Essays on development economics

Wang, Wen

2019

Wang, W. (2019). Essays on development economics. Doctoral thesis, Nanyang Technological University, Singapore.

<https://hdl.handle.net/10356/137042>

<https://doi.org/10.32657/10356/137042>

This work is licensed under a Creative Commons Attribution-NonCommercial 4.0 International License (CC BY-NC 4.0).

Downloaded on 26 Apr 2023 22:37:33 SGT

**NANYANG
TECHNOLOGICAL
UNIVERSITY**

ESSAYS ON DEVELOPMENT ECONOMICS

WANG WEN

SCHOOL OF SOCIAL SCIENCES

2019

ESSAYS ON DEVELOPMENT ECONOMICS

WANG WEN

SCHOOL OF SOCIAL SCIENCES

**A thesis submitted to the Nanyang Technological University
in partial fulfilment of the requirement for the degree of
Doctor of Philosophy**

2019

Statement of Originality

I certify that all work submitted for this thesis is my original work. I declare that no other person's work has been used without due acknowledgment. Except where it is clearly stated that I have used some of this material elsewhere, this work has not been presented by me for assessment in any other institution or University. I certify that the data collected for this project are authentic and the investigations were conducted in accordance with the ethics policies and integrity standards of Nanyang Technological University and that the research data are presented honestly and without prejudice.

January 8, 2020

Date

Wang Wen

Supervisor Declaration Statement

I have reviewed the content of this thesis and to the best of my knowledge, it does not contain plagiarised materials. The presentation style is also consistent with what is expected of the degree awarded. To the best of my knowledge, the research and writing are those of the candidate except as acknowledged in the Author Attribution Statement. I confirm that the investigations were conducted in accordance with the ethics policies and integrity standards of Nanyang Technological University and that the research data are presented honestly and without prejudice.

24 Jul 2019

.....

Date

.....

A / P James Ang

Authorship Attribution Statement

This thesis contains material from 2 papers accepted at conferences in which I am listed as an author.

Part of the content of Chapter 1 was presented as Wen Wang, “Culture and Democracy” at the *7th Kobe-NTU-Hanyang Joint Symposium in Economics*, Seoul, Korea in June 2019, the *13th Annual Meeting of the Portuguese Economic Journal*, Évora, Portugal in July 2019, and the *8th Singapore Economic Review Conference*, Singapore in August 2019.

Part of the content of Chapter 2 was presented as Wen Wang, “Novelty-seeking Traits and Knowledge Intensity” at the *94th Annual Conference of Western Economic Association International*, San Francisco, U.S.A. in July 2019.

January 8, 2020

Date

Wang Wen

Acknowledgments

I wish to express my sincere gratitude to my thesis supervisor, Professor James B. Ang, for the continuous guidance and support. This PhD would not have been achievable without his encouragement. He is an excellent role model for me during this learning journey, both as a scholar with immense knowledge and tremendous passion for research, and as a teacher who never fails to nurture and inspire students.

I would like to thank the members of my Thesis Advisory Committee, Professor Ng Yew Kwang for his guidance and valuable comments, and Professor Laura Wu, for the valuable feedback and encouragement. I would also like to thank Professor Jakob Madsen for valuable comments and feedback on the thesis chapters, Professor Tan Teck Yong for guidance on modeling, Professor Joseph D. Alba, Professor Jan F. Kiviet, Professor Huang Weihong, Professor Feng Qu, Professor Low Chan Kee for teaching us and Professor Chia Wai Mung for engaging us in academic exchange and learning opportunities. I am also grateful to friends, current and former colleagues, and classmates from both the masters and PhD programs, especially Lan and Mu, for their encouragement.

Finally, I wish to thank my parents for their unconditional love and understanding.

Contents

Statement of Originality	iii
Supervisor Declaration Statement	iv
Authorship Attribution Statement	vii
Acknowledgements	ix
Lists of Tables	xv
Lists of Figures	xvii
Executive Summary	xix
1 Culture and Democracy	1
1.1 Introduction	1
1.2 Background: wetland rice cultivation and collectivism	5
1.3 Empirical approach	7
1.3.1 Estimation methodology	7
1.3.2 Data	8
1.4 Empirical results	12
1.4.1 Main regression results	12
1.5 Robustness checks	15
1.5.1 Robustness to unobservables	15
1.5.2 Robustness to early development, alternative channels and historical institutions	15
1.5.3 Robustness to effects of some contemporary measures	19
1.5.4 Alternative measures of democracy	22
1.5.5 Region exclusion analysis	24

1.5.6	Considering alternative cereal crops	25
1.5.7	Further robustness checks	25
1.6	Conclusion	28
1.A	Additional analysis: Does rice suitability transmit to democracy through culture?	30
1.B	Variable definitions and data sources	32
1.B.1	Main dependent variable	32
1.B.2	Main explanatory variable	32
1.B.3	Control variables	32
1.B.4	Cultural dimensions	36
2	Novelty-seeking Traits and Knowledge Intensity	38
2.1	Introduction	38
2.2	Conceptual framework	41
2.3	Theoretical framework	42
2.4	Empirical approach	44
2.4.1	Estimation methodology	44
2.4.2	Data	44
2.5	Empirical results	48
2.5.1	Main regression results	48
2.6	Robustness checks	52
2.6.1	Robustness to historical effects	52
2.6.2	Robustness to effects of contemporary measures	54
2.6.3	Controlling for some population genetic measures	56
2.6.4	Controlling for some cultural effects	58
2.6.5	Alternative measures of knowledge and technology	59
2.6.6	Robustness checks using the historical average of ECI data	62
2.6.7	Repeated cross-country analysis	62
2.7	Conclusion	64
2.A	Additional analysis: region exclusion test	65
2.B	Variable definitions and data sources	65
2.B.1	Main dependent variable	65
2.B.2	Main explanatory variable	67
2.B.3	Control variables	67

3	Religiosity and technology adoption	72
3.1	Introduction	72
3.2	Empirical approach	76
3.2.1	Estimation methodology	76
3.2.2	Identification strategies	77
3.2.3	Data	80
3.3	Empirical results	83
3.3.1	Main regression results	83
3.3.2	Results of instrumental variable estimations	87
3.3.3	Results of joint estimation with Gaussian copula correction	88
3.3.4	Religiosity measures over different periods of time	90
3.4	Robustness to unobservables, culture, potential barriers to diffusion and some other contemporary measures	91
3.4.1	Sensitivity to unobservables	91
3.4.2	Religious denominations	91
3.4.3	Controlling for other cultural effects	92
3.4.4	Considering social diversity	92
3.4.5	Historical level of technology adoption	93
3.4.6	Robustness to some contemporary measures	93
3.4.7	Legal system and intellectual property rights protection	93
3.5	Further analyses	94
3.5.1	Religiosity and attitude towards science and technology	94
3.5.2	Religiosity and other measures of technology	96
3.6	Conclusion	98
3.A	Religion and its intersection with science/technology	99
3.B	Additional regression results	100
3.C	Variable definitions and data sources	101
3.C.1	Main dependent variable	101
3.C.2	Main explanatory variable	101
3.C.3	Instrumental variable	101
3.C.4	Control variables	102
3.C.5	Attitude and preference towards science and technology	103
3.C.6	Other measures of technology	104

Bibliography	113
---------------------	------------

List of Tables

1.1	Correlation matrix	6
1.2	Summary statistics	12
1.3	Main regression results	14
1.4	Controlling for effects of some historical measures	16
1.5	Controlling for some contemporary measures	20
1.6	Alternative measures of democracy	22
1.7	Region exclusion analysis	24
1.8	Analysis considering alternative cereals	26
1.9	Additional robustness checks	28
1.10	Regressing democracy on individualism using rice suitability as an instrument	31
2.1	Summary Statistics	47
2.2	Main regression Results	50
2.3	Controlling for historical effects	53
2.4	Controlling for effects of contemporary measures	55
2.5	Controlling for some genetic measures	57
2.6	Controlling for some cultural influences	59
2.7	Alternative measures of knowledge and technology	60
2.8	Robustness check using historical average of ECI	63
2.9	Repeated cross-country results	64
2.10	Sensitivity test to the exclusion of regions	65
3.1	Summary statistics	82
3.2	Main regression results (OLS estimates)	84
3.3	Sector-based regression results (OLS estimates)	86
3.4	Regression results using 2SLS estimations	89
3.5	Results of joint estimation using Gaussian copula correction	90

3.6	Regression results using religiosity measures from different periods of time	90
3.7	Robustness checks (OLS estimates)	95
3.8	Attitude and preference towards science and technology (individual-level analysis)	96
3.9	Other measures of technology	98
3.B1	Regression results using 2SLS estimations- Sectoral indices	100

List of Figures

1.1	Spatial distribution of institutionalized democracy scores (1960 - 2013) .	9
1.2	Spatial distribution of wetland rice suitability ratio	11
2.1	Spatial distribution of ECI scores for the year 2015	46
2.2	Spatial distribution of DRD4 exon III 2R and 7R allele frequency	48
2.3	The inverted U-shaped relationship between DRD4 allele frequency and ECI (unconditional)	49
2.4	The estimated inverted U-shaped association between DRD4 measure and ECI (augmented component plus residual plot)	52
3.1	Spatial distribution maps	82
3.2	Relationship between technology adoption and religiosity variation . . .	85
3.3	Relationship between sector technology adoption and religiosity variation	86

Executive Summary

This thesis comprises three self-contained essays that empirically study several topics on development economics. It seeks to uncover some potential deep-rooted factors that continue to influence contemporary economic and institutional outcomes, and to further our understanding of the differentials in economic development across countries.

Chapter 1 tests the hypothesis that societies with a history of rice farming culture tend to be less democratic today. We argue that this agricultural legacy helps explain variations in adoption of democratic institutions across countries via fostering the formation and transmission of a more collectivist culture, which in turn generates greater conformity pressures on political norms and detestation of institutional changes in society that hinder democratization. Using data on suitability of land for wetland rice cultivation as the proxy for rice farming culture, our analysis shows that there is a negative correlation between institutionalized democracy and rice farming culture at the country level. Our results indicate that an increase of one standard deviation in the rice suitability ratio leads to a decrease in the degree of institutionalized democracy by 0.138 standard deviations, holding other things equal. The estimation results remain robust when subjected to further checks and sensitivity tests. In addition, we consider alternative channels including pathogen stress, historical population density and colonization, which may also influence the degree of institutionalized democracy via the cultural dimension of collectivism / individualism. Our findings suggest that land suitability for rice cultivation remains as the major deep-rooted factor in explaining the implementation of differential political institutions across contemporary societies.

In Chapter 2, we focus on the deep-rooted determinant of knowledge intensity. Technological advancement is the key driver of economic growth. The literature suggests that some personality traits, in particular, novelty-seeking is closely associated with creativity and innovation. While novelty-seeking individuals generally are more exploratory and risk-taking, they are also susceptible to psychological disorders such as

Attention Deficit Hyperactivity Disorder (ADHD). Using country-level data on DRD4 exon III allele frequency as a proxy for the level of novelty-seeking traits in a country and the Economic Complexity Index as the measure of knowledge intensity, we examine the impact of such traits on national knowledge intensity. Our results indicate a robust inverted U-shaped relationship between the level of novelty-seeking traits and national knowledge intensity. Our findings therefore suggest that novelty-seeking traits have both positive and negative effects on knowledge intensity. Countries with moderate levels of these traits are able to exploit the advantages associated with the traits to promote technological deepening whereas both low and high levels of these traits are suboptimal for technological progress.

In Chapter 3, we document a robustly negative correlation between religiosity and the level of technology adoption in society. Heightened levels of religiosity in individuals tend to associate with more adverse attitude towards science and technology, and consequently are associated with lower levels of technology adoption in society. Using data from the World Values Survey (1981 - 2014), we construct an aggregate measure of religiosity at the country level and estimate that an increase of one standard deviation in religiosity is correlated with a decrease in the level of technology adoption by 56.4% standard deviations, holding other things constant. Additionally, we employ a measure on the historical level of pathogen prevalence as an instrumental variable for religiosity to establish the direction of causation from religiosity to technology adoption. As an alternative approach, we use Gaussian copula correction for the joint estimation of religiosity and error terms to address the issue of potential endogeneity of religiosity measure. Further analyses suggest that religiosity has similar effects on other aspects of technological progress. Our study thus contributes to the ongoing discussions on the enduring influence of religion on economic development in today's world.

Chapter 1

Culture and Democracy

1.1 Introduction

Political institutions are often deemed to play a significant role in determining economic growth and development. The question of what shape contemporary political institutions, in particular, determinants of democratization, has attracted considerable attention from both scholars and the general public. In recent years, there is a gradual shift of focus in research, from examining the proximate factors of democratization as represented by the modernization theory (see, e.g., Lipset, 1959; Boix and Stokes, 2003), to searching for more deep-rooted factors that influence the formation or nature of existing political institutions (see, e.g., Acemoglu et al., 2001, 2005; Woodberry, 2012; Giuliano and Nunn, 2013; Bentzen et al., 2017; Galor and Klemp, 2017), among these include agricultural legacies¹.

The idea that agricultural legacies matter for economic and political development has been well discussed in the literature. In his seminal work, *Guns, Germs, and Steel*, Jerad Diamond (1997) highlights the pivotal role of several major agricultural legacies in explaining the differences between societal development of countries observed today. For instance, ancient people who lived in regions with more favorable biogeographic endowment for agriculture would transit from a hunter-gatherer to agrarian society faster than those in other regions. The transition to agrarian societies brought about food security and social stability, and in turn led to a large increase in population growth and specialization of labor that further fueled technological progress and economic development in these regions (see also, Hibbs and Olsson, 2004; Olsson and Hibbs Jr,

¹Other potential deep roots include village democracy, irrigation potential and colonization.

2005). Not surprisingly, the formation of states generally followed shortly after the transition to agrarian societies in most ancient civilizations (Chanda and Putterman, 2007). The developmental head-start established during the transition continues to have an enduring effect on current social and economic outcomes (see, e.g., Nunn, 2009).

Indeed, agriculture was an indispensable part of economic life before the Industrial Revolution and the glue that held individuals together in community. As Karl Marx succinctly put, that it was through interaction and cooperation in agricultural practices that “each individual has no more torn himself off from the navel-string of his tribe or community, than each bee has freed itself from connexion with the hive” (Marx, 1977). Even when most countries have already moved from agrarian to industrialized and urbanized societies, history such as agricultural legacies, continues to shape the modern world through various channels, one of which is culture². Cultivation of different crops, in which the choice is essentially determined by the environmental suitability of the region for cultivation, requires different levels of inputs, farming techniques and coordinated efforts, consequently, generates different agricultural cultures specific to the region.

An important dimension in culture is the collectivist / individualistic cleavage. Their cultural values differ primarily in the emphasis on inter-dependence (as in collectivism) vs. independence (as in individualism) with one’s community (Markus and Kitayama, 1991, 1994). Recent studies have suggested that culture is one of the key determinants of democratization. For instance, Gorodnichenko and Roland (2015) find that an individualistic culture tends to favor democratization, while countries with a collectivist culture are less likely to adopt democracy. Nevertheless, what is the underlying historical force that drives the formation and persistence of such collectivist / individualistic culture, and in turn shapes modern political institutions, remains to be further explored.

Building on these insights, this paper tests the hypothesis that countries with a rice farming legacy are less likely to experience democratization and the main argument runs as follows: being relatively labor-intensive, rice cultivation requires concerted efforts from individuals and thus tends to enmesh people in the close-knit and enduring networks of cooperation, coordination and inter-dependence. This favors the emergence and transmission of a more collectivist culture in society. As a result, it generates strong

²As suggested by Guiso et al. (2006), culture broadly comprises “those customary beliefs and values that ethnic, religious, and social groups” possess with fairly high degrees of commonality within each group.

conformity pressures on political norms and detestation of any major institutional evolution and changes, which hinder democratization (Kateb, 1992; Brewer and Chen, 2007; Triandis, 2018).

Using data on suitability of land for wetland rice cultivation from the FAO and IIASA (2000) as the proxy for rice farming culture, and the measure of institutionalized democracy for contemporary societies from the Polity IV project, we establish a negative correlation between them. Our results indicate that an increase of one standard deviation in the rice suitability ratio leads to a decrease in the degree of institutionalized democracy by 0.138 standard deviations, holding other things equal. We further perform a battery of robustness tests to control for potential confounding effects of some early development indicators, historical institutions and contemporary measures on democracy. The results also survive when we use alternative measures of democracy or perform restricted sample estimations. Moreover, we investigate other potential channels apart from rice farming culture, which include pathogen stress, historical population density and colonization; these have been proposed in the literature as alternative theories on the historical origin of collectivist / individualistic culture, thus may also influence the degree of institutionalized democracy through this cultural dimension (Vandello and Cohen, 1999; Lange et al., 2006; Fincher et al., 2008). Our findings suggest that land suitability for rice cultivation remains as the major deep-rooted factor that explains the implementation of differential political institutions across contemporary societies. Additionally, when we consider effects of cultivation of other major cereal crops, i.e., wheat, barley and maize, or major labor-intensive crops such as pea and potato in the analysis, land suitability ratio of rice cultivation consistently shows up as a significant covariate.

Our study thus provides novel evidence on how agricultural legacies continue to influence institutional development in contemporary society. It advances the work of Gorodnichenko and Roland (2015) by pinning down the historical origin of collectivist / individualistic culture that impacts on democratization, thus furthering the understanding of deep-rooted determinants of democracy. Moreover, a potential issue with the key indicator of collectivist / individualistic culture used by Gorodnichenko and Roland (2015), i.e., the Hofstede's cultural dimensions, is that those were constructed based on survey data *ex post* and thus may be subjected to measurement errors, which would bias the estimates. Another issue lies with the identification: since democracy may influence one's cultural beliefs and causality could run in the reverse direction,

instrumental variable estimation is resorted to. In our study, we measure rice farming culture using data on land suitability for wetland rice cultivation from the FAO. These data sets are constructed using information of environmental factors, such as climatic and ecological factors that determine the overall suitability for cultivation, and therefore are sufficiently exogenous to the formation and transmission of collectivist culture and values in question. Specifically, we choose data of land suitability under rain-fed conditions with a mixed level of inputs for our analysis, to best approximate historical agricultural conditions.

This study is related to the strand of literature that attempts to identify potential deep-rooted determinants of democratization. Giuliano and Nunn (2013) demonstrate the persistent influence of historical democratic practices within ethnic groups on current political institutions. Galor and Klemp (2017) posit that genetic diversity may contribute to the formation of pre-colonial autocratic institutions in indigenous groups and have a lasting effect on contemporary national institutions. Bentzen et al. (2017) identify a positive association between irrigation and autocracy and attribute it to the monopoly power held by elites over water supply and land, as the need for irrigation exacerbates such inequality and thus creates greater resistance from the elite class towards democratization. Our research is related to that since rice farming culture entails that of irrigation. However, while their research offers an explanation from the angle of resource constraints and competition, our study focuses on how agricultural legacies as such may influence institutional development from a broader perspective of cultural formation and transmission, and our empirical results are consistent with the proposed hypothesis (see Section 1.5.7).

Our work contributes to a burgeoning body of literature that seeks to understand the role played by agricultural legacies in economic and institutional development. Using the ratio of land suitable for growing wheat relative to that for sugarcane, Easterly (2007) provides empirical evidence on the hypothesis that agricultural endowments affect inequality and consequently the development in institutions and education. Nunn and Qian (2011) focus on land suitability for growing potatoes and exploit both regional and time variations in the introduction of potatoes to the Old World, to quantify the effects of potato cultivation on population growth and urbanization in the 18th and 19th centuries through the channel of improved nutrition. Alesina et al. (2013) examine the impact of plough agriculture on historical gender division of labor and its persistence on gender norms in contemporary society via cultural transmission. Ang and Fredriksson

(2017) suggest that wheat cultivation culture is negatively correlated with the strength of family ties, which in turn may impact on other social and economic outcomes.

