	-72.375	-70.7780	-73.7208	-72.6710
Number of observations	44	44	44	44	44	44
Panel B: First-Stage for Average Risk Against Expropriation						
LOGEM4	-0.7415*** (0.1764)	-0.6581*** (0.1832)	-0.8202*** (0.1595)	-0.6964*** (0.1757)	-0.7256*** (0.1696)	-0.6518*** (0.1801)
Latitude		2.0818 (1.5173)		1.9972 (1.3193)		1.7471 (1.5816)

Table B1 continued.

	Base sample 1 (1)	Base sample 1 (2)	Base sample 1 (3)	Base sample 1 (4)	Base sample 1 (5)	Base sample 1 (6)
Panel C: Probit Regression						
Institutions (AVEXPR)	-1.7852**	-1.7766**	-2.0965**	-2.1471**	-1.8367**	-1.9603**
	(0.7408)	(0.7231)	(1.0281)	(1.0318)	(0.7868)	(0.9576)
	-0.1349***	-0.1318***	-0.1513***	-0.1453***	-0.1395***	-0.1401***
	(0.0348)	(0.0333)	(0.0427)	(0.0378)	(0.0372)	(0.0491)
Latitude		-1.8853		-3.1641		-4.5590
		(5.4593)		(4.0663)		(6.0549)
		-0.1398		-0.2141		-0.3259
		(0.4003)		(0.2579)		(0.4115)
Temperature	0.0883	-0.0450				
	(0.3005)	(0.5173)				
	0.0067	-0.0033				
	(0.0226)	(0.0383)				
Oil reserves			-1.53e-06	-2.27e-06		
			(2.06e-06)	(2.34e-06)		
			-1.11e-07	-1.53e-07		
			(1.40e-07)	(1.44e-07)		
Yellow					-0.0310	-0.9465
					(0.8639)	(1.5048)
					-0.0024	-0.0677
					(0.0656)	(0.1040)
McFadden's Pseudo R^2	0.7132	0.7163	0.7250	0.7408	0.7111	0.7263
Correctly classified	93.18%	93.18%	95.45%	95.45%	93.18%	95.45%
Log likelihood	-5.9840	-5.9182	-5.7378	-5.4083	-6.0267	-5.7107
Heteroskedasticity test	0.5357	0.7667	0.3012	-	0.6663	0.9087
Number of observations	44	44	44	44	44	44

Note: Dependent variable: MIT dummy constructed analogously to the World Bank (2013) study, using Maddison (2010) data. See Appendix Table Al for detailed variable definitions and sources. 'F-stat' is the F-statistic for the first-stage regression. The row 'Wald test of exogeneity' presents the p-value of the Wald test of exogeneity for the suspected endogenous variable and the row 'Heteroscedasticity test' presents the p-value for the test for heteroscedasticity. In panels A and C, the average marginal effects are reported below the respective coefficients. In Panel B, we only report the results regarding the average risk against expropriation, latitude, and other regressors if they are significant to save space. Standard errors are in parentheses. Significance at the 10, 5, and 1 percent levels are denoted by *, **, and ***, respectively.

Table B2. IV regression with additional controls (II), base sample 1 (ARJ).

	Base sample 1 (1)	Base sample 1 (2)	Base sample 1 (3)	Base sample 1 (4)	Base sample 1 (5)	Base sample 1 (6)	Base sample 1 (7)	Base sample 1 (8)
				Panel A: Two-Stage Probit				
Institutions (AVEXPR)	-1.9885**	-2.0354**	-1.668**	-1.8999***	-2.0429**	-1.9880**	-1.3049	-0.9970
	(0.8373)	(0.9523)	(0.7323)	(0.7098)	(0.8716)	(0.7941)	(1.2092)	(1.3368)
	-0.1607**	-0.1752**	-0.0966	-0.1262	-0.1707***	-0.1696**	-0.0840***	-0.0785*
	(0.0745)	(0.0876)	(0.0756)	(0.0934)	(0.0628)	(0.0754)	(0.0320)	(0.0447)
Latitude		1.3743		3.3940		-0.7493		-3.8280
		(3.6954)		(4.5060)		(3.3039)		(3.5219)
		0.1183		0.2255		-0.0639		-0.3016
		(0.3424)		(0.3755)		(0.2709)		(0.3874)
Ethnicity	1.9586	2.0622						
	(1.9122)	(2.1165)						
	0.1583	0.1775						
	(0.1373)	(0.1553)						
Religion (p-value)			[0.1031]	[0.1272]				
European descent					-0.0023	-0.0012		
					(0.0207)	(0.0196)		
					-0.0002	-0.0001		
					(0.0017)	(0.0016)		
F_BRIT							-1.5017	-1.6564*
							(0.9260)	(0.9538)
							-0.0967	-0.1305
							(0.0887)	(0.1225)
Wald test of exogeneity	0.6062	0.6122	0.7572	0.5979	0.6533	0.6458	0.5688	0.5198
F-stat	10.88	8.01	8.34	7.06	14.55	9.76	11.39	8.45
Log likelihood	-73.3037	-72.3047	-68.4725	-67.4382	-71.5997	-71.2400	-72.9346	-71.5127
Number of observations	44	44	44	44	44	44	44	44