The rest of the paper is organized as follows. Section 1.2 details wetland rice cultivation and the formation of collectivism under such culture. Section 1.3 describes the empirical approach, data sources and variables. Section 1.4 presents the main empirical findings. Section 1.5 presents findings of robustness checks and sensitivity tests. Lastly, Section 1.6 concludes.

1.2 Background: wetland rice cultivation and collectivism

Rice cultigens belong to genus *Oryza*, which was believably originated from Gondwana, the ancient supercontinent approximately 130 million years ago (Khush, 1997). Archaeological evidence suggests that rice farming could be traced back to 8,000 BC, placing rice among the first few cereal crops that were domesticated by ancient civilizations³ (Sweeney and McCouch, 2007). Today, rice is the primary source of nourishment and an essential staple food for approximately a third of the world's population (Khush, 1997).

Compared with other cereal crops such as wheat and barley, rice farming is highly labor-intensive and routinely requires cooperation and coordination amongst farmers during processes of planting / transplanting, tending and harvesting (Bray, 1994; Chauhan et al., 2017). Typically, the traditional method of cultivating rice requires flooding the fields in preparation for transplanting seedlings⁴. This demands careful planning for damming and channeling the water source, controlling the duration of flooding and monitoring of standing water depth. Additionally, farmers often use irrigation to supplement water supply from rainfall since rice cannot grow without abundant supply of water. Harvesting is manual and entails cutting, stacking, handling, threshing, cleaning and hauling; good harvesting techniques are effortful and crucial for grain yield maximization. The timing of transplanting is often staggered for different families, so that farmers can help each other meet peak demand for labor during planting and harvesting. It was estimated that growing rice requires more than twice

³Strictly speaking, there are two major categories of rice plants, namely the wetland and upland species. The ecological requirements of farming the two bear no resemblance to each other, as the former is a semi-aquatic plant while the latter is able to grow in dry land conditions. The production of upland rice constitutes only approximately 4% of the world's total production of rice (Muthayya et al., 2014). Our discussion is confined to wetland rice cultivation.

⁴The flooding procedure also serves as an effective control for pests and weeds.

the amount of man-hours as growing wheat (Buck et al., 1937).

Given the nature of rice cultivation, it comes as no surprise that rice-based culture fosters formation and transmission of a more collectivist culture in society, since it literally takes a village to grow rice. The hypothesis of rice farming legacy as the agricultural origin of collectivist norms was made prominent by a recent psychological study of Talhelm et al. (2014). Termed as the “rice theory of culture”, the study extends the argument of subsistence style theory and goes on to suggest that the mode of food production, in this case, the disaggregated types of farming required for different crops such as rice and wheat, are the root of stark differences in cultural orientation observed today. Rice cultivation leads to a more reciprocal and collegial relationship between farmers, efforts to reduce conflict and ultimately, to the notion that tends towards collectivism and holism. A collectivist culture places its emphasis on “group harmony, the prioritization of collective goals over personal goals, and the definition of one’s self in terms of the groups one belongs to” (Goncalo and Staw, 2006), which is in stark contrast to an individualistic culture that values self-reliance, personal freedom and achievement (Markus and Kitayama, 1991).

Table 1.1 presents the pairwise correlations between our main explanatory variable, i.e., land suitability for rice cultivation and major cultural dimensions as defined by Hofstede (1984) and Hofstede (2010). The cultural measures were constructed based on data from worldwide surveys of IBM employees’ cultural values in 1960s and subsequently expanded in 2010 to cover more countries. The Hofstede’s cultural dimensions have since been widely deployed in cross-cultural psychological studies.

Table 1.1: Correlation matrix

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
(1) <i>Rice suitability</i>	1 (143)						
(2) Individualism	-0.408*** (89)	1 (89)					
(3) Power distance	0.267* (89)	-0.664*** (89)	1 (89)				
(4) Masculinity	0.0157 (89)	0.0679 (89)	0.109 (89)	1 (89)			
(5) Uncertainty avoidance	-0.00796 (89)	-0.205 (89)	0.155 (89)	0.0393 (89)	1 (89)		
(6) Long-term orientation	-0.134 (89)	0.278* (76)	-0.150 (76)	0.0421 (76)	0.0965 (76)	1 (88)	
(7) Indulgence	0.152 (82)	0.133 (71)	-0.226 (71)	-0.0174 (71)	-0.158 (71)	-0.422*** (82)	1 (82)

Notes: This table presents the correlation between wetland rice suitability ratio and major cultural dimensions as defined by Hofstede (2010). *, ** and *** denote significance at the 10%, 5%, and 1% levels, respectively. The numbers of observations are reported in parentheses.

In line with the proposed mechanism, rice suitability is negatively correlated with individualism, which is defined as “a preference for a loosely-knit social framework in which individuals are expected to take care of only themselves and their immediate families” (Hofstede, 2010), at the 1% significance level. Rice suitability is also positively correlated with power distance, which measures “the degree to which the less powerful members of a society accept and expect that power is distributed unequally” (Hofstede, 2010), though the relationship is much weaker (at the 10% level of significance). This is not surprising since cultural traits are often inter-connected (Alesina and Giuliano, 2015). In this case, individualism and power distance are negatively correlated with each other at the 1% level of significance, and studies have shown that more individualistic societies tend to have lower power distance (see, e.g., Cox et al., 2011, for a detailed discussion). There is no significant correlation between rice suitability and other major cultural dimensions.

1.3 Empirical approach

1.3.1 Estimation methodology

To evaluate the hypothesized influence of wetland rice farming culture on the degree of institutionalized democracy, we estimate the following regression model:

$$Democracy_i = \alpha + \beta * Rice\ suitability_i + Controls_i' * \gamma + \varepsilon_i \quad (1.1)$$

where $Democracy_i$ is the average score of institutionalized democracy for country i over the period of the study, i.e., 1960 to 2013, from the Polity IV project (Marshall et al., 2014), $Rice\ suitability_i$ is defined as the percentage of arable land that is suitable or very suitable for wetland rice cultivation at the country level from the Global Agro-ecological Zones database (FAO and IIASA, 2000), $Controls_i$ is a vector of control variables and ε_i the error term. We use a set of geographical controls, which includes absolute latitude, landlocked dummy, distance to navigable river or coast, mean elevation, terrain ruggedness, temperature and precipitation in our regression analysis. These are commonly used in empirical studies of growth. In addition, continental fixed effects are included to account for potential omitted variable biases of continent-specific characteristics.

1.3.2 Data

1.3.2.1 Outcome measures

We use data on the average scores of institutionalized democracy over the period of 1960 to 2013 from the Polity IV project as the main measure of democracy at the country level. Under this framework, the extent of institutionalized democracy is measured based on its three fundamental aspects, namely, the institutions and procedures for citizens to express effective preferences, institutionalized constraints on the executive, and the assurance of civil liberties (Marshall et al., 2014). The democracy indicator classifies regimes on an additive 11-point scale ranging from zero to ten, where a greater value reflects a higher degree of institutionalized democracy.

Our main model focuses on the post-1960 era, as most former European colonies in regions of South and East Asia, Middle East and Sub-Saharan Africa had become independent sovereign states since then, and as such, have had the opportunity to develop their own sovereign institutions. We take the average score of democracy scores over the considered period 1960-2013 to ensure that Democ measures the fundamental and structural levels of democracy. Glaeser et al. (2004), for instance, argue that Polity IV's democracy measure is overly influenced by previous elections and is overly volatile and, therefore, fails to capture the structural procedures, norms and rules. Taking the average of democracy scores over five decades overcomes this concern. Figure 1.1 displays the average scores of institutionalized democracy over the period of 1960 to 2013 for all countries in our sample.

1.3.2.2 Main explanatory variable

Our main explanatory variable uses measures of the wetland rice suitability at the country level, which are constructed using crop yield data from the Global Agro-ecological Zones database (GAEZ) jointly developed by the Food and Agriculture Organization under the United Nations and the International Institute for Applied System Analysis (IIASA) (FAO and IIASA, 2000).

The estimates of crop yields in the GAEZ database are constructed following the Agro-ecological Zones (AEZ) methodology. AEZ is a framework intended for rational planning of land use, by using information on land resources and assessment of biophysical restrictions and potentials. Based on information pertaining to characteristics of respective crops, the FAO first identifies the required environmental conditions for their cultivations. Next, environmental conditions that are relevant to agricultural

Figure 1.1: Spatial distribution of institutionalized democracy scores (1960 - 2013)

Notes: The score of institutionalized democracy measures the relative degree of democracy of countries. Larger sizes of the dots represent higher degrees of democracy.
Source: The Polity IV project (Marshall et al., 2014).

production, which include climatic, soil and topographic conditions, are standardized under the AEZ framework. The climatic data are taken from the database developed by the Climate Research Unit at the University of East Anglia, while the land data are from the FAO/UNESCO Digital Soil Map of the World and terrain data are from the database compiled by the U.S. Geological Survey EROS Data Center. Following that, the FAO performs crop modeling and environmental matching algorithms to compute the maximum potential and agronomically attainable crop yields for basic land resources units, i.e., grid cells of 0.5 degrees by 0.5 degrees, under different levels of inputs and management conditions respectively (Fischer et al., 2000).

For crop suitability, data are derived by computing the percentage of maximum crop yield attainable in each basic land resources unit. The FAO further categorizes suitability data into five groups: the very suitable (80 - 100%), suitable (60 - 80%), moderately suitable (40 - 60%), marginally suitable (20 - 40%) and not suitable (0 - 20%). Country-level measures of crop suitability are constructed by data aggregation.

In this study, we define the wetland rice suitability ratio as the percentage of arable land that is suitable or very suitable for wetland rice cultivation in a given country, thus

our measure covers land with yields of at least 60% of the maximum attainable value. Furthermore, we choose data under rain-fed conditions with a mixed level of inputs, to best approximate historical agricultural conditions: the rain-fed conditions ensure that the source of water supply is less likely to be affected by human intervention on that, for instance, efficacy of different irrigation methods; while the mixed level of inputs reflects the level of inputs and farming technologies deployed to best approximate the actual practice. Under the mixed level of inputs, the extent of land with cultivation potential for rain-fed crops is estimated by assuming that the most suitable land is deployed under high-level input farming, moderately suitable land under intermediate-level input farming and marginally suitable land under low-level input management. Thus, data on cultivation potential under mixed-level inputs reflect the land suitability for cultivation under the assumption of effective resource allocation⁵.

Although the AEZ framework uses data on environmental conditions for the period of 1960 to 1996 in its computation, it is suggested that the crop suitability ratios derived as such should serve as good proxies for historical levels of crop suitability (see, e.g., Nunn and Qian, 2011, for a detailed discussion). Essentially, the major environmental conditions concerned with agricultural production, such as climatic and soil conditions, do not evolve drastically over time. Thus, in our case, the main explanatory variable, i.e., rice suitability ratio, should be sufficiently exogenous to the formation and transmission of collectivist culture and values.

While we do not have historical data on rice production to perform a direct test on how well our measure captures the historical conditions of rice farming, a quick check using data on current rice production from the FAO database indicates that there is a strong positive correlation between rice suitability ratio and current rice production, as the correlation coefficient is 0.58 and statistically significant at the 1% level⁶. Together with the foregoing discussion on rice suitability ratio, this strong correlation offers further reassurance that our rice suitability measure would serve as a reasonably good

⁵There are four types of input levels available in the GAEZ database, namely, the low, intermediate, high and mixed levels of inputs. The first three are self-explanatory and correspond to the assumption that the respective level of farming management and application of production techniques (e.g., use of fertilizers, pest controls etc) deployed to all arable lands, whereas the mixed level of inputs involves a mixture of low-, intermediate- and high-levels of inputs deployed on lands with different extents of suitability as aforementioned.

⁶We use the FAO data on rice production and take the average of annual rice production over the period of this study (i.e., 1960 - 2013). Following the approach of Nunn and Qian (2011), we normalize both variables, i.e., rice production and land suitability for growing rice by the total area of arable land of a country.

proxy for rice farming legacy. Figure 1.2 displays data of the wetland rice suitability ratio for all countries in our sample.

Figure 1.2: Spatial distribution of wetland rice suitability ratio

Notes: The wetland rice suitability ratio is the share of arable land that is suitable or very suitable for wetland rice cultivation of countries. Larger sizes of the dots represent higher suitability ratios of rice cultivation.

Source: The Global Agro-ecological Zones database (FAO and IIASA, 2000).

Summary statistics of the main variables are presented in Table 1.2. Our data cover a total of 143 countries across five continents.

Table 1.2: Summary statistics

Variable	Observed	Mean	SD	Minimum	Maximum
Institutionalized democracy score	143	4.18	3.48	0.00	10.00
Rice suitability ratio	143	0.05	0.08	0.00	0.59
Absolute latitude	143	0.30	0.19	0.01	0.71
Landlocked (dummy)	143	0.24	0.43	0.00	1.00
Distance to coast or river (10 ³ km)	143	0.34	0.44	0.01	2.39
Elevation	143	176.07	191.78	2.39	1096.50
Terrain ruggedness (index)	143	1.29	1.23	0.04	6.74
Temperature	143	18.05	8.24	-7.82	28.25
Precipitation	143	86.73	59.67	2.91	259.95

Notes: Refer to the Appendix for descriptions of all variables.

1.4 Empirical results

1.4.1 Main regression results

Our hypothesis is that rice cultivation culture tends to foster collectivism and hence reduces the degree of institutionalized democracy in contemporary societies. This is supported by results of main regressions on the relationship between the rice suitability ratio (hereinafter referred to as “rice suitability”) and the average score of institutionalized democracy over the period of our study (hereinafter referred to as “Democ”) at the country level, as presented in Table 1.3. Standardized beta coefficients are reported for the ease of interpretation, as all variables are normalized to have a mean of zero and a standard deviation of one.

The results of estimating Eq. 1.1 are shown in Table 1.3. The coefficient of rice suitability is negative and statistically significant in the regression in the first column when continental dummies are excluded from the regressions. This finding is unaltered when continent fixed effects are added in Column (2), where the specification in Column (2) is taken as the baseline regression in subsequent analyses. Of the control variables, only the coefficients of distance to river, elevation and precipitation are consistently significant. The coefficients of latitude and precipitation are significantly positive in Column (1) but their significance weakens or disappears once continent fixed effects are included in the regressions and the standard errors are clustered between continents. This suggests that autocracy is not more common in countries close to the equator because of absolute latitude per se, but rather because they are located in some continents which are more prone to having autocratic regimes. This result is consistent with the findings of Madsen et al. (2015) and Acemoglu et al. (2019) that democratization waves

tend to cluster geographically.

Results of the baseline model suggest that an increase of one standard deviation ($SD = 0.08$) in the rice suitability ratio leads to a decrease in the degree of institutionalized democracy by 0.138 standard deviations ($SD = 3.48$), holding all other variables constant. Therefore, variations in rice suitability are able to explain a reasonable portion of differences in the degree of democracy across countries. For instance, Cambodia has a relatively high rice suitability ratio of 0.40 (the maximum value is 0.59) and a low Democ score of 1.6 (the maximum value is 10). Based on the baseline estimation, if Cambodia were to have a lower rice suitability ratio such as 0.17, which is similar to that of Brazil, then the estimated Democ score would be 4.7, which is substantially higher than the current Democ score. The estimated Democ score is also similar to the actual score of Brazil, which is 5.1.

In the Appendix, we also report results of regressing democracy on individualism using rice suitability as an instrument of individualism, which serves to check the extent to which rice suitability affects democracy through culture. From the results, it can be inferred that rice suitability, among other factors, has been a fundamental deterrent for the formation of democratic institutions as rice suitability promoted collectivist cultures (see Table 1.10).

Table 1.3: Main regression results

Dep. Var. = Democ	(1)	(2)
		Baseline specification
<i>Rice suitability</i>	-0.189*** (-2.614)	-0.138** (-4.554)
Latitude	0.587*** (-3.365)	0.408 (-1.741)
Landlocked	0.007 (-0.098)	-0.004 (-0.040)
Distance to river	-0.373*** (-5.509)	-0.274*** (-5.675)
Elevation	0.209*** (-2.696)	0.209** (-4.394)
Terrain ruggedness	-0.132* (-1.817)	-0.05 (-1.396)
Temperature	-0.147 (-0.802)	-0.059 (-0.344)
Precipitation	0.519*** (-5.165)	0.347* (-2.697)
Continent FE	No	Yes
Cluster s.e.	No	Yes
R-squared	0.518	0.585
Observations	143	143

Notes: This table reports the correlation between rice suitability and institutional democracy score. Standardized beta coefficients of regressions are presented, heteroskedasticity robust standard errors are used and *t* statistics are reported in parentheses. *, ** and *** denote significance at the 10%, 5% and 1% levels, respectively. The intercept estimates are included in the regressions but not shown. The continent dummies are Asia, Europe, America, Oceania and Africa.

1.5 Robustness checks

1.5.1 Robustness to unobservables

In this subsection, we examine the influence of unobservables to further establish the robustness of our main results, using the approach of coefficient stability proposed by Altonji et al. (2005) and subsequently developed by Oster (2017). The coefficient of proportionality, denoted as δ , measures the relative ratio of the impact of unobservables to that of observables such that the unobservables would have an equally important impact as observables on the coefficient of the main explanatory variable. Specifically, under the assumption of proportional selection on observables, δ is calculated as follows:

$$\delta = \frac{(\beta^* - \tilde{\beta})(\tilde{R} - \hat{R})}{(R_{max} - \tilde{R})(\tilde{\beta} - \hat{\beta})} \quad (1.2)$$

where $\tilde{\beta}$ and $\tilde{R}$ are the estimated coefficient of interest and the R-squared value associated with the controlled regression respectively, $\hat{\beta}$ and $\hat{R}$ for the uncontrolled regression, and β^* and R_{max} for the hypothetical regression that includes both observed and unobserved controls. Oster (2017) derives a bounded value of $R_{max} = 1.3\tilde{R}$ for this method and argues that a value of δ greater than one (at $\beta^* = 0$) would indicate that results are robust to omitted variable bias. In our case, the estimated δ is 1.15, thus suggesting that the OLS effect in the baseline model is unlikely to be driven by the unobservables.

1.5.2 Robustness to early development, alternative channels and historical institutions

The literature on long-run comparative economic development suggests that early historical conditions may play an important role in explaining current social and economic outcomes (Nunn, 2009). Table 1.4 presents the regression results when considering these historical factors.