Table B2 continued.

	Base sample 1 (1)	Base sample 1 (2)	Base sample 1 (3)	Base sample 1 (4)	Base sample 1 (5)	Base sample 1 (6)	Base sample 1 (7)	Base sample 1 (8)
Panel B: First-Stage for Average Risk Against Expropriation								
LOGEM4	-0.7564*** (0.1697)	-0.6469*** (0.1829)	-0.6710*** (0.1532)	-0.5616*** (0.1715)	-0.5821*** (0.1760)	-0.5414*** (0.1829)	-0.7138*** (0.1750)	-0.5930*** (.18922)
Latitude		1.9708 (1.3845)		1.7812 (1.3461)		1.0636 (1.4049)		2.001 (1.3496)
European descent					0.0148** (0.0063)	0.0130* (0.0068)		

Table B2 continued.

	Base sample 1 (1)	Base sample 1 (2)	Base sample 1 (3)	Base sample 1 (4)	Base sample 1 (5)	Base sample 1 (6)	Base sample 1 (7)	Base sample 1 (8)
Panel C: Probit Regression								
Institutions (AVEXPR)	-1.9800**	-2.0399**	-1.5312**	-1.6668**	-1.9162**	-1.8520**	-1.7735**	-1.6941**
	(0.8709)	(0.9769)	(0.6685)	(0.7576)	(0.9116)	(0.8522)	(0.8172)	(0.7318)
	-0.1281**	-0.1316***	-0.0830**	-0.0888**	-0.1427***	-0.1358***	-0.1113***	-0.0999***
	(0.0394)	(0.0467)	(0.0340)	(0.0355)	(0.0433)	(0.0410)	(0.0339)	(0.0247)
Latitude		0.7221		1.9394		-1.2928		-3.1955
		(4.2288)		(4.3533)		(3.3997)		(4.1928)
		0.0466		0.1034		-0.0005		-0.1885
		(0.2723)		(0.2334)		(0.0013)		(0.2407)
Ethnicity	2.1642	2.3156						
	(1.9294)	(2.1210)						
	0.1401	0.1494						
	(0.1122)	(0.1229)						
Religion (p-value)			[0.5926]	[0.6299]				
European descent					-0.0076	-0.0063		
					(0.0178)	(0.0177)		
					-0.0006	-0.0948		
					(0.0013)	(0.2461)		
F_BRIT							-1.0991	-1.385
							(0.9192)	(1.1243)
							-0.0690	-0.0817
							(0.0490)	(0.0603)
McFadden's Pseudo R^2	0.7496	0.7503	0.7822	0.7870	0.7159	0.7194	0.7493	0.7649
Correctly classified	95.45%	95.45%	95.45%	95.45%	95.45%	93.18%	95.45%	95.45%
Log likelihood	-5.2244	-5.2095	-4.5444	-4.4442	-5.9270	-5.8549	-5.2301	-4.9046
Heteroskedasticity test	0.9047	0.7176	-	0.9734	0.6883	0.7176	0.3059	0.5805
Number of observations	44	44	44	44	44	44	44	44

Note: See Notes of Table B1.

Table B3. IV regression with additional controls (I), base sample 2 (ARJ).