In the upper panel of Table 1.4, we consider several indicators of early development, as well as alternative channels or other potential deep roots of collectivism that consequently may impact on contemporary democracy. First, we control for effects of some major agricultural legacies, namely, the timing of agricultural transition and biogeography in Columns (1) and (2). The timing of agricultural transition captures the estimated time elapsed since the transition from hunter-gatherer to agrarian societies, and has been shown to influence long-run economic comparative development via institutions (Putterman, 2004; Ang, 2013); while biogeographic endowment is closely associated

Table 1.4: Controlling for effects of some historical measures

Dep. Var. = Democ	Indicators on early development / alternative channels					
	(1)	(2)	(3)	(4)	(5)	(6)
<i>Rice suitability</i>	-0.154*** (-5.301)	-0.158** (-4.467)	-0.143** (-4.207)	-0.140*** (-4.802)	-0.151*** (-5.225)	-0.137** (-4.496)
Timing of agricultural transition	-0.136 (-0.748)					
Biogeography		-0.093 (-0.307)				
Population density in 1500 AD (log)			0.030 (0.229)			
Pathogen stress				-0.071 (-0.431)		
Human settlement history					0.123 (1.344)	
Predicted genetic diversity						0.017 (0.378)
Baseline control	Yes	Yes	Yes	Yes	Yes	Yes
Continent FE	Yes	Yes	Yes	Yes	Yes	Yes
R-squared	0.588	0.580	0.591	0.625	0.588	0.585
Observations	141	133	142	131	142	143

Dep. Var. = Democ	Historical institutions and colonization						
	(7)	(8)	(9)	(10)	(11)	(12)	(13)
<i>Rice suitability</i>	-0.154*** (-4.873)	-0.139** (-4.127)	-0.139** (-3.502)	-0.105** (-2.892)	-0.142*** (-5.078)	-0.227** (-3.782)	-0.130** (-4.267)
State history up to 1500 AD	-0.080 (-0.557)						
State history up to 1500 AD (ancestry-adjusted)		0.025 (0.165)					
Traditional local democratic practices			0.072 (1.522)				
European language (% of population)				0.309* (2.299)			
Colony indicator					0.200*** (8.207)		
Duration of colonization						-0.364 (-0.818)	
Indicator of colonial origin:							
Spanish							-0.038 (-0.889)
British							0.153* (2.651)
French							-0.055 (-0.690)
Portuguese							-0.016 (-0.461)
Other European							-0.002 (-0.045)
Baseline control	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Continent FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
R-squared	0.586	0.585	0.585	0.629	0.586	0.485	0.613
Observations	138	143	134	118	135	84	143

Notes: This table reports the standardized beta coefficients of regressions under different specifications. Robust standard errors are clustered at the continent level and *t* statistics are reported in parentheses. *, ** and *** denote significance at the 10%, 5% and 1% levels, respectively. The intercept estimates are not shown. The baseline controls used under the full specification include absolute latitude, landlocked dummy, distance to navigable river or coast, mean elevation, terrain ruggedness, temperature and precipitation. The continent dummies are Asia, Europe, America, Oceania and Africa.

with agricultural productivity and has a persistent impact on economic development since prehistory (Olsson and Hibbs Jr, 2005).

Second, we include the natural logarithm of population density in 1500 AD, which is a basic measure for the level of prosperity for an economy in the Malthusian epoch in Column (3). Vandello and Cohen (1999) suggest that high population density tends to associate with a greater need for coordination and lower incidence of conflicts, thus promoting the formation of collectivism. Accordingly, the population density in 1500 AD could be another deep root of collectivism in addition to rice suitability.

Third, we control for the historical pathogen prevalence in Column (4). The historical pathogen prevalence index covers nine diseases including leishmania, schistosoma, trypanosoma, malaria, filaria, leprosy, dengue, typhus and tuberculosis (Schaller and Murray, 2008). The literature on the pathogen prevalence theory suggests that high parasite stress, together with the associated high level of communicable diseases, discourage natives from interacting with people from external areas. This renders societies more conservative and exclusive, and in turn favors the formation of collectivism (Fincher et al., 2008; Thornhill and Fincher, 2014), which serves as a defense strategy against pathogens in this case. Thus, historical pathogen stress could also be a potential deep root of collectivism. Furthermore, Acemoglu et al. (2001) find that disease burden may play an important role in shaping the path of institutional development.

Fourth, Ahlerup and Olsson (2012) argue that the duration of human settlement in a given area is a key determinant of ethnolinguistic diversity, which have significant impacts on economic and political development in contemporary societies. We account for the effect of this variable in Column (5).

Fifth, Galor and Klemp (2017) conjecture that genetic diversity contributes to the formation of pre-colonial autocratic institutions in indigenous groups, and consequently influences contemporary national institutions. We control for the predicted genetic diversity in Column (6), which is an index constructed using information on expected heterozygosity at the population level and the ancient migration distance from East Africa by Ashraf and Galor (2013).

Our regression results show that none of these additional controls are statistically significant, while effects of rice suitability on Democ remain robust in all cases. This further supports our hypothesis that rice cultivation is the major deep root of collectivism, whose effects on the institutionalized democracy persist in contemporary societies.

In the lower panel of Table 1.4, we examine the effects of indigenous institutions

prior to major colonization and the influence of colonization, respectively. In Column (7), we consider a measure for historical state development from 1 AD to 1500 AD in the analysis, which captures information pertaining to the history of nationhood and state capacity and has been shown to be a key determinant of economic development in 1500 AD by Putterman (2008). Additionally, as Putterman and Weil (2010) suggest that the history of a population's ancestors may have stronger predictive power for current economic outcomes, compared with that of their present residences, we control for the ancestry-adjusted version of this historical variable in Column (8), which reflects the state history of contemporary populations' ancestors. This is done by pre-multiplying the original variable with the ancestry matrix created by Putterman and Weil (2010), which contains the proportion of people in each country in 2000 AD that was originated from various source countries in 1500 AD. Accordingly, the constructed variable represents the weighted average of its contemporary population's ancestral historical variables. In Column (9), we control for the weighted average of traditional local democratic practices within ethnic groups at the country level, as Giuliano and Nunn (2013) suggest that the effects of historical democratic practices within ethnic groups on political institutions persist until today. Local democratic practices measure whether local village headmen were chosen by election and/or informal consensus rather than by less democratic succession rules. All measures of state history or traditional democratic practices are not significant in this case.

Next, we consider the potential influence of colonization on democracy. The general consensus is that colonization is a deep determinant of modern economic development and political institutions, and its effects are persistent (see, e.g., Acemoglu et al., 2001, 2002; Olsson, 2009). In Column (10), we control for the share of population that speaks European languages, which serves as a proxy for the influence of European colonization; while in Column (11) we use an indicator variable, which takes the value of one if the country was a colony and zero if otherwise. Both are statistically significant and positively correlated with *Democ*.

Furthermore, we also consider the impact of duration of colonization in Column (12), as Olsson (2009) suggests that the duration of colonization may matter for the degree of democracy in contemporary societies. In our case, the duration of colonization does not appear to have a significant effect on *Democ*.

Lastly, we zoom in on the impact brought by different origins of colonizers on contemporary democracy, using an indicator variable that distinguishes the origin of

colonizers as Spanish, British, French, Portuguese or other European in Column (13). As suggested by Lange et al. (2006), European colonizers differed largely in their colonial strategies and governance, which in turn left different colonial legacies pertaining to institutions. Particularly, institutions such as legal systems that were introduced to the colonies, remained in force even after the independence of the colonies. Our regression results indicate that only the coefficient of British colonial origin appears to be significant and positively correlated with *Democ*, while the rest are not significant. This finding is in line with findings of Lange et al. (2006), which show that the British colonizer tended to be more market-oriented compared with others, and this in turn led to a more conducive environment for germination of democratic ideas in British colonies.

The results reported in Table 1.4 indicate that the negative correlation between rice suitability and *Democ* remains highly significant at the 5% level with inclusion of indicators on early development and historical institutions.

1.5.3 Robustness to effects of some contemporary measures

This subsection reports regression results that consider the effects of various contemporary measures, which include income levels, urbanization, education, availability of natural resources, measures of social diversity and religious adherence as control variables in Table 1.5.

First, Lipset (1959) posits that economic modernization begets democracy, since “economic development carries with it the political correlate of democracy”. Boix (2003) suggests that education plays an important role in the establishment of institutionalized democracy. Therefore, we consider GDP per capita (in natural logarithm) in the starting year of our study, 1960, the extent of urbanization in 1960 using the percentage of population residing in cities, which is another key indicator of economic development and modernization, as well as the average years of schooling in 1960 as our control variables in Columns (1) to (3), respectively. In this case, urbanization and education are positively correlated with *Democ*, while the coefficient of GDP per capita is not statistically significant.

Second, a large body of literature on political economy suggests that economic and fiscal dependence on natural resources such as oil and minerals, tends to create authoritarian regimes and hinder democratization (see, e.g., Mahdavy and Cook, 1970; Ross, 2001; Wantchekon, 2002; Jensen and Wantchekon, 2004). We therefore also

Table 1.5: Controlling for some contemporary measures

Dep. Var. = Democ	(1)	(2)	(3)	(4)	(5)
<i>Rice suitability</i>	-0.202** (-3.516)	-0.118* (-2.315)	-0.167** (-3.818)	-0.148** (-4.249)	-0.149*** (-7.652)
GDP per capita (log) (1960)	-0.011 (-0.107)				
Urbanization (1960)		0.328** (3.678)			
Education (1960)			0.477** (4.006)		
Oil rents (% of GDP)				-0.217* (-2.482)	
Mineral rents (% of GDP)					-0.095 (-1.858)
Baseline control	Yes	Yes	Yes	Yes	Yes
Continent FE	Yes	Yes	Yes	Yes	Yes
R-squared	0.683	0.624	0.677	0.620	0.592
Observations	85	143	123	143	143
Dep. Var. = Democ	(6)	(7)	(8)	(9)	(10)
<i>Rice suitability</i>	-0.148** (-4.158)	-0.138** (-4.492)	-0.093** (-3.112)	-0.140** (-4.464)	-0.142** (-3.797)
Diamond production	0.086 (1.473)				
Ethnic fractionalization		0.005 (0.058)			
Language fractionalization			0.125* (2.448)		
Religion fractionalization				0.068 (0.820)	
Population shares of religious adherents:					
Protestants (%)					0.218*** (5.097)
Muslims (%)					-0.185 (-1.555)
Catholics (%)					0.118 (1.016)
Buddhists (%)					0.001 (0.019)
Baseline control	Yes	Yes	Yes	Yes	Yes
Continent FE	Yes	Yes	Yes	Yes	Yes
R-squared	0.590	0.583	0.604	0.589	0.639
Observations	143	142	139	143	143

Notes: This table reports the standardized beta coefficients of regressions under different specifications. Robust standard errors are clustered at the continent level and *t* statistics are reported in parentheses. *, ** and *** denote significance at the 10%, 5% and 1% levels, respectively. The intercept estimates are not shown. The baseline controls used under the full specification include absolute latitude, landlocked dummy, distance to navigable river or coast, mean elevation, terrain ruggedness, temperature and precipitation. The continent dummies are Asia, Europe, America, Oceania and Africa.

account for potential confounding effects of natural resource rents in our analysis. In Columns (4) to (6), we control for the average oil rents (as a share of GDP), mineral rents, and an indicator for diamond production over the period of the study. Only oil rents appear to be negatively correlated with *Democ*, while coefficients of mineral rents and the diamond indicator are not statistically significant.

Third, we consider the diversity measures from Alesina et al. (2003) including ethnic, linguistic and religious fractionalization in our regressions (in Columns (7) to (9)). Diversity measures are commonly controlled in research that attempts to explain cross-country differences in economic development (see, e.g., Brock and Durlauf, 2001). However, past studies have reported mixed results on the potential influences of population diversity on institutions and economic outcomes. For instance, while Easterly and Levine (1997) suggest that ethnic and linguistic diversity is negatively correlated to per capita GDP growth, Alesina and Ferrara (2005) suggest that such effects of population diversity are conditional on other societal features. Our regression results show that linguistic fractionalization appears to be positively correlated with *Democ*, while other measures are not statistically significant.

Fourth, we examine the influence of religions on democracy. The important role of religions on modern economic development is often highlighted by scholars. For instance, Max Weber (1930) hails the ethics of ascetic Protestantism as the source of capitalistic spirit. Adherents of the same religion also tend to share similar cultural values, which in turn influence the formation of political and social institutions. Building on these insights, we control for the population share of adherents to the major religions, which include Protestantism, Islam, Catholicism and Buddhism in Column (10). Our results show that there is a positive correlation between the population share of Protestants and *Democ*, while estimates of other religious influences are not statistically significant. This is consistent with findings in the current literature, which suggest that Protestantism helps promulgate moral and cultural values such as individualism that are conducive to the formation of democracy (Bruce, 2004; Woodberry, 2012).

The hypothesized negative relationship between rice suitability and *Democ* remains robust in all the above cases. Our main findings are not undermined by including these contemporary controls in the analysis.

1.5.4 Alternative measures of democracy

To further ascertain the robustness of our results, we consider a number of alternative measures of democracy in this subsection and Table 1.6 reports the regression findings.

Table 1.6: Alternative measures of democracy

	(1)	(2)	(3)	(4)	(5)
Dep. Var. =	Democ 1800-2013	Democ 1900-2013	Democ 1930-2013	Democ 1985-2013	Polity2 1960-2013
<i>Rice suitability</i>	-0.153** (-4.599)	-0.144** (-3.995)	-0.145** (-4.057)	-0.099* (-2.502)	-0.103** (-3.431)
Baseline control	Yes	Yes	Yes	Yes	Yes
Continent FE	Yes	Yes	Yes	Yes	Yes
R-squared	0.442	0.526	0.543	0.605	0.548
Observations	143	142	142	143	142
	(6)	(7)	(8)	(9)	(10)
Dep. Var. =	Polity2 1800-2013	Electoral democracy 1960-2013	Liberal democracy 1960-2013	Deliberative democracy 1960-2013	Support vector machine democracy index 1981-2011
<i>Rice suitability</i>	-0.128** (-3.613)	-0.083* (-2.172)	-0.109* (-2.544)	-0.094* (-2.428)	-0.105** (-3.435)
Baseline control	Yes	Yes	Yes	Yes	Yes
Continent FE	Yes	Yes	Yes	Yes	Yes
R-squared	0.393	0.615	0.618	0.571	0.647
Observations	142	142	142	142	143

Notes: This table reports the standardized beta coefficients of regressions under different specifications. Robust standard errors are clustered at the continent level and *t* statistics are reported in parentheses. *, ** and *** denote significance at the 10%, 5% and 1% levels, respectively. The intercept estimates are not shown. The baseline controls used under the full specification include absolute latitude, landlocked dummy, distance to navigable river or coast, mean elevation, terrain ruggedness, temperature and precipitation. The continent dummies are Asia, Europe, America, Oceania and Africa. In the upper panel, the dependent variables are the historical averages of Democ scores from the year 1800 to 2013 in Column (1), from the year 1900 to 2013 in Column (2), from the year 1930 to 2013 in Column (3), from the year 1985 to 2013 in Column (4), and the historical average of Polity2 scores from the year 1960 to 2013 in Column (5), respectively. In the lower panel, the dependent variables are the historical average of Polity2 scores from the year 1800 to 2013 in Column (6), the average of the electoral democracy index (V-Dem) from the year 1960 to 2013 in Column (7), the average of the liberal democracy index (V-Dem) from the year 1960 to 2013 in Column (8), the average of the deliberative democracy index (V-Dem) from the year 1960 to 2013 in Column (9), and the average of support vector machine (SVM) democracy index from the year 1981 to 2011 in Column (10), respectively.

First, we use the average of Democ scores from the year 1800 to 2015 in Column (1), as 1800 is the earliest year that data on Democ are available. In Columns (2) to (4), we deploy historical averages of Democ with different starting years, namely, 1900, 1930 and 1985 as our dependent variables in the analysis. Overall, our measure of rice suitability remains statistically significant at least at the 10% level, regardless of different periods of time specified for the averages of Democ. This confirms that our main empirical results are unlikely to be driven by the choice of Democ for a specific period as the dependent variable. The effects of rice suitability on institutionalized

democracy are persistent over time.

Next, we deploy the polity2 index from the Polity IV project (Marshall et al., 2014) as an alternative measure of contemporary institutions. The polity2 index is a composite index commonly used in quantitative research on general regime effects. We use the average of the polity2 index for the same period as our baseline model (i.e., 1960 to 2013) in Column (5) and the historical average for the period of 1800 to 2013 in Column (6), respectively.

Additionally, we consider several alternative indices of democracy from the Varieties of Democracy (V-Dem) database, namely, the electoral democracy, liberal democracy and deliberative democracy to verify our main results in Columns (7) to (9). These indices place emphasis on different fundamental principles of modern democracy and measure the quality of these aspects. First, electoral democracy falls under the umbrella of representative democracy, which is an indirect democracy whereby the people exercise their rights of political participation through their elected representatives. This index measures the quality of democracy by the extent of suffrage, cleanness and effectiveness of election process, freedom of expression and presence of an independent media. Next, the measure of liberal democracy focuses on the protection of civil liberties and properties by constitutions, strong rule of law, executive constraints as well as the extent of electoral democracy. Lastly, the index of deliberative democracy emphasizes the importance of deliberation in polity decision-making process and measures the quality of discussions at all levels that lead to political decisions (Coppedge et al., 2015).

Finally, we include a novel democracy index generated using support vector machine (SVM) algorithms by Gründler and Krieger (2016) in the robustness checks. SVM is a type of machine learning models commonly used for classification and regression analysis. This methodology was introduced by Gründler and Krieger (2016) into the quantification of democracy to mitigate issues of arbitrariness and potential measurement problems associated with the aggregation rules of most existing democracy indices (see Munck and Verkuilen, 2002; Treier and Jackman, 2008, for a detailed discussion). Using core elements of democracy, which include political participation, political competition and civil rights as the input attributes, and major democracy indices as the information source and training data, Gründler and Krieger (2016) calculate the optimal support vector regression function and derive the SVM democracy index by iterations. The constructed democracy index thus captures information on existing major democracy indices. We use data on the SVM index of all its available years (i.e., 1981 to 2011)

in the analysis in Column (10).

In all the above cases, the negative correlation between rice suitability ratio and the degree of institutionalized democracy remains robust to different definitions of democracy measures. The coefficient of rice suitability is at least statistically significant at the 10% level for alternative measures of democracy scores.

1.5.5 Region exclusion analysis

To rule out the possibility that our main results may be driven by inclusion of a particular region, we perform sensitivity tests by excluding different regions in our analysis, and Table 1.7 presents the results. In the upper panel, we sequentially exclude observations that belong to a continent; while in the lower panel, we sequentially exclude large countries, which are China, India, Japan, the U.S., Canada and Russia, respectively. The coefficient of rice suitability remains statistically significant in all restricted sample estimations, suggesting that our main findings are robust to different regional coverage.

Table 1.7: Region exclusion analysis

	(1)	(2)	(3)	(4)	(5)	(6)
Dep. Var. = Democ	Exclude Asia	Exclude Europe	Exclude Africa	Exclude Oceania	Exclude N. America	Exclude S. America
<i>Rice suitability</i>	-0.134** (-3.353)	-0.166** (-4.024)	-0.147** (-3.318)	-0.163*** (-9.269)	-0.113* (-2.415)	-0.212* (-2.291)
Baseline control	Yes	Yes	Yes	Yes	Yes	Yes
Continent FE	Yes	Yes	Yes	Yes	Yes	Yes
R-squared	0.657	0.438	0.544	0.587	0.623	0.614
Observations	106	110	97	140	122	124
	(7)	(8)	(9)	(10)	(11)	(12)
Dep. Var. = Democ	Exclude China	Exclude India	Exclude Japan	Exclude United States	Exclude Canada	Exclude Russia
<i>Rice suitability</i>	-0.135** (-4.343)	-0.125** (-3.715)	-0.132** (-3.707)	-0.139** (-4.476)	-0.139** (-4.539)	-0.136*** (-4.628)
Baseline control	Yes	Yes	Yes	Yes	Yes	Yes
Continent FE	Yes	Yes	Yes	Yes	Yes	Yes
R-squared	0.587	0.601	0.593	0.584	0.577	0.589
Observations	142	142	142	142	142	142

Notes: This table reports the standardized beta coefficients of regressions under different specifications. Robust standard errors are clustered at the continent level and *t* statistics are reported in parentheses. *, ** and *** denote significance at the 10%, 5% and 1% levels, respectively. The intercept estimates are not shown. The baseline controls used under the full specification include absolute latitude, landlocked dummy, distance to navigable river or coast, mean elevation, terrain ruggedness, temperature and precipitation. The continent dummies are Asia, Europe, America, Oceania and Africa. In the upper panel, we sequentially exclude observations that belong to Asia, Europe, Africa, Oceania, North America and South America; while in the lower panel, we sequentially exclude China, India, Japan, the U.S., Canada and Russia, respectively.