	Base sample 2 (1)	Base sample 2 (2)	Base sample 2 (3)	Base sample 2 (4)	Base sample 2 (5)	Base sample 2 (6)	Base sample 2 (7)	Base sample 2 (8)
				Panel A: Two-Stage Probit				
Institutions (AVEXPR)	-1.5788*** (0.4251) -0.1804*** (0.0210)	-1.6390*** (0.4708) -0.1754*** (0.0371)	-1.5723*** (0.4192) -0.1808*** (0.0317)	-1.5405*** (0.4204) -0.1909*** (0.0322)	-1.2849*** (0.2791) -0.1807*** (0.0193)	-1.2602*** (0.2841) -0.1906*** (0.0211)	-1.2640*** (0.3307) -0.1340*** (0.0298)	-1.4485*** (0.4297) -0.1516*** (0.0507)
Latitude		-3.0273 (3.6701) -0.3240 (0.3432)		0.9898 (2.3193) 0.1226 (0.3144)		1.2856 (2.0034) 0.1945 (0.3178)		3.3179 (2.1634) 0.3473 (0.2149)
Temperature	-0.1810 (0.1832) -0.0207 (0.0226)	-0.4195 (0.3357) -0.0449 (0.0345)						
Oil reserves			-7.52e-07 (6.80e-07) -8.65e-08 (7.47e-08)	-8.18e-07 (5.67e-07) -1.01e-07 (7.56e-08)				
Yellow					-0.0657 (0.5471) -0.0092 (0.0778)	0.1541 (0.6548) 0.0233 (0.0982)		
Malaria							9.8243 (12.6294) 1.0417 (1.1541)	14.3031 (25.6367) 1.4972 (2.1251)
Wald test of exogeneity	0.3132	0.5128	0.3119	0.3543	0.1192	0.1147	0.1086	0.0935
F-stat	14.10	8.93	14.48	13.25	10.26	7.18	16.49	10.93
Log likelihood	-112.1201	-109.4710	-111.8006	-105.5063	-118.3685	-115.3595	-107.1288	-103.7272
Number of observations	60	59	60	59	61	60	59	58

Table B3 continued.

	Base sample 2 (1)	Base sample 2 (2)	Base sample 2 (3)	Base sample 2 (4)	Base sample 2 (5)	Base sample 2 (6)	Base sample 2 (7)	Base sample 2 (8)
Panel B: First-Stage for Average Risk Against Expropriation								
LEGOR_FR	-1.4030*** (0.3425)	-1.2974*** (0.3656)	-1.3325*** (0.3442)	-0.9146*** (.3440)	-1.4562*** (0.3521)	-1.1834*** (0.3761)	-1.5855*** (0.3316)	-1.3645*** (0.3535)
Latitude		0.4962 (1.4707)		3.0816*** (1.0059)		2.4280* (1.4781)		1.6528 (1.1635)
Temperature	-0.3045*** (0.0991)	-0.2798** (0.1352)						
Oil reserves			-6.36e-07*** (2.01e-07)	-7.11e-07*** (1.88e-07)				
Malaria							-1.7268*** (0.4893)	-1.3527** (0.5397)

Table B3 continued.

	Base sample 2 (1)	Base sample 2 (2)	Base sample 2 (3)	Base sample 2 (4)	Base sample 2 (5)	Base sample 2 (6)	Base sample 2 (7)	Base sample 2 (8)
Panel C: Probit Regression								
Institutions (AVEXPR)	-1.5604***	-1.5965***	-1.5546***	-1.4583**	-1.2396***	-1.2011***	-1.2654***	-1.538***
	(0.4759)	(0.5153)	(0.4620)	(0.4679)	(0.3255)	(0.3513)	(0.3765)	(0.5058)
	-.1581***	-.1520***	-0.1565***	-0.1462***	-0.1536***	-0.1502***	-0.1218***	-0.1387***
	(0.0215)	(0.0237)	(0.0192)	(0.0261)	(0.0156)	(0.0237)	(0.0133)	(0.0194)
Latitude		-3.9698		-1.0449		-0.2168		3.5132
		(3.4254)		(1.7685)		(2.4656)		(3.0129)
		-0.3780		-0.1048		-0.0271		0.3170
		(0.3065)		(0.1763)		(0.3085)		(0.2569)
Temperature	-0.0769	-0.4058						
	(0.1960)	(0.3631)						
	-0.0078	-0.0386						
	(0.0197)	(0.0323)						
Oil reserves			-5.98e-07	-5.96e-07				
			(7.96e-07)	(6.85e-07)				
			-6.02e-08	-5.97e-08				
			(7.83e-08)	(6.67e-08)				
Yellow					0.4622	0.3814		
					(0.5621)	(0.8568)		
					0.0573	0.0477		
					(0.0687)	(0.1066)		
Malaria							14.2406	29.9839
							(19.4791)	(40.6137)
							1.3704	2.7051
							(1.8693)	(3.6221)
McFadden's Pseudo R^2	0.6424	0.6532	0.6441	0.6352	0.5796	0.5631	0.6780	0.6907
Correctly classified	91.67%	89.83%	91.67%	93.22%	90.16%	90.00%	93.22%	94.83%
Log likelihood	-11.2143	-10.3346	-11.1595	-10.8704	-13.8129	-13.7014	-10.4098	-9.5448
Heteroskedasticity test	0.9126	0.9484	0.9844	0.9954	0.8842	0.9620	0.9449	0.9905
Number of observations	60	59	60	59	61	60	59	58

Note: See Notes of Table B1.