1.5.6 Considering alternative cereal crops

In this subsection, we also examine the potential effects of cultivation cultures of other major cereal crops, which include wheat, barley and maize in Table 1.8. According to data from the FAO, together with rice, these crops were estimated to provide approximately 80% of total calories supplied by known cereals in the world for the year of 1961 (see Mayshar et al., 2015; Ang and Fredriksson, 2017, for more information). Compared with barley or wheat cultivation, maize cultivation shares similar characteristics with that of wetland rice, as both are relatively labor-intensive and require greater efforts of coordination amongst farmers. However, unlike rice, maize is a New World crop and its presence was confined to the American continent before 1500 AD, the Columbian exchange. Given its relatively short length of cultivation history in the rest of the world, maize would not be a suitable candidate for our purpose to better understand the formation of collectivist culture through agricultural legacies, in search for the historical roots of democracy.

Column (1) repeats results of the baseline model for the ease of comparison, while in Columns (2) to (4), we replace rice suitability with that of wheat, barley and maize as the main explanatory variable, respectively. All coefficients of suitability ratios for alternative cereal crops are not statistically significant.

We further perform a horserace analysis on the four cereal crops, where rice suitability is paired with wheat, barley or maize suitability respectively in Columns (5) to (7), and include all four suitability variables in the regression in Column (8). In all cases, only the coefficient of rice suitability is statistically significant at the 5% level, and its magnitude remains qualitatively unaffected.

1.5.7 Further robustness checks

In this subsection, we present findings of additional tests that further establish the robustness of our main hypothesis in Table 1.9. We consider effects of environmental factors related to rice production, agricultural techniques deployed in rice farming, cultivation of other labor-intensive crops, as well as using rice suitability ratios under different levels of inputs in the analysis.

First, we consider environmental conditions that are closely related to rice production and thus are potential confounders to our main explanatory measure, rice suitability. In Columns (1) and (2), we control for the overall soil suitability for agriculture and variation in precipitation, respectively. Our results remain qualitatively

Table 1.8: Analysis considering alternative cereals

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Dep. Var. = Democ	Baseline							
<i>Rice suitability</i>	-0.138** (-4.554)				-0.140** (-4.177)	-0.140** (-4.156)	-0.141** (-4.572)	-0.146** (-4.170)
Wheat suitability		-0.027 (-0.222)			-0.038 (-0.321)			-0.133 (-0.169)
Barley suitability			-0.027 (-0.211)			-0.038 (-0.307)		0.050 (0.058)
Maize suitability				0.105 (1.385)			0.107 (1.334)	0.121 (1.624)
Baseline control	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Continent FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
R-squared	0.585	0.575	0.575	0.583	0.585	0.585	0.594	0.596
Observations	143	143	142	142	143	142	142	142

Notes: This table reports the standardized beta coefficients of regressions under different specifications. Robust standard errors are clustered at the continent level and *t* statistics are reported in parentheses. *, ** and *** denote significance at the 10%, 5%, and 1% levels, respectively. The intercept estimates are not shown. The baseline controls used under the full specification include absolute latitude, landlocked dummy, distance to navigable river or coast, mean elevation, terrain ruggedness, temperature and precipitation. The continent dummies are Asia, Europe, America, Oceania and Africa.

unchanged with the inclusion of these environmental factors.

Next, we control for potential effects of different agricultural techniques. In Column (3), we consider a historical measure of plough use, since ploughing is typically required to prepare for rice planting. Additionally, Alesina et al. (2013) suggest that ploughing practice influences social outcomes such as gender norms via cultural transmission. Our regression results suggest that the use of plough does not have a significant effect on Democ.

A recent study by Bentzen et al. (2017) has suggested that irrigation may be a potential deep-root of autocracy that works through the channel of resource competition. As aforementioned, rice cultivation culture entails irrigation practices, thus the effect of rice farming on Democ could be confounded by that of irrigation. We address this issue by considering the historical measure of irrigation potential constructed by Bentzen et al. (2017) in Column (4). The coefficient of rice suitability remains statistically significant at the 1% level when controlling for irrigation potential.

Our main argument builds on the fact that the labor-intensive nature of rice farming gives rise to collectivism that acts as an adaption mechanism. To check whether cultivation of other labor-intensive crops would have similar effects like rice on fostering the emergence of collectivist culture, and consequently on impeding democratization, we also control for the effect of cultivation of other major labor-intensive crops identified by Fouka and Schlaepfer (2017), which include potato, maize and bean in Column (5).

Following the approach of Fouka and Schlaepfer (2017), we take the average of land suitability ratios for cultivation of these crops, weighed by the estimated labor intensity required for each crop and normalized by the sum of land suitability ratios of all three crops. The regression results indicate that effects of rice suitability on Democ remain highly significant while that for suitability of other labor-intensive crops is insignificant.

Lastly, our main explanatory measure is the rice suitability ratio under rain-fed conditions and mixed-level inputs; the mixed level of inputs is chosen to reflect the land suitability for rice cultivation under the assumption of effective resource allocation. To further test the robustness of our main hypothesis, we replace the main explanatory measure with land suitability ratios for wetland rice cultivation under rain-fed conditions, but with a low level or an intermediate level of inputs in Columns (6) and (7), respectively. This allows us to conduct counter-factual checks on the rice suitability, i.e., on the assumption that when resources are limited and only low- or intermediate-level inputs can be deployed to all arable lands; assuming the use of lower levels of inputs also further isolates the influence of farming technology from the rice suitability ratio, and in turn makes the rice suitability measures more exogenous since democracy might open up opportunities for innovation and technological diffusion. Our main results survive as coefficients of both alternative measures of rice suitability are negative and statistically significant at the 5% level. In line with the assumption of limited resources, the magnitudes of their coefficients decrease notably when lower levels of inputs are deployed, compared to that in the baseline model, since the cultivation potential of rice in these cases could only be partially realized.

Table 1.9: Additional robustness checks

Dep. Var. = Democ	(1)	(2)	(3)	(4)	(5)	(6)	(7)
<i>Rice suitability</i>	-0.132** (-3.315)	-0.124*** (-5.282)	-0.129* (-2.741)	-0.161*** (-5.862)	-0.164*** (-13.810)		
Soil suitability	0.072 (0.409)						
Precipitation variation		-0.055 (-0.437)					
Plough use			-0.074 (-0.328)				
Irrigation potential				-0.221 (-1.793)			
Other labor-intensive crops					-0.003 (-0.036)		
Rice suitability ratio (low inputs)						-0.092** (-2.797)	
Rice suitability ratio (intermediate inputs)							-0.116** (-3.506)
Baseline control	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Continent FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
R-squared	0.590	0.585	0.586	0.610	0.578	0.580	0.582
Observations	142	143	143	142	124	142	142

Notes: This table reports the standardized beta coefficients of regressions under different specifications. Robust standard errors are clustered at the continent level and *t* statistics are reported in parentheses. *, ** and *** denote significance at the 10%, 5% and 1% levels, respectively. The intercept estimates are not shown. The baseline controls used under the full specification include absolute latitude, landlocked dummy, distance to navigable river or coast, mean elevation, terrain ruggedness, temperature and precipitation. The continent dummies are Asia, Europe, America, Oceania and Africa.

1.6 Conclusion

This research has hypothesized that a rice growing legacy has deterred the evolution of democratic institutions. Through centuries, or even millenniums, collaboration between workers to build elaborate irrigation infrastructure, distribute water, dredge, and drain in rice growing areas, has strengthened the emergence of collective values; a basic engagement in order to survive. Over generations, the rice growing mode of production has resulted in collectivist trails with the associated obedience and insufficient self-expression, which in turn has deterred the formation of democratic institutions.

Based on a large cross-country dataset, our evidence suggests that contemporary democracy is negatively affected by a rice growing agricultural legacy. We found support for the hypothesis that countries that in the past have depended on rice cultures are more likely to have developed autocratic institutions than countries that have relied on other staple food cultures. These results were robust to the inclusion of confounding factors such as geographical characteristics, other main suitability crops, the main determinants of democracy highlighted in the literature, and religious adherence. Furthermore, the

results were robust to the sequential exclusion of continents and different timing and alternative measures of the outcome variable (democracy).

The finding that rice suitability has been an effective deterrent of the formation democratic institutions does not necessarily imply that alternative agricultural legacies have promoted democratization. In fact, we found that land suitability for other major cereals such as wheat, barley and maize have insignificant effects on current democracy. Following the lead of Acemoglu et al. (2001, 2005) and Woodberry (2012) the literature has found that the colonial past and trade have been influential for the formation of democratic institutions in former colonies. However, these results still do not answer the question as to why the western countries were forerunners for the development of democratic institutions. Furthermore, the oft-held view that autocratic institutions are more likely in countries close to the equator is not supported in our regressions with rice suitability included as a regressor, suggesting that autocracy is not an outcome of being close to equator but being located in a continent that is prone to autocratic institutions.

Chapter 1 Appendix

This appendix provides details and data sources of all variables employed by the empirical analyses in the current chapter.

1.A Additional analysis: Does rice suitability transmit to democracy through culture?

Instead of regressing democracy directly on rice suitability, we regress democracy on individualism using rice suitability as an instrument of individualism in this section in order to check the extent to which rice suitability affects democracy through culture. We use the individualism index of Hofstede (2010) for this purpose. Unfortunately, the limited geographical scope of the Hofstede index reduces the number of observations to 89, where many of the excluded observations included in our baseline sample has a high rice suitability index; thus, significantly reducing the efficiency of our IV regressions relative to that of the baseline regression.

The first- and second-stage regression results are reported in Table 1.10. The coefficient of rice suitability in the first stage is significantly negatively associated with Hofstede's index of individualism ($F = 10.5$), suggesting that rice suitability contributes to the formation of collectivism. The second-stage results in Column (1) of Table 1.10 shows a significantly positive relationship between individualism and democracy – a result that is consistent with the finding of Gorodnichenko and Roland (2015). From these results, it can be inferred that rice suitability, among other factors, has been a fundamental deterrent for the formation of democratic institutions as rice suitability promoted collectivist cultures.

Table 1.10: Regressing democracy on individualism using rice suitability as an instrument

	(1)	(2)
	2nd stage	1st stage
	Dep. Var. = Democ	Dep. Var. = Individualism
<i>Individualism</i>	1.106** (2.624)	
<i>Rice suitability</i>		-0.174*** (-3.247)
Latitude	-0.351 (-1.242)	0.873*** (3.446)
Landlocked	-0.144 (-1.143)	0.115 (1.62)
Distance to river or coast	0.23 (0.982)	-0.451*** (-4.253)
Elevation	-0.25 (-1.120)	0.411** (2.627)
Terrain ruggedness	0.08 (0.527)	-0.163* (-1.733)
Temperature	-0.006 (-0.027)	0.027 (0.102)
Precipitation	0.289*** (3.186)	0.185* (1.738)
Continent FE	Yes	Yes
R-squared	0.582	0.674
Observations	89	89

Notes: This table reports the standardized beta coefficients of regressions under different specifications. Robust standard errors are clustered at the continent level and t statistics are reported in parentheses. *, ** and *** denote significance at the 10%, 5% and 1% levels, respectively. The intercept estimates are not shown. The continent dummies are Asia, Europe, America, Oceania and Africa.

1.B Variable definitions and data sources

1.B.1 Main dependent variable

Our main dependent variable is the average scores of institutionalized democracy over the period of 1960 to 2013 from the Polity IV project (Marshall et al., 2014). The democracy indicator classifies regimes on an additive 11-point scale ranging from zero to ten, where a greater value reflects a higher degree of institutionalized democracy. Since the data are not available back to 1900 for some of the countries, their averages are based on the available data over the specified period.

1.B.2 Main explanatory variable

Our main explanatory variable is the country-level wetland rice suitability ratio, defined as the percentage of arable land that is suitable or very suitable for wetland rice cultivation in a given country. The measure is constructed using crop yield data from the Global Agro-ecological Zones database (GAEZ) jointly developed by the Food and Agriculture Organization under the United Nations and the International Institute for Applied System Analysis (FAO and IIASA, 2000).

The rice suitability ratio in this paper is constructed using land suitability data under rain-fed conditions with a mixed level of inputs. We choose the rain-fed conditions to ensure that the source of water supply is less likely to be affected by human intervention on that, for instance, efficacy of different irrigation methods. Additionally, the mixed level of inputs reflects the level of inputs and farming technologies deployed to best approximate the actual practice. Under the mixed level of inputs, the extent of land with cultivation potential for rain-fed crops is estimated by assuming that the most suitable land is deployed under high-level input farming, moderately suitable land under intermediate-level input farming and marginally suitable land under low-level input management. Thus, data on cultivation potential under mixed-level inputs reflect the land suitability for cultivation under the assumption of effective resource allocation.

1.B.3 Control variables

1.B.3.1 Geographic controls

Landlocked: A dummy variable indicating 1 if a country is fully enclosed by land and 0 otherwise, from the CIA world fact book.

Distance to coast or river: the average distance to the nearest coast or navigable river (10^3km), from the Geographically based economic data (G-ECON) project.

Terrain ruggedness: an index measures terrain irregularities of a country, from Nunn and Puga (2012).

Elevation: the average elevation of a country above sea level, from the G-ECON project.

Absolute latitude: the absolute latitude of a country, from La Porta et al. (1999).

Precipitation: the intertemporal average monthly precipitation of a country in mm over the period of year 1961–1990, from Ashraf and Galor (2013).

Mean temperature: the average monthly temperature of a country over the period of year 1961–1990, from the G-ECON project.

Notes: CIA: Central Intelligence Agency.

1.B.3.2 Historical controls

Timing of agricultural transition: the number of years (in thousands) elapsed in 2000 AD, since the transition to agriculture was estimated to occur, from Putterman and Trainor (2006).

Biogeography: the first principal component of the standardized numbers of domestic wild plants, from Hibbs and Olsson (2004) and Olsson and Hibbs Jr (2005).

Population density in 1500 AD: the estimated population density in 1500 AD, from Acemoglu et al. (2002).

Predicted genetic diversity: an index calculated using information on expected heterozygosity at population level and ancient migration distance from East Africa, from Ashraf and Galor (2013). The ancestry adjusted version is constructed using the migration matrix from Putterman and Weil (2010) to make it compatible for contemporary populations.

Pathogen stress: a standardized historical pathogen prevalence index, which covers nine diseases, from Schaller and Murray (2008).

Human settlement history: the historical duration of human settlement (in million years), from Ahlerup and Olsson (2012).

State history up to 1500 AD: an index covering the state history for the period of 1 AD to 1500 AD, from Putterman (2004). The ancestry-adjusted version is constructed using the ancestry matrix created by Putterman and Weil (2010).

Traditional local democratic practices: the weighted average of traditional local demo-

cratic practices at the country level, from Giuliano and Nunn (2013).

European language (% of population): the share of population that speaks European languages in a given country, from Easterly and Levine (2016).

Colony indicator: an indicator variable that equals to one if the country was a colony and zero if otherwise, from Olsson (2009).

Duration of colonization: the number of years under colonial rule, from Olsson (2009).

Indicator of colonial origin: a set of indicator variables capturing whether the colonizer was Spain, Britain, France, Portugal or other European countries, from Nunn and Puga (2012).

1.B.3.3 Contemporary controls

GDP per capita in 1960 (log): log of GDP per capita for the year 1960 converted to constant 2005 international dollar using PPP rates, from the World Development Indicators.

Urbanization: the percentage of total population living in urban areas as defined by national statistical offices for the year 1960, from the World Development Indicators.

Education: the average of years of schooling for the population aged 15 and above and that of the population aged 25 and above for the year 1960, from Barro and Lee (2013).

Oil rents (% of GDP): the difference between the value of crude oil production at regional prices and total costs of production as the percentage of GDP, from the World Development Indicators. We take the average of rents for the period of 1960 to 2013.

Mineral rents (% of GDP): the difference between of production for a stock of minerals at world prices and their total costs of production as the percentage of GDP, from the World Development Indicators. Minerals included in the calculation are tin, gold, lead, zinc, iron, copper, nickel, silver, bauxite, and phosphate. We take the average of rents for the period of 1960 to 2013.

Diamond production: an indicator variable that takes the value of one if there is production of diamond and zero if otherwise, from Gilmore et al. (2005).

Ethnic fractionalization: the probability that two randomly selected people come from different ethnic groups, from Alesina et al. (2003).

Language fractionalization: the probability that two randomly selected people speaks different ethnic languages, from Alesina et al. (2003).

Religion fractionalization: the probability that two randomly selected people have different religious beliefs, from Alesina et al. (2003).

Protestants (%) / Muslims (%) / Catholics (%) / Buddhists (%): the population share that follows the religion in a given country for the year 1970, from McCleary and Barro (2006).

1.B.3.4 Alternative measures of Democracy

Polity2 index: a composite index commonly used in quantitative studies for investigation of general regime effects, from the Polity IV project (Marshall et al., 2014).

Electoral democracy: an index that measures the quality of democracy by the extent of suffrage, cleanness and effectiveness of election process, freedom of expression and presence of an independent media, from the Varieties of Democracy (V-Dem) database (Coppedge et al., 2015).

Liberal democracy: an index that measures the quality of democracy by focusing on the protection of civil liberties and properties by constitutions, rule of law, executive constraints as well as the extent of electoral democracy, from the V-Dem database (Coppedge et al., 2015).

Deliberative democracy: an index that measures the quality of democracy by the quality of discussions at all levels that lead to political decisions, from the V-Dem database (Coppedge et al., 2015).

SVM democracy index: an index generated by support vector machine (SVM) algorithm using information from existing major democracy indices, from Gründler and Krieger (2016).

1.B.3.5 Other major crop suitability ratios

Wheat / barley / maize suitability ratio: the country-level suitability ratio of a given cereal crop is defined as the percentage of arable land that is suitable or very suitable for its cultivation in a given country, under rainfed conditions with mixed inputs. The measures are constructed using crop yield data from the GAEZ (FAO and IIASA, 2000).

1.B.3.6 Other controls

Soil suitability: a measure constructed based on information of soil carbon density and soil pH of an index of land suitability, from Ashraf and Galor (2013).

Precipitation variation: the standard deviation of precipitation of a country over the period of year 1961–1990, from the G-ECON project.

Plough use: a country-level, population-weighted measure of historical animal plow cultivation practice, from Alesina et al. (2013).

Irrigation potential: the percentage of arable land where irrigation can at least double the output of agricultural production, from Bentzen et al. (2017).

Other labor-intensive crops: the average of land suitability ratios for cultivation of potato, maize and bean, weighed by the estimated labor intensity required for each crop from Fouka and Schlaepfer (2017) and normalized by the sum of suitability ratios of all three crops. Data on land suitability ratios are from the GAEZ (FAO and IIASA, 2000).

Rice suitability ratio (with low inputs): the percentage of arable land that is suitable or very suitable for the cultivation of wetland rice in a given country, under rain-fed conditions with low inputs, from the GAEZ (FAO and IIASA, 2000).

Rice suitability ratio (with intermediate inputs): the percentage of arable land that is suitable or very suitable for the cultivation of wetland rice in a given country, under rain-fed conditions with intermediate inputs, from the GAEZ (FAO and IIASA, 2000).

1.B.4 Cultural dimensions

Individualism: a Hofstede index defined as “a preference for a loosely-knit social framework in which individuals are expected to take care of only themselves and their immediate families”, from Hofstede (2010) and the Hofstede Centre.

Power distance: a Hofstede index defined as “the degree to which the less powerful members of a society accept and expect that power is distributed unequally”, from Hofstede (2010) and the Hofstede Centre.

Masculinity: a Hofstede index defined as “a preference in society for achievement, heroism, assertiveness, and material rewards for success. Society at large is more competitive”, from Hofstede (2010) and the Hofstede Centre.

Uncertainty avoidance: a Hofstede index defined as “the degree to which the members of a society feel uncomfortable with uncertainty and ambiguity”, from Hofstede (2010) and the Hofstede Centre.

Long-term orientation / pragmatism: a Hofstede index that measures a society’s “links with its own past while dealing with the challenges of the present and the future”, from Hofstede (2010) and the Hofstede Centre.

Indulgence versus Restraint: a Hofstede index defined as “a society that allows rela-

tively free gratification of basic and natural human drives related to enjoying life and having fun”, Hofstede (2010) and the Hofstede Centre.

Chapter 2

Novelty-seeking Traits and Knowledge Intensity

2.1 Introduction

For modern economies, “technology is knowledge” (Mokyr, 2005b). Economic historian, Joel Mokyr (2005a) has long recognized the importance of knowledge for sustained technological progress and attributed the success of the Industrial Revolution to the accumulation of “useful knowledge” in society. What determines a country’s knowledge intensity is thus among the central questions addressed in the vast literature on economic development. As innovative knowledge is the engine for technological advancement, past studies generally focus on social and technological capabilities, institutional environment, cultural influence or other enablers of knowledge creation to account for differences of knowledge intensity across countries (see, e.g., Shane, 1992; Fagerberg et al., 2010; Bénabou et al., 2015b). Only a few studies have examined the potential effect of deep-rooted factors, such as the impact of inherent differences between populations on technological development (see, e.g., Spolaore and Wacziarg, 2009; Ashraf and Galor, 2013).

In this study, we focus on the role of novelty-seeking trait, which is a personality trait associated with the propensity of being excited in response to novel stimuli (Cloninger, 1994), in explaining technological deepening. Genetic studies suggest that novelty-seeking behaviour is linked to individual differences in certain neurotransmitter activities in the brain, specifically, the dopamine signalling pathways determined by genes such as Dopamine receptor D4 (DRD4) exon III gene (see, e.g., Benjamin et al., 1996; Ebstein et al., 1996). Additionally, studies find that while novelty-seeking individ-

uials generally are more exploratory and risk-taking, they are also more susceptible to psychological disorders, especially Attention Deficit Hyperactivity Disorder (ADHD) (see, e.g., Cloninger et al., 1994; Kuhnén and Chiao, 2009).

Building on the existing scientific findings, we investigate the impact of novelty-seeking traits on knowledge intensity and our main argument runs as follows. Novelty-seeking traits and the associated exploratory and risk-taking characteristics are closely linked to creativity and drive knowledge creation (Schweizer, 2006), thus contributing positively to technological progress. However, as novelty-seeking individuals are generally more impulsive and capricious in their behavior due to related disorders like ADHD, a high representation of these traits in the population tends to reduce cooperation, lower work efficiency and overall productivity in society. Therefore, high levels of these traits in the population can have countervailing effects on technological development. They can enhance knowledge intensity on one hand by expanding countries' innovation capability, while retarding technological development by reducing cooperation and productivity on the other. If we further assume diminishing marginal returns for both novelty-seeking traits and those without, i.e., the beneficial effects of novelty-seeking traits that are linked to promoting knowledge creation dominate at low levels of these traits while the deleterious ones prevail at high levels, our hypothesis would predict an inverted U-shaped relationship between novelty-seeking traits and knowledge intensity. There exists an optimal level of novelty-seeking traits in a society such that the level of knowledge intensity is maximized.

Using the Economic Complexity Index constructed by Hidalgo and Hausmann (2009) as the measure of knowledge intensity and the measure of novelty-seeking traits from Gören (2017), we establish an inverted U-shaped relationship between them. We further perform a battery of robustness checks to rule out the potential confounding effects of early development and contemporary indicators, some major genetic measures pertaining to both indigenous and contemporary populations, as well as different cultural dimensions on knowledge intensity. Compared to other measures of knowledge, ECI is a novel measure that estimates countries' productive capabilities by assessing the diversity and sophistication of their exported goods. Countries that can produce and export more products and more complex products are deemed to have higher levels of knowledge intensity. As a robustness check for our outcome measure, we also use various definitions of knowledge and technology measures to ascertain the effects of novelty-seeking traits on that. Additionally, we exploit the time dimension of

data on knowledge intensity to study the trend of shift of the estimated optimal level of novelty-seeking traits for knowledge intensity in recent decades.

Our study thus provides novel evidence on the underlying mechanism of how novelty-seeking traits affect economic development. A better understanding of how long-term development is linked to novelty-seeking traits is important because inherent differences between populations as such, are one of the most deep-rooted determinants in comparative economic development. Evidenced by our empirical findings, novelty-seeking traits exhibit persistent effects on knowledge intensity, which are not captured by other fundamental determinants such as geography, history, cultural influence or institutions.

This study is related to the strand of literature on determinants of knowledge and technological innovation. Fagerberg et al. (2010) conduct an extensive review on studies of innovation capabilities and national innovation system approach to answer how these enablers may influence innovation and technological progress. In addition, several works examine effects of national culture and social values on innovation outcomes. Using invention patents or trademarks per capita as proxies for innovation, Shane (1992, 1993) and Taylor and Wilson (2012) find a positive correlation between Hofstede's individualism index and national level innovation. Gorodnichenko and Roland (2017) further present, both theoretically and empirically, that more individualistic countries have more innovation and consequently higher long-run growth, and argue that it is primarily driven by social status rewards associated with innovation in such culture. Bénabou et al. (2015b) show that religiosity is associated with adverse attitudes towards science and technology, and thus is negatively correlated with technological innovation.

Our work contributes to a small but growing body of literature that explores the roles played by genetic and population variations in economic development. Galor and Moav (2002) hypothesize that individuals differ genetically in their preference for quality and quantity of children, and those with a higher valuation for offspring quality, i.e., traits that are complementary to the growth process, were selected for and progressively increased their share in the population during the Malthusian Epoch. This in turn contributed to the subsequent transition to industrialization. Galor and Michalopoulos (2012) focus on the selection of entrepreneurial traits in the course of development and argue that risk-tolerant, growth promoting traits would be selected for in initial stages of development, while risk-adverse traits would gain the evolutionary advantage during mature stages of development. Ashraf and Galor (2013) propose two

countering effects of genetic diversity on economic development. The benefit of genetic diversity is linked to promoting competition and driving technological advancement; however, high levels of genetic heterogeneity reduce trust and cooperation, and are thus disruptive to socio-economic order and aggregate productivity. This research is related to that of Gören (2017), which identifies a reduced-form association between novelty-seeking traits and comparative economic development measured using the natural logarithm of real GDP per capita in the year 2000. Differing from Gören (2017), we empirically test and establish technological progress as a causal mechanism of how these traits affect economic development.

The rest of the paper is organized as follows. Section 2.2 provides a detailed discussion on novelty-seeking traits and technological advancement, Section 2.3 puts forth a theoretical model illustrating the hypothesized inverted U-shaped relationship between novelty-seeking traits and knowledge intensity. Section 2.4 describes the empirical approach, data sources and variables. Section 2.5 presents the main empirical findings. Section 2.6 presents several robustness checks. Lastly, Section 2.7 concludes.

2.2 Conceptual framework

Novelty-seeking trait is a personality trait associated with the propensity of being excited in response to new stimuli (Cloninger, 1994). Previous works report that novelty-seeking individuals tend to engage in more exploratory and risk-taking activities. For example, Chen et al. (1999) and Matthews and Butler (2011) document the link between human ancient migration patterns and DRD4 gene variants associated with novelty-seeking, and posit that migratory societies select for genetic variants conferring exploratory propensity. Other studies find that individuals carrying novelty-seeking gene variants are more risk-taking in making financial decisions, compared with those without (Dreber et al., 2009, 2011; Kuhnen and Chiao, 2009). These findings suggest that novelty-seeking traits may be closely linked to economic activities such as innovation and entrepreneurial activities, which are beneficial to technological advancement.

However, it is worth noting that novelty-seeking individuals are also more susceptible to psychological disorders, especially Attention Deficit Hyperactivity Disorder (ADHD) (Cloninger et al., 1994). The DRD4 gene variants associated with novelty-seeking and ADHD probably were part of human adaptation that was selected for by ancient human migration processes (Chen et al., 1999). The main components of ADHD, for instance, hyper level of activity and attention diversion, no longer confer adaptive

advantages to the affected individuals in modern society. Rather, the impulsivity and capriciousness in their behavior are often perceived as disruptive and undesirable in the highly ordered contemporary societies (Williams and Taylor, 2006). Adults with ADHD persistently grapple with work-related problems. For instance, Coetzer and Trimble (2010) find that individuals with adult attention deficit experience greater difficulty in performing routine tasks and teamwork, and warrant more attention to cooperative conflict management. Other studies consistently document the association of ADHD with impaired work performance and increased absenteeism (Kessler et al., 2005; Matza et al., 2005; De Graaf et al., 2008; Kessler et al., 2009; Küpper et al., 2012; Halbesleben et al., 2013; Coetzer and Gibbison, 2016). Therefore, high levels of novelty-seeking traits, which lead to reduced cooperation and productivity, may impede technological development.

In view of the potentially positive and adverse effects of novelty-seeking traits on technological advancement, we hypothesize an inverted U-shaped relationship between the level of novelty-seeking traits in society and knowledge intensity. Assuming diminishing returns for both novelty-seeking traits and those without, at low levels, novelty-seeking traits are the limiting factor of countries' innovation capability and thus we expect a positive association between these traits and knowledge intensity. On the contrary, at relatively high levels of novelty-seeking traits, the costs of reduced cooperation and productivity due to ADHD and its co-morbidities become more prominent and outweigh the benefits of these traits, so knowledge intensity will be negatively associated with these traits. As both low and high levels of novelty-seeking traits are suboptimal, intermediate levels of such traits would be conducive for countries to reap the maximum benefits of these traits in knowledge creation, promote technological deepening and attain high levels of knowledge intensity.

In the next section, we provide a concise theoretical model illustrating the above-mentioned relationship. Our empirical findings strongly support this conjecture and remain robust when subjected to additional controls and further tests.

2.3 Theoretical framework

We use the following model to illustrate how variations in the levels of novelty-seeking traits could explain differences in knowledge intensity among countries. Consider an economy containing a mass of N individuals, whereby G of them carrying the gene variants associated with novelty-seeking behavior and the rest, $T = N - G$

are without. Further, we assume that the production of final goods requires inputs from both groups of workers, G , T as well as capital, K , and follows the Cobb-Douglas functional form:

$$Y = A(\mathbf{x}) G^\alpha T^\beta K^\delta \quad (2.1)$$

where $\mathbf{x}$ is a vector of cultural, institutional, human capital/education and all other factors that may influence the productivity, $A(\mathbf{x})$ captures these effects, elasticity parameters, $\alpha, \beta, \delta > 0$ and $\alpha + \beta + \delta = 1$ for constant returns to scale. The two groups of workers, G and T , differing in their personality outcomes and consequently work performance, are associated with different output elasticities, α and β , respectively.

Let τ denote the frequency of novelty-seeking gene in the economy, $\tau \in (0, 1)$, and correspondingly $G = \tau N$ and $T = (1 - \tau)N$. Rewrite the production function as:

$$Y = A(\mathbf{x}) \tau^\alpha (1 - \tau)^\beta [N^{\alpha+\beta} K^\delta] \quad (2.2)$$

Therefore, the level of technology, denoted as $a(\mathbf{x}, \tau)$, is a function of τ and $\mathbf{x}$:

$$a(\mathbf{x}, \tau) = A(\mathbf{x}) \tau^\alpha (1 - \tau)^\beta \quad (2.3)$$

Taking the first-order condition of $a(\mathbf{x}, \tau)$ with respect to τ :

$$\frac{\partial a(\mathbf{x}, \tau)}{\partial \tau} = a(\mathbf{x}, \tau) \left[\frac{\alpha - (\alpha + \beta) \tau}{\tau (1 - \tau)} \right] \quad (2.4)$$

Thus, Equation (2.3) predicts an inverted U-shaped relationship between the level of novelty-seeking traits and the level of technology as follows:

$$\frac{\partial a(\mathbf{x}, \tau)}{\partial \tau} = \begin{cases} > 0 & , \text{ if } \tau < \frac{\alpha}{\alpha + \beta} \\ < 0 & , \text{ if } \tau > \frac{\alpha}{\alpha + \beta} \end{cases} \quad (2.5)$$

As diminishing returns (α and β) are assumed for both novelty-seeking traits, τ , and those without, $(1 - \tau)$, beneficial effects of novelty-seeking traits that is linked to promoting knowledge creation will dominate at low levels of these traits whereas deleterious effects due to reduced cooperation and lower productivity would prevail at high levels. There exists an optimal level of gene frequency, $\tau = \frac{\alpha}{\alpha + \beta}$, where $\frac{\alpha}{\alpha + \beta} \in (0, 1)$,

such that it maximizes the level of technology.

Though we use the Cobb-Douglas form for simplicity in illustration, our result is applicable to more general functional forms with diminishing marginal returns for both novelty-seeking traits, τ , and those without, $(1 - \tau)$.

2.4 Empirical approach

2.4.1 Estimation methodology

Our main hypothesis predicts an inverted U-shaped relationship between novelty-seeking traits and knowledge intensity at the country level. We estimate the following regression model to evaluate the relationship:

$$Knowledge\ intensity_i = \alpha + \beta_1 DRD4_i + \beta_2 (DRD4_i)^2 + Controls_i' \gamma + \varepsilon_i \quad (2.6)$$

where $Knowledge\ intensity_i$ is the Economic Complexity Index for the year of 2015 developed by Hidalgo and Hausmann (2009) for country i , $DRD4_i$ and $(DRD4_i)^2$ is the combined DRD4 exon III 2R and 7R allele frequency at the country level from Gören (2017), our proxy for novelty-seeking traits and its quadratic term, $Controls_i$ is a vector of control variables and ε_i the error term. We use a set of geographical factors, which includes island dummy, landlocked dummy, distance to navigable river, terrain ruggedness, mean elevation, absolute latitude, soil suitability, precipitation and mean and variation of temperature as control variables in regressions. These are commonly used in empirical studies of growth. In addition, continental fixed effects are included to account for potential omitted variable biases of continent-specific characteristics. We expect β_1 to be positive and β_2 negative for the hypothesized inverted U-shaped relationship between the level of novelty-seeking traits and the knowledge intensity of a country, which correspond to the potential benefits and costs associated with the level of novelty-seeking traits in a society.

2.4.2 Data

2.4.2.1 Outcome measures

We use data from the Economic Complexity Index (ECI) for the year of 2015 developed by Hidalgo and Hausmann (2009) as the measure of knowledge intensity. Hidalgo and Hausmann (2009) contend that the economic sophistication of countries reflects the

division of labour and depends crucially on the “non-tradable inputs” required for production (Hausmann et al., 2014). They therefore attempt to assess countries’ knowledge intensity by estimating the amount of tacit knowledge contained in the exported goods. Emphasizing the importance of productive capabilities, i.e., the “modularized chunks of embedded knowledge” that connect a country and its exported goods, Hausmann et al. (2014) argue that production of sophisticated goods requires more productive capabilities; on the other hand, a country that can produce sophisticated goods is also able to produce different types of goods, simply by combining its productive capabilities in different ways. Therefore, ECI captures the impact of factors that determine the diversity and sophistication of the products in countries’ export baskets and predicts that countries with high levels of knowledge intensity can produce and export more products and more complex products. Such examples include Japan and Germany, which are capable of producing sophisticated goods like metalworking machine-tools, smart robots and medical imaging equipment; whereas countries ranked low in the ECI tend to export fewer products and those with low complexity, for instance, raw minerals, oil and agricultural products.

Unlike other common exogenous measures of knowledge intensity, such as percentage of employment in knowledge-intensive services or ratio of high-tech and medium-high-tech output to total manufactures output, ECI is an endogenously defined measure and therefore requires no *a priori* definition of knowledge intensity as the former do. Under this framework, a country’s knowledge intensity is estimated by averaging the knowledge intensity of all products it can export, and conversely, the knowledge intensity embedded in a product is estimated by averaging the knowledge intensity of countries that can export it. The definitions of knowledge intensity for countries and products are circular, and the model is solved by iterative methods. Technical details of deriving ECI are available in Appendix.

As a novel measure of knowledge intensity, ECI has gained increasing attention for its high predictive power on economic outcomes. Hausmann et al. (2014) argue that tacit knowledge, which is difficult to transfer, is the main constraint of economic growth. Compared with indicators on quality of governance and measures of human capital or competitiveness, ECI is shown to be a superior predictor of future economic growth, owing to its ability to reflect a country’s level of development. Hartmann et al. (2017) establish a robust and negative correlation between ECI and income inequality and suggest that a country’s knowledge intensity impacts on both the generation and

Figure 2.1: Spatial distribution of ECI scores for the year 2015

Notes: The ECI measures the relative levels of knowledge intensity of countries. Darker shades and larger sizes of the dots represent higher levels of ECI scores.

Source: The Atlas of Economic Complexity.

distribution of its income.

While ECI has proved to be a useful measure for knowledge intensity, one potential drawback of the current ECI measure is the omission of services sector from the estimation framework. This could result in an underestimation of overall knowledge intensity for service-based economies. However, unlike trade data on goods, quality data on services sector at disaggregated levels are hardly available. Ensuing studies that attempt to estimate economic complexity with inclusion of services data, can only conduct the analysis at the aggregated level (see, e.g., Stojkoski et al., 2016). This limitation renders their estimate of economic complexity less informative and consequently an inferior predictor for future economic growth, compared with the ECI constructed by Hidalgo and Hausmann (2009). Therefore, ECI is still preferred to other versions of economic complexity estimates for the purpose of measuring knowledge intensity. Figure 2.1 displays the scores of ECI (2015) for all countries in our sample.

2.4.2.2 Main explanatory variable

Our main explanatory variable uses the combined DRD4 exon III 2R and 7R allele frequency constructed by Gören (2017) as the novelty-seeking measure in a country. The DRD4 exon III gene, with variants containing different lengths of repeats ranging

from 2 to 11 (i.e., 2R - 11R), is often referred as the “adventure gene” for its association with human novelty-seeking traits. Among these alleles, the 2R and 7R variants share similar evolutionary and functional attributes, and primarily differ from the rest in their responses to dopamine in neurotransmission. It is posited that the blunted responses generated by the 2R and 7R variants need higher concentration of dopamine and this in turn leads to personality outcomes such as novelty-seeking or ADHD (Swanson et al., 2000; Wang et al., 2004). There is a large body of literature documenting the positive association of the 7R variant with such personality outcomes and increasing evidence showing similar association of the 2R variant with novelty-seeking traits or ADHD (Benjamin et al., 1996; Ebstein et al., 1996, 1997; Strobel et al., 1999; Tomitaka et al., 1999; Leung et al., 2005; Reist et al., 2007; Leung et al., 2017).

Based on the existing scientific evidence, Gören (2017) constructs a measure on the level of novelty-seeking traits in a society using the combined frequencies of DRD4 exon III 2R and 7R alleles at the country level. This is done by matching ethnicity data from Alesina et al. (2003) to DRD4 exon III genome data on the indigenous populations compiled by Gören (2016), to account for the multi-ethnic nature of contemporary countries. Figure 2.2 displays the combined DRD4 exon III 2R and 7R allele frequency for all countries in the sample.

Summary statistics of the variables are presented in Table 2.1. Our data cover a total of 114 countries across five continents.