Table B4. IV regression with additional controls (II), base sample 2 (ARJ).

	Base sample 2 (1)	Base sample 2 (2)	Base sample 2 (3)	Base sample 2 (4)	Base sample 2 (5)	Base sample 2 (6)	Base sample 2 (7)	Base sample 2 (8)
Panel A: Two-Stage Probit								
Institutions (AVEXPR)	-1.4602*** (0.3528) -0.1705*** (0.0236)	-1.4697*** (0.3545) -0.1839*** (0.0221)	-1.3444*** (0.2997) -0.1725*** (0.0544)	-1.3306*** (0.3039) -0.1852*** (0.0629)	-1.3951*** (0.2966) -0.1965*** (0.0227)	-1.3824*** (0.3002) -0.1939*** (0.0352)	-1.3407*** (0.3490) -0.1513*** (0.0240)	-1.1309** (0.5047) -0.1148*** (0.0370)
Latitude		1.6210 (1.7810) 0.2028 (0.2519)		1.0659 (2.1724) 0.1483 (0.3480)		-0.9442 (1.9079) -0.1324 (0.2529)		-2.5981 (2.9457) -0.2638 (0.2969)
Ethnicity	0.2046 (1.3189) 0.0239 (0.1513)	0.3454 (1.2413) 0.04321 (0.1525)						
Religion (p-value)			[0.2020]	[0.1322]				
European descent					0.0123* (0.0064) 0.0017 (0.0011)	0.0142* (0.0074) 0.0020 (0.0013)		
F_BRIT							-0.9482 (0.6349) -0.1070* (0.0613)	-1.3740* (0.7330) -0.1395** (0.0659)
Wald test of exogeneity	0.1636	0.1579	0.3972	0.4069	0.1243	0.2200	0.6151	0.9684
F-stat	11.89	10.12	6.29	6.40	21.88	14.73	10.64	9.98
Log likelihood	-101.2966	-98.7887	-115.81751	-110.67757	-110.5498	-107.5906	-116.0741	-111.3340
Number of observations	57	57	61	60	61	60	60	59

Table B4 continued.

	Base sample 2 (1)	Base sample 2 (2)	Base sample 2 (3)	Base sample 2 (4)	Base sample 2 (5)	Base sample 2 (6)	Base sample 2 (7)	Base sample 2 (8)
				Panel B: First-Stage for Average Risk Against Expropriation				
LEGOR_FR	-1.3691*** (0.3220)	-1.0699*** (0.3348)	-1.1122** (0.4749)	-0.8710* (0.4579)	-1.2632*** (0.3125)	-1.1723*** (0.3236)	-2.2067*** (0.4950)	-1.7587*** (0.5284)
Latitude		2.4105** (1.0536)		2.8499*** (1.0714)		0.0251 (1.1466)		2.5101** (1.1130)
Ethnicity	-1.6798*** (0.5866)	-1.0961* (0.6166)						
European descent					0.0198*** (0.0042)	0.0206*** (0.0049)		
F_BRIT							-0.9898** (0.5026)	-0.7548 (0.5124)

Table B4 continued.

	Base sample 2 (1)	Base sample 2 (2)	Base sample 2 (3)	Base sample 2 (4)	Base sample 2 (5)	Base sample 2 (6)	Base sample 2 (7)	Base sample 2 (8)
Panel C: Probit Regression								
Avexpr	-1.4234***	-1.3901***	-1.2146***	-1.1579***	-1.3467***	-1.2754***	-1.3017***	-1.1430***
	(0.4124)	(0.4249)	(0.3613)	(0.3753)	(0.3565)	(0.3610)	(0.3788)	(0.3933)
	-0.1423	-0.1392***	-0.1365***	-0.1312***	-0.1668***	-0.1534***	-0.1431***	-0.1160***
	(0.0188)	(0.0217)	(0.0228)	(0.0272)	(0.0170)	(0.0235)	(0.0171)	(0.0227)
Latitude		-0.5512		-0.6259		-1.9183		-2.5197
		(1.8481)		(1.7607)		(1.9247)		(2.2044)
		-0.0552		-0.0709		-0.2308		-0.2557
		(0.1857)		(0.1994)		(0.2253)		(0.2148)
Ethnicity	0.9080	0.7577						
	(1.3927)	(1.4882)						
	0.0908	0.0759						
	(0.1363)	(0.1467)						
Religion (p-value)			[0.4652]	[0.5231]				
European descent					0.0050	0.0079		
					(0.0069)	(0.0079)		
					0.0006	0.0010		
					(0.0008)	(0.0009)		
F_BRIT							-1.0576*	-1.3663*
							(0.5968)	(0.7105)
							-0.1162**	-0.1387**
							(0.0581)	(0.0619)
McFadden's Pseudo R^2	0.6385	0.6400	0.6150	0.6012	0.5775	0.5768	0.6223	0.6387
Correctly classified	91.23%	91.23%	91.80%	91.67%	90.16%	91.67%	86.67%	88.14%
Log likelihood	-10.6048	-10.5602	-12.6504	-12.5073	-13.8843	-13.2718	-12.3125	-11.2413
Heteroskedasticity test	0.7713	0.7677	0.3791	-	0.5835	0.7898	0.9236	0.9827
Number of observations	57	57	61	60	61	60	60	59

Note: See Notes of Table B1.

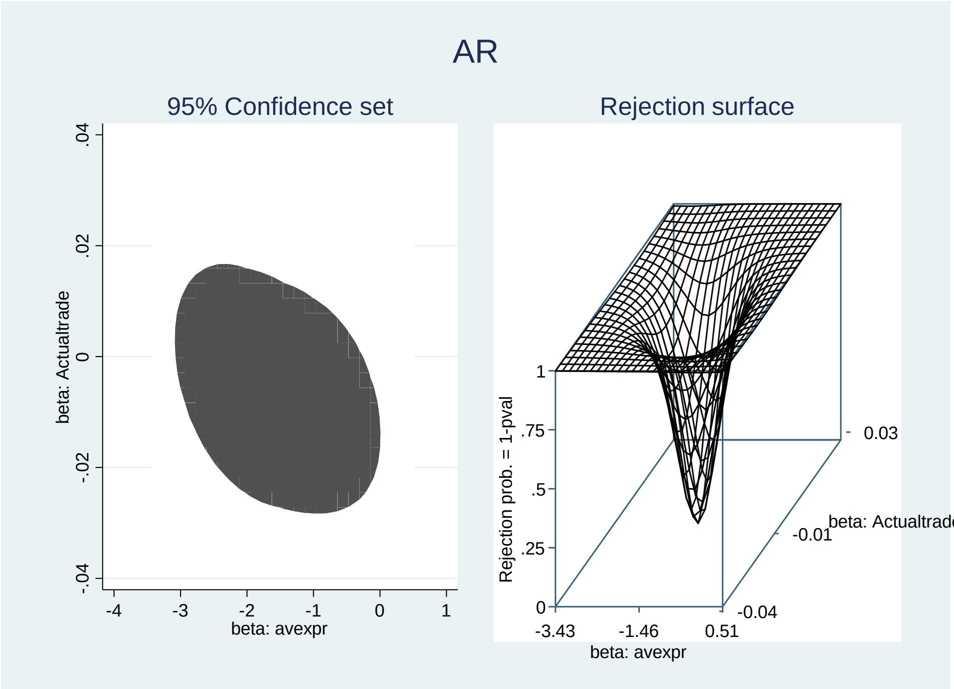
Table B5. Probit regression with institutions and human capital (ARJ).