Table 2.1: Summary Statistics

Variable	Observed	Mean	SD	Minimum	Maximum
ECI of year 2015	114	0.05	1.02	-2.19	2.47
DRD4 R2R7 allele frequency	114	0.24	0.07	0.10	0.46
Island (dummy)	114	0.08	0.27	0.00	1.00
Landlocked (dummy)	114	0.18	0.38	0.00	1.00
Distance to river (km)	114	5.91e+05	5.66e+05	3.06e+04	2.96e+06
Absolute latitude	114	0.33	0.19	0.01	0.71
Terrain ruggedness (Index)	114	1.19	0.95	0.02	4.76
Elevation	114	161.29	199.55	6.40	1096.50
Soil suitability	114	0.58	0.20	0.00	0.95
Precipitation	114	87.91	60.13	2.91	259.95
Temperature	114	17.10	8.36	-7.82	27.31
Temperature variation	114	5.08	3.44	0.43	14.72

Notes: Refer to the Appendix for descriptions of all variables.

Figure 2.2: Spatial distribution of DRD4 exon III 2R and 7R allele frequency

Notes: The DRD4 measure is the combined DRD4 exon III 2- and 7-repeat allele frequency at the country-level. Darker shades and larger sizes of the dots represent higher levels of the combined DRD4 2R7R allele frequency.

Source: Gören (2017).

2.5 Empirical results

2.5.1 Main regression results

Figure 2.3 displays the scatter plot and the fitted quadratic relationship for DRD4 exon III 2R and 7R allele frequency and ECI. This graphical presentation supports our hypothesis of an inverted U-shaped relationship between the two variables.

Results of the main regressions on the relationship between DRD4 exon III 2R and 7R allele frequency (hereinafter referred to as “DRD4 measure”) and ECI for the year of 2015 (hereinafter referred to as “ECI”) are presented in Table 2.2. We cluster the standard errors by continents when continental fixed effects are included, as countries within the same continent tend to share similar characteristics in development patterns and technological environment, accordingly, it is reasonable to assume that observations are independent across continents but not within. The baseline regression results also survive without clustering the standard errors by continents.

Additionally, we perform the inverted U-shaped tests on DRD4 measure to ascertain the hypothesized inverted U-shaped relationship, following the approach of Lind and Mehlum (2010). To reject the null hypothesis of a monotone or U-shaped relationship

Figure 2.3: The inverted U-shaped relationship between DRD4 allele frequency and ECI (unconditional)

Notes: This figure displays the association between DRD4 measure and ECI using an unconditional quadratic fit.

Table 2.2: Main regression Results

	(1)	(2)	(3)	(4)
Dep. Var. = ECI			Full specification	Full specification [Beta coefficient]
<i>DRD4</i>	1.171 (0.938)	13.904** (3.245)	28.356*** (11.438)	1.838 *** (11.438)
<i>DRD4 squared</i>		-24.004** (-3.375)	-48.566*** (-15.080)	-1.645 *** (-15.080)
Island			0.683* (2.551)	0.181 * (2.551)
Landlocked			0.009 (0.024)	0.003 (0.024)
Distance to river			0.000 (1.611)	0.148 (1.611)
Latitude			0.784 (0.734)	0.148 (0.734)
Terrain ruggedness			-0.039 (-1.063)	-0.036 (-1.063)
Elevation			-0.001* (-2.562)	-0.261* (-2.562)
Soil suitability			0.013 (0.099)	0.002 (0.099)
Precipitation			0.001 (0.263)	0.040 (0.263)
Temperature			-0.054** (-2.953)	-0.443** (-2.953)
Temperature variation			0.023 (0.212)	0.078 (0.212)
Continent FE	Yes	Yes	Yes	Yes
R-squared	0.457	0.476	0.608	0.608
Inverted U test	-	0.0173	0.000515	0.000515
Observations	114	114	114	114
Optimal DRD4 frequency	-	0.290	0.291	0.291

Notes: This table reports the correlation between DRD4 measure and ECI. Robust standard errors are clustered at the continent level and *t* statistics are reported in parentheses. *, ** and *** denote significance at the 10%, 5%, and 1% levels, respectively. The intercept estimates are not shown. The inverted U-shaped test is performed following the approach of Lind and Mehlum (2010) and the overall *p*-value is reported. The baseline controls used under the full specification include island dummy, landlocked dummy, distance to navigable river, absolute latitude, terrain ruggedness, mean elevation, soil suitability, precipitation and mean and variation of temperature. The continent dummies are Asia, Europe, America, Oceania and Africa.

and establish an inverted U-shaped relationship, it requires that the turning point lie within the observed data range, the coefficient of the quadratic term be negative and significant, and the slopes at both ends of the data range be statistically significant and have the expected sign. This ensures that the slope is increasing at a decreasing rate before the turning point and decreasing at an increasing rate after that.

Column (1) shows the bivariate association between ECI and DRD4 measure conditional on controlling for continental fixed effects. The influence of DRD4 measure on ECI appears to be positive but is not statistically significant. When the squared term of DRD4 measure is added in the regression in Column (2), both coefficients of DRD4 measure and its squared term are statistically significant at the 5% level, indicating a strong inverted U-shaped relationship between DRD4 measure and ECI as hypothesized. We further include various geographical controls related to location, topography, land productivity and climatic conditions into the regression in Column (3). The coefficients of DRD4 measure and its squared term remain statistically significant at the 1% level, supporting our main hypothesis of the inverted U-shaped relationship between DRD4 measure and ECI. The standardized beta coefficients for the full specifications are presented in Column (4) for the ease of interpretation.

In the baseline estimation, the optimal DRD4 frequency is estimated to be 0.291¹. ECI is positively associated with the DRD4 measure when the gene frequency is smaller than the turning point, showing that the benefits of novelty-seeking traits on knowledge creation dominate over the cost of ADHD at low levels of these traits. Above the optimal point, a higher level of gene frequency is associated with a lower ECI, indicating that the cost of ADHD associated with high levels of novelty-seeking traits outweigh the benefits. Figure 2.4 displays the estimated hump-shaped association between DRD4 measure and ECI under baseline specifications, using the augmented component-plus-residual plot².

¹The turning point is estimated by taking the first-order condition with respect to DRD4 in Equation 1.1.

²Additionally, we undertake continental Jack-Knife resampling regressions to rule out the possibility that our main results may be driven by inclusion of a particular continent (see Table 2.10 in the Appendix).

Figure 2.4: The estimated inverted U-shaped association between DRD4 measure and ECI (augmented component plus residual plot)

Notes: This figure display the association between DRD4 measure and ECI when fitted by a least-square quadratic estimator under the baseline specifications. The estimated ECI on the vertical axis is constructed using the augmented component, i.e., the fitted values of ECI predicted by DRD4 measure and its squared term, plus the residuals from the baseline model, following the approach of Ashraf and Galor (2013).

2.6 Robustness checks

2.6.1 Robustness to historical effects

The important role of early development in explaining current economic outcomes is often highlighted in the literature on comparative economic development (Nunn, 2009). Table 2.3 presents the regression results when considering these historical factors.

First, the timing of agricultural transition and biogeography (in Columns (1) and (2)) are closely related to agricultural productivity, which in turn influence the historical level of technology adoption and advancement (Olsson and Hibbs Jr, 2005; Putterman, 2008). Second, we include the natural logarithm of population density in 1500 AD, which is a basic measure for the level of prosperity of an economy in the Malthusian epoch, into the regression in Column (3). Furthermore, the average level of technology adoption in 1500 AD is controlled in the analysis in Column (4), as the persistent

Table 2.3: Controlling for historical effects

Dep. Var. = ECI	Historical effects (unadjusted)					
	(1)	(2)	(3)	(4)	(5)	(6)
<i>DRD4</i>	1.836*** (9.229)	1.792*** (12.539)	1.990*** (17.101)	1.894*** (11.113)	1.858*** (10.306)	2.338*** (9.141)
<i>DRD4 squared</i>	-1.643*** (-10.890)	-1.623*** (-26.476)	-1.769*** (-19.443)	-1.842*** (-6.292)	-1.669*** (-8.759)	-2.275*** (-6.541)
Timing of agricultural transition	-0.005 (-0.038)					-0.348* (-2.756)
Biogeography		0.062 (0.192)				0.131 (0.552)
Population density in 1500 AD (log)			0.191* (2.684)			0.209 (1.597)
Technology adoption in 1500 AD				0.192 (1.823)		-0.244 (-1.119)
State history up to 1500 AD					0.147 (2.095)	0.276** (3.585)
Baseline control	Yes	Yes	Yes	Yes	Yes	Yes
Continent FE	Yes	Yes	Yes	Yes	Yes	Yes
R-squared	0.608	0.603	0.625	0.673	0.616	0.709
Inverted U test	0.00084	0.000389	0.000564	0.0102	0.00273	0.00561
Observations	114	105	114	89	110	88

Dep. Var. = ECI	Ancestry-adjusted historical effects					
	(7)	(8)	(9)	(10)	(11)	(12)
<i>DRD4</i>	1.827*** (14.325)	1.838*** (11.067)	1.645*** (9.718)	1.741*** (10.388)	1.837*** (14.451)	1.611*** (13.913)
<i>DRD4 squared</i>	-1.631*** (-19.393)	-1.647*** (-15.016)	-1.422*** (-6.456)	-1.591*** (-8.836)	-1.634*** (-11.724)	-1.438*** (-8.463)
Timing of agricultural transition	0.044 (0.405)					-0.167 (-0.889)
Biogeography		-0.032 (-0.225)				-0.182 (-1.764)
Population density in 1500 AD (log)			0.285*** (5.729)			0.301*** (5.881)
Technology adoption in 1500 AD				0.225*** (8.816)		0.157** (3.155)
State history up to 1500 AD					0.140* (2.725)	-0.021 (-0.259)
Baseline control	Yes	Yes	Yes	Yes	Yes	Yes
Continent FE	Yes	Yes	Yes	Yes	Yes	Yes
R-squared	0.608	0.608	0.665	0.644	0.619	0.692
Inverted U test	0.000649	0.000418	0.0109	0.00257	0.00154	0.00409
Observations	114	114	114	114	114	114

Notes: This table reports the standardized beta coefficients of regressions under different specifications; robust standard errors are clustered at the continent level and *t* statistics are reported in parentheses. *, ** and *** denote significance at the 10%, 5%, and 1% levels, respectively. The intercept estimates are not shown. The inverted U-shaped test is performed following the approach of Lind and Mehlum (2010) and the overall *p*-value is reported. All regressions include the baseline controls consisting of island dummy, landlocked dummy, distance to navigable river, absolute latitude, terrain ruggedness, mean elevation, soil suitability, precipitation and mean and variation of temperature. The continent dummies are Asia, Europe, America, Oceania and Africa. Following the approach of Putterman and Weil (2010), ancestry-adjusted historical variables are constructed to capture the history of contemporary populations' ancestors for the analysis in Columns (7) to (12) in the lower panel.

influence of the historical level of technology adoption on comparative economic development has been documented by previous studies (Comin et al., 2010). Lastly, a measure for historical state development from 1 AD to 1500 AD, which captures information pertaining to the history of nationhood and state capacity and has been shown to be a key determinant of economic development in 1500 AD by Putterman (2008), is considered in the analysis in Column (5). The analysis in Column (6) includes all the above-mentioned historical factors in the regression.

Additionally, we report the regression results when all historical variables are ancestry-adjusted to reflect the history of contemporary populations' ancestors in Columns (7) to (12), as Putterman and Weil (2010) suggest that the history of a population's ancestors may have stronger predictive power for current economic outcomes, compared with that of their present residences. Country-level historical variables are pre-multiplied by the ancestry matrix created by Putterman and Weil (2010), which contains the proportion of people in each country in 2000 AD that was originated from various source countries in 1500 AD. The constructed variables thus represent the weighted average of its contemporary population's ancestral historical variables.

The results reported in Table 3 indicate that the inverted U-shaped relationship between DRD4 measure and ECI remains highly significant at the 1% level with inclusion of these early development indicators. Individually, for unadjusted historical effects, only the coefficient of the population density in 1500 AD appears to be significant and positively correlated with ECI, while for the ancestry-adjusted version, the population density in 1500 AD, the level of technology adoption in 1500 AD and early state institutions appear to be positively correlated with ECI; the rest are not statistically significant.

2.6.2 Robustness to effects of contemporary measures

This subsection reports regression results that consider the effects of various contemporary measures, which include income levels, education, research and development (R&D) expenditure, adoption of information and communications technology (ICT) and the quality of governance as control variables in Table 2.4.

First, the level of technology is closely related to the standards of living (Fagerberg et al., 2010), thus we include GDP per capita in the year 2000 (natural logarithm) as a control in our analysis (in Column (1)).

In addition, education (in the year 2000) influences the nation's human capital

Table 2.4: Controlling for effects of contemporary measures

Dep. Var. = ECI	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
<i>DRD4</i>	0.895*** (34.796)	1.321*** (10.713)	1.685** (4.059)	0.810** (2.847)	0.779** (3.025)	1.060*** (7.654)	1.316*** (9.268)	1.682*** (11.466)	1.434*** (9.540)	1.038** (3.438)
<i>DRD4 squared</i>	-0.731*** (-8.852)	-1.109*** (-13.593)	-1.529** (-3.039)	-0.716* (-2.557)	-0.647* (-2.609)	-0.911*** (-7.228)	-1.155*** (-10.995)	-1.474*** (-21.152)	-1.279*** (-12.249)	-0.914** (-4.124)
In GDP per capita in 2000	0.615*** (5.729)									
Years of schooling		0.385** (3.653)								
R&D expenditure (share of GDP)			0.477*** (9.147)							
Internet user				0.483** (3.723)						
Fixed telephone subscriptions					0.668*** (5.277)					
Mobile cellular subscriptions						0.512*** (5.528)				
Property rights protection							0.460*** (8.191)			
Executive constraints								0.364*** (8.644)		
Democracy									0.508*** (4.664)	
Corruption perceptions										0.384** (3.882)
Baseline control	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Continent FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
R-squared	0.757	0.681	0.772	0.707	0.756	0.728	0.748	0.668	0.706	0.755
Inverted U test	0.0297	0.000319	0.0585	0.0615	0.0785	0.0105	0.00126	0.000713	0.00164	0.0169
Observations	112	100	62	113	114	114	113	113	114	80

Notes: This table reports the standardized beta coefficients of regressions under different specifications; robust standard errors are clustered at the continent level and *t* statistics are reported in parentheses. *, **, and *** denote significance at the 10%, 5%, and 1% levels, respectively. The intercept estimates are not shown. The inverted U-shaped test is performed following the approach of Lind and Mehlum (2010) and the overall *p*-value is reported. All regressions include the baseline controls consisting of island dummy, landlocked dummy, distance to navigable river, absolute latitude, terrain ruggedness, mean elevation, soil suitability, precipitation and mean and variation of temperature. The continent dummies are Asia, Europe, America, Oceania and Africa.

accumulation while the R&D expenditure (in the year 2000) captures R&D efforts at the national level; these are deemed as important determinants of technological progress in endogenous growth models and thus are included in our analysis (in Columns (2) to (3)).

Furthermore, following the “capability” literature on innovation (Fagerberg et al., 2010), which suggests that ICTs play an important role in providing technological capabilities for innovation, we consider the utilization rate of several major ICTs in the year 2000 including the Internet, fixed line and mobile phone in our regressions (in Columns (4) to (6)).

Lastly, governance and institutions also serve as enabling factors that facilitate technological progress in society (Fagerberg et al., 2010). Indicators for quality of governance, which include property rights protection, executive constraints, adoption of

democratic system, and corruption perceptions, are controlled individually in Columns (7) to (10). The coefficients of DRD4 terms remain statistically significant in all cases.

2.6.3 Controlling for some population genetic measures

Genes and culture represent two streams of human inheritance in society. We next consider the possible influence of these deep-root factors on knowledge intensity in the following subsections.

Given that our proxy for novelty-seeking traits in society is DRD4 exon III 2R and 7R combined allele frequency, a genetic measure, we investigate potential confounding effects of several major genetic measures, namely the predicted genetic diversity (both the ancestry-adjusted and unadjusted versions), the current genetic distance to the U.S., the current blood distance to the U.K. and to the U.S., and the genetic distance to the U.K. in 1500 AD in Table 2.5.

The predicted genetic diversity is calculated using information on expected heterozygosity at population level and ancient migration distance from East Africa by Ashraf and Galor (2013). The ancestry-adjusted version makes it compatible with contemporary populations. This measure mainly captures information on overall intra-population genetic variations. Additionally, Ashraf and Galor (2013) show that the predicted genetic diversity (ancestry-adjusted) has an inverted U-shaped effect on long-run comparative economic development. We therefore consider both the linear and quadratic specifications of this measure in our analysis (in Columns (2) and (3)).

The genetic distance to the U.S. measures the genetic divergence between the country's population to that of the U.S., which is considered as the current technological frontier, a measure that has been shown to have certain predictive power in income difference among countries (Spolaore and Wacziarg, 2009). The Nei genetic distance measure assumes that genetic differences arise from mutation and genetic drifts, whereas the fixation index genetic distance (F_{ST}) measure assumes no mutation in populations. We consider both measures in our analysis (in Columns (4) and (5)). Additionally, F_{ST} is more than a measure of genetic divergence: it is directly linked to the variance in inter-population allele frequency, and contrariwise, to the degree of intra-population similarity (Holsinger and Weir, 2009).

Furthermore, we employ the current Mahalanobis distance between the frequency of blood types in a given country and that in the U.S. or the U.K. (denoted as "blood distance") as measures of genetic dissimilarity (in Columns (6) and (7)). The U.S.

Table 2.5: Controlling for some genetic measures

Dep. Var. = ECI	Baseline	Measures on contemporary population					
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
<i>DRD4</i>	1.838*** (11.438)	1.842*** (10.936)	1.737*** (7.233)	1.778*** (6.596)	1.759*** (6.206)	1.663*** (6.626)	1.644*** (6.471)
<i>DRD4 squared</i>	-1.645*** (-15.080)	-1.658*** (-13.495)	-1.555*** (-5.406)	-1.587*** (-7.726)	-1.570*** (-7.339)	-1.439*** (-5.151)	-1.431*** (-5.170)
Predicted genetic diversity (ancestry adjusted)		-0.032 (-0.240)	2.374 (0.441)				
Predicted genetic diversity squared (ancestry adjusted)			-2.416 (-0.455)				
F _{ST} genetic distance to the U.S.				-0.030 (-0.187)			
Nei genetic distance to the U.S.					0.001 (0.007)		
Blood distance to the U.K.						-0.207*** (-6.637)	
Blood distance to the U.S.							-0.154** (-4.056)
Baseline control	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Continent FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
R-squared	0.608	0.608	0.610	0.617	0.617	0.606	0.601
Inverted U test	0.000515	0.00138	0.0156	0.00208	0.00255	0.0167	0.0153
Observations	114	114	114	111	111	110	110
Dep. Var. = ECI	Measures on indigenous population						
	(8)	(9)	(10)	(11)			
<i>DRD4</i>	1.823*** (9.555)	1.710*** (10.072)	1.845*** (5.934)	1.854*** (5.159)			
<i>DRD4 squared</i>	-1.639*** (-14.663)	-1.551*** (-14.225)	-1.640*** (-7.183)	-1.647*** (-5.927)			
Predicted genetic diversity (unadjusted)	-0.096 (-0.332)	3.676 (0.946)					
Predicted genetic diversity squared (unadjusted)		-3.746 (-1.007)					
F _{ST} genetic distance to the U.K. (1500 AD)			0.051 (0.244)				
Nei genetic distance to the U.K. (1500 AD)				0.060 (0.264)			
Baseline control	Yes	Yes	Yes	Yes			
Continent FE	Yes	Yes	Yes	Yes			
R-squared	0.608	0.614	0.614	0.608			
Inverted U test	0.000796	0.000571	0.00301	0.00453			
Observations	114	114	112	111			

Notes: This table reports the standardized beta coefficients of regressions under different specifications; robust standard errors are clustered at the continent level and *t* statistics are reported in parentheses. *, ** and *** denote significance at the 10%, 5%, and 1% levels, respectively. The intercept estimates are not shown. The inverted U-shaped test is performed following the approach of Lind and Mehlum (2010) and the overall *p*-value is reported. All regressions include the baseline controls consisting of island dummy, landlocked dummy, distance to navigable river, absolute latitude, terrain ruggedness, mean elevation, soil suitability, precipitation and mean and variation of temperature. The continent dummies are Asia, Europe, America, Oceania and Africa. we control for the effect of genetic measures pertaining to contemporary populations in Columns (2) to (7) in the upper panel, and that of genetic measures pertaining to indigenous populations in Columns (8) to (11) in the lower panel.

and the U.K. are countries with the most individualistic culture in the world, and Gorodnichenko and Roland (2017) suggest that blood distances to these countries are negatively correlated to innovation.