	Base sample (1)	Base sample (2)	Base sample (3)	Base Sample (4)	Base Sample (5)	Base Sample (6)	Base Sample (7)	Base Sample (8)	Base Sample (9)	Base Sample (10)	Base Sample (11)
Human capital	-0.7121***		-0.7509***		-0.6095**		-0.5664**		-0.4525	-0.5469	-0.4390
	(0.2601)		(0.2985)		(0.2447)		(0.0263)		(0.3488)	(0.3929)	(0.3396)
	-0.0718***		-0.0750***		-0.0548***		-0.0507**		-0.0341	-0.0395	-0.0330
	(0.0165)		(0.0197)		(0.0185)		(0.0209)		(0.0239)	(0.0245)	(0.0229)
Institutions (AVEXPR)		-1.5681***		-1.6857***		-1.5367***		-1.6731**	-1.0287*	-1.1790*	-0.9795*
		(0.4495)		(0.5583)		(0.4724)		(0.6813)	(0.5485)	(0.6249)	(0.5884)
		-0.1486***		-0.1577***		-0.1450***		-0.1570***	-0.0776**	-0.0852**	-0.0735**
		(0.0250)		(0.0354)		(0.0270)		(0.0510)	(0.0335)	(0.0337)	(0.0374)
Latitude			0.8055	1.0971			-1.0284	1.0886		2.6206	
			(2.8729)	(2.7181)			(3.3520)	(3.5857)		(3.4602)	
			0.0805	0.1026			-0.0920	0.1021		0.1893	
			(0.2852)	(0.0253)			(0.3000)	(0.3360)		(0.2415)	
Continent dummies	No	No	No	No	Yes	Yes	Yes	Yes	No	No	Yes
Log likelihood	-7.9472	-7.0800	-7.9078	-6.9968	-7.2898	-7.0408	-7.2434	-6.9937	-5.7727	-5.4670	-5.7240
McFadden's Pseudo R^2	0.5800	0.6259	0.5821	0.6303	0.6148	0.6279	0.6172	0.6304	0.6949	0.7111	0.6975
Correctly classified	90.48%	95.24%	90.45%	95.24	92.86%	95.24%	92.86%	95.24%	95.24%	95.24%	95.24%
Heteroscedasticity test	0.9963	0.1181	(b)	0.1767	0.7430	(b)	(b)	0.1157	0.1895	(b)	(b)
Observations	42	42	42	42	42	42	42	42	42	42	42

Note: Dependent variable: MIT dummy constructed analogously to the World Bank (2013) study, using Maddison (2010) data. See Appendix Table Al for detailed variable definitions and sources. The row 'Heteroscedasticity test' presents the p-value for the test for heteroscedasticity. We report both, the coefficients and, below, the average marginal effects. The respective standard errors are in parentheses. Significance at the 10, 5, and 1 percent levels are denoted by *, **, and ***, respectively. (b) Convergence not achieved.

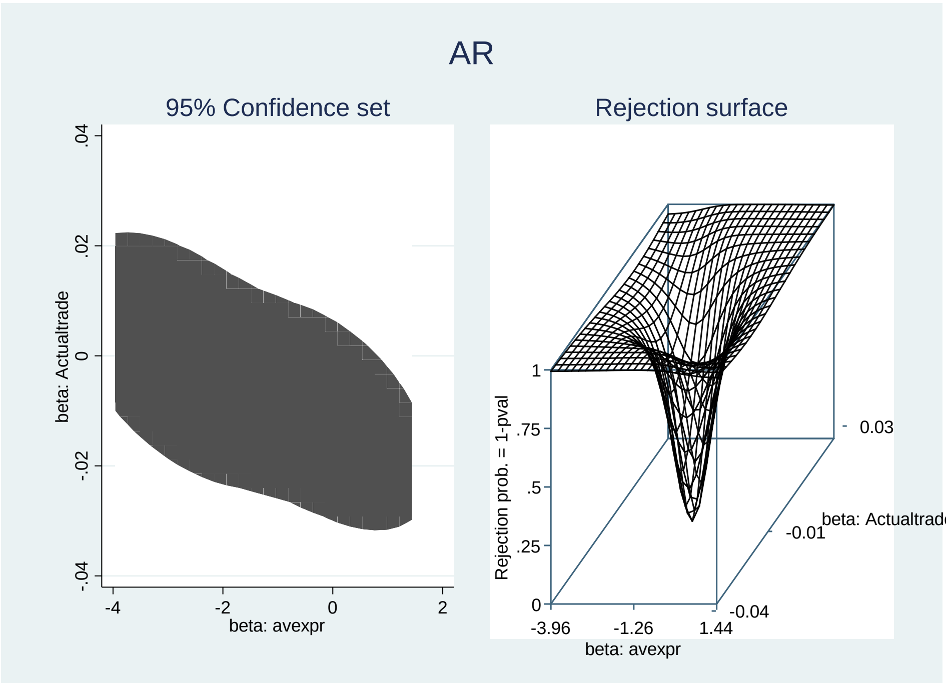
Appendix C. RST – AR test results

Figure C1. AR test results of Table 11, Column (2) (RST).



Notes: 'avexpr' denotes the average risk against expropriation and 'Actualtrade' denotes the actual trade share.

Figure C2. AR test results of Table 11, Column (3) (RST).



Notes: 'avexpr' denotes the average risk against expropriation and 'Actualtrade' denotes the actual trade share.