Lastly, we control for the effect of genetic measures pertaining to indigenous populations, which are more exogenous in nature. These include the predicted genetic diversity (unadjusted) and both the F_{ST} and Nei genetic distance to the U.K. in 1500 AD in Columns (8) to (11).

Table 2.5 presents the regression results when these controls are considered. For current genetic measures, predicted genetic diversity and the genetic distance to the U.S. do not appear to have a significant effect on ECI (in Columns (2) to (5)), while the blood distance to the U.S. or the U.K. is negatively correlated to ECI, which is consistent with findings of Gorodnichenko and Roland (2017). All coefficients of genetic measures pertaining to indigenous populations are not statistically significant. Our main results are not undermined by the inclusion of these genetic measures. The estimates of DRD4 terms remain largely unaffected and statistically significant at the 1% level, indicating that our main analysis is unlikely to be confounded by the overall intra- and inter-population genetic variations, divergence or relatedness.

2.6.4 Controlling for some cultural effects

Culture is often believed to influence creativity, and consequently technological innovation, which is the product of creativity. We investigate effects of various cultural dimensions discussed by Hofstede (2010), which include individualism, power distance, masculinity, uncertainty avoidance, long-term orientation and indulgence on the knowledge intensity. In addition, we control for religiosity, which has been shown to link to adverse attitudes towards science and technology, and consequently impacts technological innovation negatively (Bénabou et al., 2015b).

Table 2.6 reports the regression results when controlling for different cultural dimensions respectively. Our results show that there is a positive correlation between individualism and ECI (in Column (2)) but a negative relationship between power distance and ECI (in Column (3)). This is consistent with findings in the current literature, which postulate that individualism promotes innovation by the social status rewards associated with innovation in such culture, and culture with low power distance also promotes innovation by granting individuals more autonomy and freedom (Shane, 1992, 1993; Gorodnichenko and Roland, 2011; Taylor and Wilson, 2012; Gorodnichenko and

Roland, 2017). The regression result also confirms that religiosity has a negative effect on ECI (in Column (8)), although its coefficient is less significant (at the 10% level), compared with that of individualism or power distance. Estimates of other cultural dimensions are not statistically significant (in Columns (4) to (7)). In all cases, estimates of DRD4 measure and its squared term remain statistically significant when different cultural influences are considered.

Table 2.6: Controlling for some cultural influences

Dep. Var. = ECI	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Baseline							
<i>DRD4</i>	1.838*** (11.438)	1.189** (2.791)	1.332** (2.970)	1.607** (2.899)	1.571** (3.198)	1.776** (4.096)	1.657** (3.009)	2.105*** (5.521)
<i>DRD4 squared</i>	-1.645*** (-15.080)	-0.993** (-2.828)	-1.087** (-2.917)	-1.391* (-2.754)	-1.318** (-3.140)	-1.404** (-3.772)	-1.402* (-2.290)	-1.855*** (-4.970)
Individualism		0.371*** (5.390)						
Power distance			-0.167** (-3.052)					
Masculinity				0.150 (1.469)				
Uncertainty avoidance					0.082 (0.590)			
Long-term orientation						0.250 (1.640)		
Indulgence							0.216 (1.173)	
Religiosity								-0.262* (-2.451)
Baseline control	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Continent FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
R-squared	0.608	0.712	0.688	0.690	0.679	0.651	0.672	0.644
Inverted U test	0.000515	0.0252	0.0266	0.0366	0.0222	0.0234	0.107	0.0414
Observations	114	84	84	84	84	83	78	69

Notes: This table reports the standardized beta coefficients of regressions under different specifications; robust standard errors are clustered at the continent level and *t* statistics are reported in parentheses. *, ** and *** denote significance at the 10%, 5%, and 1% levels, respectively. The intercept estimates are not shown. The inverted U-shaped test is performed following the approach of Lind and Mehlum (2010) and the overall *p*-value is reported. All regressions include the baseline controls consisting of island dummy, landlocked dummy, distance to navigable river, absolute latitude, terrain ruggedness, mean elevation, soil suitability, precipitation and mean and variation of temperature. The continent dummies are Asia, Europe, America, Oceania and Africa.

2.6.5 Alternative measures of knowledge and technology

To further ascertain the robustness of our results, we consider a number of alternative measures of knowledge and technology in this subsection. Table 2.7 reports the regression findings on different measures of knowledge and technology.

First, we consider two alternative measures of knowledge intensity in our analysis. Knowledge-intensive services play a vital role in innovation processes by serving as both sources of knowledge creation and facilitators of knowledge transfer in organizations and industries. Thus, it is an important indicator of knowledge production

Table 2.7: Alternative measures of knowledge and technology

	(1)	(2)	(3)	(4)
Dep. Var. =	Employment in knowledge-intensive services (% of workforce)	High-tech and medium-high-tech output (% of total output)	Scientific and technical journal articles per capita	R&D expenditure (% of GDP)
<i>DRD4</i>	1.156*** (8.913)	2.262*** (4.738)	1.677** (4.034)	2.202*** (4.983)
<i>DRD4 squared</i>	-1.223*** (-10.787)	-2.047*** (-4.688)	-1.543*** (-4.939)	-2.032*** (-5.694)
Baseline control	Yes	Yes	Yes	Yes
Continent FE	Yes	Yes	Yes	Yes
R-squared	0.739	0.370	0.634	0.426
Inverted U test	0.000804	0.00818	0.0114	0.0051
Observations	90	85	114	98
	(5)	(6)	(7)	(8)
Dep. Var. =	Global innovation index (2017)	TFP (2000)	Technology adoption (2000)	ISO9001 quality certificates
<i>DRD4</i>	1.844** (4.572)	1.075*** (5.439)	1.628** (2.901)	0.877*** (6.262)
<i>DRD4 squared</i>	-1.728*** (-4.856)	-1.134*** (-6.354)	-1.600** (-4.011)	-0.803** (-3.942)
Baseline control	Yes	Yes	Yes	Yes
Continent FE	Yes	Yes	Yes	Yes
R-squared	0.657	0.733	0.648	0.501
Inverted U test	0.00589	0.00501	0.0342	0.0386
Observations	102	78	104	102

Notes: This table reports the standardized beta coefficients of regressions under different specifications; robust standard errors are clustered at the continent level and *t* statistics are reported in parentheses. *, ** and *** denote significance at the 10%, 5%, and 1% levels, respectively. The intercept estimates are not shown. The inverted U-shaped test is performed following the approach of Lind and Mehlum (2010) and the overall *p*-value is reported. All regressions include the baseline controls consisting of island dummy, landlocked dummy, distance to navigable river, absolute latitude, terrain ruggedness, mean elevation, soil suitability, precipitation and mean and variation of temperature. The continent dummies are Asia, Europe, America, Oceania and Africa. In the upper panel, the dependent variables are employment in knowledge-intensive services (as percentage of workforce) for the year 2015 in Column (1), high-tech and medium-high-tech output as a percentage of total manufactures output for the year 2014, based on the Organization for Economic Cooperation and Development (OECD) classification of Technology Intensity Definition in Column (2), the average number of scientific and technical journal publications per capita during the years 2005 to 2015 in Column (3), and the average research and development expenditure (as percentage of GDP) during the years 2005 to 2015 in Column (4), respectively. In the lower panel, the dependent variables are Global innovation index (GII) scores for the year 2017 in Column (5), the estimated TFP relative to the U. S. for the year 2000 in Column (6), the average level of technology adoption in the year 2000 in Column (7), and the number of ISO 9001:2015 certificates for quality management systems issued (per billion PPP\$ GDP) in the year 2015 in Column (8), respectively.

and impact. The regression using the employment in knowledge-intensive services (as percentage of workforce) in 2015 as the dependent variable is presented in Column (1). High-tech and medium-high-tech industries are categorized based on their direct R&D intensity³, under the classification of Technology Intensity defined by the Organization for Economic Co-operation and Development (OECD). These industries often serve as a driver of productivity and regional economic development. We use the ratio of high-tech and medium-high-tech output to total manufactures output in 2014 as an indicator of technology intensity in Column (2).

Next, we deploy indicators on knowledge creation as alternative measures. Using data from the World Bank, we report the regression results where the dependent variables are the average number of scientific and technical journal publications per capita and the average research and development expenditure (as a percentage of GDP) during the years 2005 to 2015 (in Columns (3) to (4)), respectively. These are common measures of knowledge creation adopted in empirical studies. Scientific and technical journal publications capture information on quality of a country's science base and R&D expenditure reflects national efforts and resource allocation to innovation.

Additionally, we use data from the 2017 Global Innovation Index (GII) to verify our main results (in Column (5)). GII is an annual ranking on countries' innovation success and capacities published by Cornell University, INSEAD and the World Intellectual Property Organization. It uses both subjective and objective data to assess the innovation performance and contains indicators on different aspects of innovation covering political environment, education, infrastructure and knowledge creation.

Finally, we include several measures related to technological capacity in the robustness checks. In Columns (6) and (7), we consider Total Factor Productivity (TFP) relative to the U.S. in the year 2000 and a measure of the average level of technology adoption in each country for the year 2000 from Comin et al. (2010) as the dependent variable, respectively. In addition, we adopt the number of ISO 9001:2015 certificates for quality management systems issued (per billion PPP \$ GDP) in 2015 as an alternative measure in the analysis in Column (8). The ISO 9000 standards are a family of international standards that provide guidance and tools pertaining to quality management system for companies and organizations. Though it is procedural in nature, the ISO certification reflects an emphasis on high production quality by enterprises and thus serves as a

³R&D intensity is calculated using the ratio of R&D expenditures to gross output.

useful indicator of production capability.

In all the above cases, the inverted U-shaped relationship between DRD4 measure and the level of knowledge or technology remains robust to different definitions of knowledge and technology measures. Coefficients of DRD4 measure and its squared terms are at least statistically significant at the 5% level for alternative measures of knowledge and technology.

2.6.6 Robustness checks using the historical average of ECI data

In this section, we conduct robustness checks by deploying the historical data of ECI from the years 1966 to 2015 to test our main hypothesis. Table 2.8 reports the regression results when we use the historical average of ECI scores from the year 1966 to 2015 in columns (1) to (5) and the ten-year average of ECI scores in columns (6) to (10). Overall, DRD4 measure and its squared term remain statistically significant at the 5% level, regardless of the different lengths or periods of time specified for the average of ECI. This confirms that our main empirical results are unlikely to be driven by the choice of ECI of a specific period as the dependent variable. The effects of novelty-seeking traits on knowledge intensity are persistent over time.

Interestingly, the estimated optimal DRD4 frequency appears to increase gradually from 0.241 to 0.284 through the decades (see Columns (6) to (10)). A possible explanation is that globalization and economic growth have made societies more open and tolerant in general (Friedman, 2006). Though still highly orderly in nature, societies today are becoming more accepting of the seemingly disruptive and unpredictable behavior of novelty-seeking individuals, and thus are able to unleash more of their innovative potential and benefit from knowledge creation. More studies would be necessary to ascertain the underlying mechanisms.

2.6.7 Repeated cross-country analysis

Lastly, we conduct a repeated cross-country analysis using the yearly, the five-year average and the ten-year average of ECI data from the year 1966 to 2015, respectively. This approach allows us to have more observations in the regression analysis and Table 2.9 reports the results. The DRD4 measure and its squared term remain statistically significant at the 1% level, with or without the inclusion of geographical controls and continental fixed effects.

Table 2.8: Robustness check using historical average of ECI

	(1)	(2)	(3)	(4)	(5)
Dep. Var. = ECI	1966 - 2015	1976 - 2015	1986 - 2015	1996 - 2015	2006 - 2015
DRD4	1.388*** (8.809)	1.391*** (8.692)	1.345*** (7.987)	1.446*** (10.614)	1.724*** (13.822)
DRD4 squared	-1.330*** (-12.204)	-1.323*** (-11.524)	-1.269*** (-8.968)	-1.350*** (-11.139)	-1.587*** (-16.808)
Baseline control	Yes	Yes	Yes	Yes	Yes
Continent FE	Yes	Yes	Yes	Yes	Yes
R-squared	0.698	0.689	0.671	0.661	0.660
Inverted U test	0.00115	0.00119	0.00272	0.0023	0.000852
Observations	110	110	110	110	110
Optimal DRD4 frequency	0.273	0.275	0.277	0.280	0.284
	(6)	(7)	(8)	(9)	(10)
Dep. Var. = ECI	1966 - 1975	1976 - 1985	1986 - 1995	1996 - 2005	2006 - 2015
DRD4	0.693*** (7.966)	0.935** (3.746)	1.097** (3.468)	1.145*** (6.673)	1.724*** (13.822)
DRD4 squared	-0.776*** (-7.034)	-0.983*** (-5.549)	-1.062** (-4.203)	-1.089*** (-6.897)	-1.587*** (-16.808)
Baseline control	Yes	Yes	Yes	Yes	Yes
Continent FE	Yes	Yes	Yes	Yes	Yes
R-squared	0.729	0.724	0.659	0.650	0.660
Inverted U test	0.00378	0.0185	0.0192	0.00616	0.000852
Observations	91	91	108	110	110
Optimal DRD4 frequency	0.241	0.257	0.270	0.275	0.284

Notes: This table reports the standardized beta coefficients of regressions under different specifications; robust standard errors are clustered at the continent level and *t* statistics are reported in parentheses. *, ** and *** denote significance at the 10%, 5%, and 1% levels, respectively. The intercept estimates are not shown. The inverted U-shaped test is performed following the approach of Lind and Mehlum (2010) and the overall *p*-value is reported. All regressions include the baseline controls consisting of island dummy, landlocked dummy, distance to navigable river, absolute latitude, terrain ruggedness, mean elevation, soil suitability, precipitation and mean and variation of temperature. The continent dummies are Asia, Europe, America, Oceania and Africa. The dependent variables are the historical average of ECI scores from the year 1966 to 2015 in Columns (1) to (5) in the upper panel and the ten-year average of ECI scores from the year 1966 to 2015 in Columns (6) to (10) in the lower panel.

Table 2.9: Repeated cross-country results

	(1)	(2)	(3)	(4)	(5)	(6)
Dep. Var. =	ECI yearly	ECI yearly	ECI five-year average	ECI five-year average	ECI ten-year average	ECI ten-year average
<i>DRD4</i>	1.041*** (13.727)	1.135*** (9.616)	1.045*** (6.309)	1.152*** (10.196)	1.065*** (4.723)	1.186*** (9.186)
<i>DRD4 squared</i>	-1.122*** (-15.853)	-1.119*** (-16.294)	-1.122*** (-7.268)	-1.133*** (-16.438)	-1.139*** (-5.423)	-1.159*** (-13.668)
Baseline control	No	Yes	No	Yes	No	Yes
Continent FE	No	Yes	No	Yes	No	Yes
Time FE	Yes	Yes	Yes	Yes	Yes	Yes
R-squared	0.052	0.642	0.052	0.652	0.054	0.656
Inverted U test	0.00000	0.00108	0.000000	0.000891	0.00001	0.00113
Observations	4937	4937	1001	1001	510	510

Notes: This table reports the standardized beta coefficients of regressions under different specifications; robust standard errors are clustered at the continent level and *t* statistics are reported in parentheses. *, **, and *** denote significance at the 10%, 5%, and 1% levels, respectively. The intercept estimates are not shown. The inverted U-shaped test is performed following the approach of Lind and Mehlum (2010) and the overall *p*-value is reported. All regressions include the baseline controls consisting of island dummy, landlocked dummy, distance to navigable river, absolute latitude, terrain ruggedness, mean elevation, soil suitability, precipitation and mean and variation of temperature. The continent dummies are Asia, Europe, America, Oceania and Africa. The dependent variables are: the yearly ECI scores, the five-year average ECI scores and the ten-year average ECI scores for the period of year 1966-2015.

2.7 Conclusion

This paper attempts to explicate differences in the knowledge intensity across countries by investigating the potential influence of inherent variations among populations. Specifically, we focus on the level of novelty-seeking traits in a country and establish technological progress as a mechanism of how these traits affect economic development. Our empirical findings support the main hypothesis of an inverted U-shaped relationship between the level of novelty-seeking traits and knowledge intensity at the country level, confirming that such traits have both positive and negative effects on technological development. Countries with intermediate levels of novelty-seeking traits are in a better position to further knowledge creation and technological deepening, whereas both low and high levels of these traits are suboptimal for technological progress. It would be interesting to investigate the impact of novelty-seeking traits on knowledge intensity at the firm level to see how these traits affect entrepreneurial and innovation activities in organizations. Future research may also shed light on possible mechanisms of the gradual increase in the optimal DRD4 frequency for knowledge intensity in recent decades.

Chapter 2 Appendix

This appendix provides details and data sources of all variables employed by the empirical analyses in the current paper.

2.A Additional analysis: region exclusion test

We undertake continental Jack-Knife resampling regressions to rule out the possibility that our main results may be driven by inclusion of a particular continent. In Table 2.10, we sequentially exclude observations that belong to a continent. The coefficients of DRD4 and its squared term remain statistically significant in all restricted sample estimations, together with significant results from the inverted U-shaped tests, suggesting that our main findings are not driven by a particular continent.

Table 2.10: Sensitivity test to the exclusion of regions

	(1)	(2)	(3)	(4)	(5)	(6)
Dep. Var. = ECI	Full sample	Exclude Asia	Exclude America	Exclude Europe	Exclude Oceania	Exclude Africa
<i>DRD4</i>	1.838*** (11.438)	0.885* (2.753)	1.697*** (14.736)	2.318*** (35.879)	1.889*** (10.165)	1.949** (4.311)
<i>DRD4 squared</i>	-1.645*** (-15.080)	-0.813* (-2.979)	-1.549*** (-12.524)	-2.098*** (-9.590)	-1.721*** (-14.323)	-1.564** (-3.964)
Baseline control	Yes	Yes	Yes	Yes	Yes	Yes
Continent FE	Yes	Yes	Yes	Yes	Yes	Yes
R-squared	0.608	0.728	0.648	0.459	0.600	0.577
Inverted U test	0.000515	.0389	0.00394	0.0127	0.00181	0.0256
Observations	114	82	92	81	111	90

Notes: This table reports the standardized beta coefficients of regressions under different specifications; robust standard errors are clustered at the continent level and *t* statistics are reported in parentheses. *, ** and *** denote significance at the 1%, 5%, and 10% levels, respectively. The intercept estimates are not shown. The inverted U-shaped test is performed following the approach of Lind and Mehlum (2010) and the overall *p*-value is reported. All regressions include the baseline controls consisting of island dummy, landlocked dummy, distance to navigable river, absolute latitude, terrain ruggedness, mean elevation, soil suitability, precipitation and mean and variation of temperature. The continent dummies are Asia, Europe, America, Oceania and Africa.

2.B Variable definitions and data sources

2.B.1 Main dependent variable

Our main dependent variable is national knowledge intensity. We measure it using the Economic Complexity Index (ECI) for the year of 2015 developed by Hidalgo and Hausmann (2009). The following materials on the derivation of ECI are mainly based

on Hidalgo and Hausmann (2009), Hausmann et al. (2014) and Hartmann et al. (2017).

ECI is computed using information on the bipartite networks constructed from trade data, wherein countries and the products they export are connected. The bipartite networks can be described using a matrix, M_{cp} , where its element takes the value 1 if country c exports product p with revealed comparative advantage (RCA) greater or equal to one, and zero if otherwise. In addition, concepts of diversity and ubiquity are introduced to define the degree of nodes in the bipartite network, where diversity of a country is the number of different types of products in its export basket and ubiquity of a product the number of countries that can produce and export it. These terms are calculated as follows:

$$\text{Diversity of country } c = k_{c,0} = \sum_p M_{cp} \quad (2.7)$$

$$\text{Ubiquity of product } p = k_{p,0} = \sum_c M_{cp} \quad (2.8)$$

Using the information of diversity and ubiquity, coupled with the method of reflections, the knowledge intensity of country c (i.e., country complexity) for the $(N + 1)^{th}$ iteration can be obtained as follows:

$$k_c(N + 1) = \frac{1}{k_{c,0}} \sum_p M_{cp} k_p(N), \quad N \geq 0 \quad (2.9)$$

Similarly, the knowledge intensity of a product p (i.e., product complexity) for the $(N + 1)^{th}$ iteration is as follows:

$$k_p(N + 1) = \frac{1}{k_{p,0}} \sum_c M_{cp} k_c(N), \quad N \geq 0 \quad (2.10)$$

By repeated substitutions of k_c and k_p using Equations 2.9 and 2.10, the following expression that links the knowledge intensity of two countries, c and c' is obtained:

$$k_c(N) = \sum_{c'} \tilde{M}_{cc'} k_{c'}(N - 2) \quad (2.11)$$

where

$$\tilde{M}_{cc'} = \sum_p \frac{M_{cp} M_{c'p}}{k_{c,0} k_{p,0}}$$

The matrix $\tilde{M}_{cc'}$ connects countries exporting similar products while discounting

common products using the inverse of the ubiquity of a product and normalized by the diversity of a country. The solution of economic complexity is obtained from Equation 2.11 in the form of the eigenvector associated with $\tilde{M}_{cc'}$ when $k_c(N) = k_{c'}(N-2) = 1$. The eigenvector of $\tilde{M}_{cc'}$ associated with the largest eigenvalue is a vector of ones, which is not informative. Therefore, the eigenvector associated with the second largest eigenvalue, denoted as $\vec{K}$, which contains the largest amount of variance in the system is used instead. The standardized $\vec{K}$ thus constitutes the economic complexity measured:

$$ECI = \frac{\vec{K} - \langle \vec{K} \rangle}{stdev(\vec{K})} \quad (2.12)$$

We use ECI data from the year 2015 in the main analysis, which are obtained from the website of the Atlas of Economic Complexity hosted at Harvard University (<http://atlas.cid.harvard.edu/rankings/>). The historical ECI data from the year 1966 to 2015 are obtained from the website of the Observatory of Economic Complexity at MIT (<https://atlas.media.mit.edu/en/rankings/country/eci/>).

2.B.2 Main explanatory variable

Our main explanatory variable is the presence of novelty-seeking traits in a society. It is measured using the country-level DRD4 exon III 2- and 7-repeat combined allele frequency compiled by Gören (2017). Existing literature documents a substantial amount of evidence on the positive association of the 2R and 7R variants with novelty-seeking traits or ADHD (Benjamin et al., 1996; Ebstein et al., 1996, 1997; Strobel et al., 1999; Tomitaka et al., 1999; Leung et al., 2005; Reist et al., 2007; Leung et al., 2017). The novelty-seeking measure is constructed by matching ethnicity data from Alesina et al. (2003) to DRD4 exon III genome data on the indigenous populations from Gören (2016), and thus consists of ethnicity-weighted DRD4 exon III 2- and 7-repeat allele frequency at the country level.

2.B.3 Control variables

2.B.3.1 Geographic controls

Island: A dummy variable indicating 1 if a country is an island and 0 otherwise, from the CIA world fact book.

Landlocked: A dummy variable indicating 1 if a country is fully enclosed by land and 0 otherwise, from the CIA world fact book.

Distance to navigable river: the average distance to the nearest navigable river (km), from the Geographically based economic data (G-ECON) project.

Terrain ruggedness: an index measures terrain irregularities of a country, from Nunn and Puga (2012).

Elevation: the average elevation of a country above sea level, from the G-ECON project.

Absolute latitude: the absolute latitude of a country, from La Porta et al. (1999).

Soil suitability: a measure constructed based on information of soil carbon density and soil pH of an index of land suitability, from Ashraf and Galor (2013).

Precipitation: the intertemporal average monthly precipitation of a country in mm over the period of year 1961–1990, from Ashraf and Galor (2013).

Mean temperature: the average monthly temperature of a country over the period of year 1961–1990, from the G-ECON project.

Variation of temperature: the standard deviation of temperature of a country over the period of year 1961–1990, from the G-ECON project.

Notes: CIA: Central Intelligence Agency.

2.B.3.2 Historical controls

Timing of agricultural transition: the number of years (in thousands) elapsed in 2000 AD, since the transition to agriculture was estimated to occur, from Putterman and Trainor (2006).

Biogeography: the first principal component of the standardized numbers of domestic wild plants, from Hibbs and Olsson (2004); Olsson and Hibbs Jr (2005).

Population density in 1500 AD: the estimated population density in 1500 AD, from Acemoglu et al. (2002).

Technology adoption in 1500 AD: the average level of technology adoption in 1500 AD, from Comin et al. (2010).

State history up to 1500 AD: an index covering the state history for the period of 1 AD to 1500 AD, from Putterman (2004).

Notes: ancestry-adjusted historical variables are constructed using the global migration matrix from Putterman and Weil (2010).

2.B.3.3 Contemporary controls

GDP per capita in 2000 (log): log of GDP per capita for the year 2000 converted to constant 2005 international dollar using PPP rates, from World Development Indicators WDI (2012).

Years of Schooling: the average of years of schooling for the population aged 15 and above and that of the population aged 25 and above for the year 2000, from Barro and Lee (2013).

R&D expenditure: Research and development expenditure as percentage of GDP for the year 2000, from the World Bank.

Internet user: individuals using the Internet as percentage of population for the year 2000, from the World Bank.

Fixed telephone subscriptions: the number of fixed telephone subscriptions per 100 people for the year 2000, from the World Bank.

Mobile cellular subscriptions: the number of mobile cellular subscriptions per 100 people for the year 2000, from the World Bank.

Property rights protection: the property rights index for the year 2000, from the Quality of Government dataset of Teorell et al. (2015).

Executive constraints: the average scores of executive constraints from the Polity IV project for the period of year 1960-2009, and is rescaled to 0-1 from 1-7, from Gurr et al. (2010).

Democracy: the Voice and Accountability index developed as one of six Worldwide Governance Indicators ("WGI"), which measures the extent to which a country's citizens are able to participate in selecting their government, freedom of expression, freedom of association and a free media, from Kaufmann et al. (2011).

Corruption perceptions index: the corruption perceptions index for the year 2000, from the Quality of Government dataset of Teorell et al. (2015).

2.B.3.4 Genetic controls

Predicted genetic diversity: An index calculated using information on expected heterozygosity at population level and ancient migration distance from East Africa, from Ashraf and Galor (2013). The ancestry adjusted version is constructed using the migration matrix from Putterman and Weil (2010) to make it compatible for contemporary populations.

F_{ST} genetic distance: the relative F_{ST} genetic distance to the U.S. (weighted), from Spolaore and Wacziarg (2009).

Nei genetic distance: the relative Nei genetic distance to the U.S. (weighted), from Spolaore and Wacziarg (2009).

Blood distance to the U.K.: the Mahalanobis distance between the frequency of blood types in a given country and that in the U.K., from Gorodnichenko and Roland (2017).

Blood distance to the U.S.: the Mahalanobis distance between the frequency of blood types in a given country and that in the U.S., from Gorodnichenko and Roland (2017).

F_{ST} genetic distance to the U.K. (1500 AD): the relative F_{ST} genetic distance to the U.K. in 1500 AD, from Spolaore and Wacziarg (2009).

Nei genetic distance to the U.K. (1500 AD): the relative Nei genetic distance to the U.K. in 1500 AD, from Spolaore and Wacziarg (2009).

2.B.3.5 Cultural controls

Individualism: a Hofstede index defined as “a preference for a loosely-knit social framework in which individuals are expected to take care of only themselves and their immediate families”, from Hofstede (2010) and the Hofstede Centre.

Power distance: a Hofstede index defined as “the degree to which the less powerful members of a society accept and expect that power is distributed unequally”, from Hofstede (2010) and the Hofstede Centre.

Masculinity: a Hofstede index defined as “a preference in society for achievement, heroism, assertiveness, and material rewards for success. Society at large is more competitive”, from Hofstede (2010) and the Hofstede Centre.

Uncertainty avoidance: a Hofstede index defined as “the degree to which the members of a society feel uncomfortable with uncertainty and ambiguity”, from Hofstede (2010) and the Hofstede Centre.

Long-term orientation / pragmatism: a Hofstede index that measures a society’s “links with its own past while dealing with the challenges of the present and the future”, from Hofstede (2010) and the Hofstede Centre.

Indulgence versus Restraint: a Hofstede index defined as “a society that allows relatively free gratification of basic and natural human drives related to enjoying life and having fun”, Hofstede (2010) and the Hofstede Centre.

Religiosity: the first principal component of the religiosity variable constructed from the longitudinal data of WVS (year 1981 - 2014) based on five questions related to

religiosity on (1) whether you are a religious person; (2) the importance of religion in one's life; (3) belief in God; (4) the importance of God; and (5) church attendance, from the World Values Survey (WVS) Database (2015).

2.B.3.6 Alternative measures of knowledge and technology

Employment in knowledge-intensive services: the employment in knowledge-intensive services as percentage of workforce for the year 2015, from the Global Innovation Index website.

High-tech and medium-high-tech output: high-tech and medium-high-tech output as a percentage of total manufactures output, based on the Organisation for Economic Cooperation and Development (OECD) classification of Technology Intensity Definition for the year 2014, from the Global Innovation Index website.

Scientific and technical journal articles per capita: the average value of scientific and technical journal articles per capita from the year 2005-2015, from the World Bank.

R&D expenditure: the average value of research and development expenditure as percentage of GDP for the year of 2005-2015, from the World Bank.

Global Innovation Index (GII) for the year 2017: an index based on the annual ranking on countries' innovation success and capacities published by Cornell University, INSEAD and the World Intellectual Property Organization, from the Global Innovation Index website (<https://www.globalinnovationindex.org/>).

TFP in 2000 AD: the estimated total factor productivity relative to the U.S. for year 2000, from the United Nations Industrial Development Organization website: (<https://www.unido.org/data1/wpd/Index.cfm>).

Technology adoption in 2000 AD: the average adoption level of technology in each country for the year 2000, from Comin et al. (2010).

Number of ISO 9001:2015 certificates for quality management systems: the number of ISO 9001:2015 certificates for quality management systems issued (per billion PPP\$ GDP) for the year 2015, from the Global Innovation Index website.

Chapter 3

Religiosity and technology adoption

3.1 Introduction

Do religion and science/technology inevitably conflict? This question is of perennial intellectual interest to theologians, philosophers, scientists and historians of the field. From the anti-printing decree issued by the Muslim ruler in the sixteenth century to the current resistance and objection to vaccination or genetically modified crops by religious communities in the United States, religious practices, which serve as a type of informal institution (North, 1991), continue to affect individuals' economic attitudes and behaviors (Guiso et al., 2003) and consequently, the economic, political and social outcomes in contemporary society.

Religion, as suggested by Iannaccone (1998), is “a shared set of beliefs, activities and institutions premised upon faith in supernatural forces”. The origin of religion is often indistinguishable from that of mythology; the latter consisted of sacred tales, cultural beliefs and values specific to the tribal groups in ancient times. Its origin lies in meeting people's desires to understand and explain the world, which is probably one of the few commonalities that it shares with science. In this regard, Bertrand Russell (1930) once considers religion “as a disease born of fear and as a source of untold misery to the human race”. A simple yet gross distinction between religion and science is that religion deals with both the natural and the supernatural world while science only concerns with the former. In addition, while religious doctrines are taken as the truth, scientific hypotheses are subjected to empirical tests for verification (De Cruz, 2018).

Historically, religion and science/technology were often entangled in seemingly

irreconcilable conflicts¹. One of the most prominent cases of their interaction would be the dispute arisen from creationism and the Darwinian evolutionary theory. While creationism, which often is the fundamental part of doctrines in various religious denominations, asserts that the world and human beings are created by the (super)divine force, Charles Darwin's theory of evolution, published in 1895, states that all species arise and develop under the force of natural selection, i.e., survival of the fittest, and consequently challenges the corresponding religious account. Not surprisingly, the theory of evolution was initially greeted by religious communities with vehement opposition, repression and ridicule.

Nonetheless, with new evidence on the evolutionary theory being unveiled, the reception by various religious communities has gradually evolved over years. Additionally, religious authorities had since adopted a more open approach to science in general. For instance, Pope Pius XII took a neutral stance on research of human evolution (Pio et al., 1950). Pope John Paul II accepted the theory of evolution on the physical form of human beings but not for soul, which he deemed as a separate process of creation (Paul, 1996). Earlier this century, the Church of England even publicly endorsed the theory of evolution and issued statements to apologize to Charles Darwin for the initial rejection of the theory (Brown, 2008). In view of the development in religious authorities' reaction and attitudes towards scientific theories as such, a relevant question to ask would be whether religion still impacts on science/technology as before, and if so, to what extent the influence of religiosity on science/technology would be in today's context.

In contemporary societies, technology continues to be the centerpiece of economic development. An important aspect of technological progress that is worth examining is technology adoption, which is the decision (by an organization) to acquire and use an innovation. Comin et al. (2008) show that the differentials in technology adoption are able to explain the cross-country variations in total factor productivity and income levels. Similar findings documented by Comin and Hobijn (2010) indicate that differentials in the lags of technology adoption can account for a sizable portion of income differences across countries. Despite the significance of technology adoption in economic development, studies that attempt to identify potential deep-rooted determinants that could account for the differentials in current levels of technology adoption, are scanty (see,

¹Given the link and distinction between religion and science, scholars have been keen to examine and characterize their past interactions; there have been several typologies proposed by historians on the relationship between religion and science, see Section 3.A in the Appendix for more information.

e.g., Comin et al., 2010; Ang, 2015).

Against this background, this paper investigates the effects of religiosity on technology adoption in contemporary countries. It hypothesizes that heightened levels of religiosity in individuals tend to associate with more adverse attitude towards science and technology, as well as correlating with lower levels of technology adoption in society. Using data from the World Values Survey (1981 - 2014), we construct an aggregate measure of religiosity at the country level and estimate that an increase of one standard deviation in religiosity is associated with a decrease in the level of technology adoption by 56.4% standard deviations, holding other things constant. Additionally, we employ a measure on the historical level of pathogen prevalence as an instrumental variable (IV) for religiosity to establish the direction of causation from religiosity to technology adoption. Results of IV estimations are consistent with that of OLS estimations reported earlier, as coefficients of religiosity in all cases are negative and statistically significant at the 1% level. Using an alternative approach to address the endogeneity issue by deploying joint estimation of the religiosity measure and the structural error term with Gaussian copula correction does not change the main results and thus lends further support to our main hypothesis.

Furthermore, the results survive when controlling for potential confounding effects of culture, barriers to technology diffusion, social diversity and institutions on technology adoption in robustness tests. Further analyses confirm our hypothesis on the link between individuals' level of religiosity and their preference towards science and technology, which potentially could be the mediation channel as these would collectively influence a society's level of technology adoption. The findings also suggest that religiosity has similar effects on other aspects of technological progress at the country level, for instance, the level of knowledge intensity and total factor productivity.

Our work thus contributes to a vast body of literature that discusses the important and enduring influence of religion on technological progress and economic growth. A century ago, Max Weber (1930) hails the ethics of ascetic Protestantism as the source of capitalistic spirit, as certain aspects of the Protestant teachings helped generate economic attitudes that are conducive to capitalism. More recently, Guiso et al. (2003) suggest that religiosity foster economic attitudes such as trust that are conducive to economic growth. Barro and McCleary (2003) examine the influence of religious beliefs (in god or hell) on economic performance conditional on the level of church attendance, which is perceived as inputs of production of religious beliefs. Using Ramadan fasting

as a natural experiment, Campante and Yanagizawa-Drott (2015) conclude that religious practices may impact economic performance negatively through the channel of labor supply choices.

This study is related to Bénabou et al. (2015a,b), which document negative correlations between religiosity and individuals' attitude towards innovation in society, as well as between religiosity and a country's innovative capacity measured by (log) number of patents per capita. However, it differs from the literature in the following important ways.

First, unlike Bénabou et al. (2015a,b) who focus on attitudes toward science and technology as their economic outcome, this paper argues that religiosity is related to the extent to which one has adopted a new technology. Asking a respondent his view on science and technology is different from inquiring if he has adopted a new technology. For the purpose of this study, we argue that the latter is more relevant since attitudes will influence the decision on technology. Hence, by focusing on the actual action taken we can better identify the effect of religious beliefs and behaviors and thereby reduce the possibility of getting spurious correlations.

Second, despite the argument presented above, we cannot interpret the estimated effect of religiosity in a causal sense since the estimates may still be biased due to failure to account for some unobserved potential confounders and measurement errors in proxies for religiosity. To this end, drawing on the parasite-stress theory of values advanced by Thornhill and Fincher (2014), we provide fresh instrumented evidence using the extent to which a society is exposed to pathogen stress as an instrument for religiosity. In high pathogen environments, people are expected to practice ethical principles founded in their religious traditions whereas those under low pathogen stress should display less religious values. Hence, the intensity of pathogen exposure is expected to have a positive effect on the extent of religiosity. As an alternative approach, we use Gaussian copula correction for the joint estimation of religiosity and error terms to address the issue of potential endogeneity of religiosity measure. The main results survive in both approaches.

This research is also related to the strand of literature that examine the differentials in technology adoption or diffusion across countries. Comin et al. (2010) demonstrate the persistent influence of historical levels of technology adoption on the current practice of technology adoption. Using rice cultivation suitability index as the proxy for individualistic culture, Ang (2015) suggests that an individualistic culture tends

to foster adoption of new technologies in society. Relatedly, Geroski (2000) provides a detailed survey that attempts to explain the differentials in diffusion rates of new technology by focusing on models such as epidemic model and alternatives like probit model. Additionally, Fogli and Veldkamp (2012) discuss the impact of social networks on technology diffusion, by using germ stress as the IV for social networks, based on the premise that social networks evolved to fit both economic and epidemiological environment.

The rest of the paper is organized as follows. Section 3.2 describes the empirical approach, data sources and variables. Section 3.3 presents the main empirical findings on the influence of religiosity on technology adoption. Section 3.4 presents findings of sensitivity tests on unobservables and robustness checks when controlling for other cultural influences, potential barriers to technological diffusion and some contemporary measures. Section 3.5 provides further analyses by exploring the relationship between religiosity and individuals' attitude and preference towards science and technology, as well as examining the impact of religiosity on other aspects of technological progress. Lastly, Section 3.6 concludes.

3.2 Empirical approach

3.2.1 Estimation methodology

To evaluate the influence of religiosity on technology adoption in society, we estimate the following regression model:

$$Technology\ adoption_i = \alpha + \beta * Religiosity_i + Controls_i' * \gamma + \varepsilon_i \quad (3.1)$$

where $Technology\ adoption_i$ is the technology adoption index for country i for the year 2000 from Comin et al. (2010), $Religiosity_i$ is a measure of religiosity constructed using the World Values Survey data (all six waves spanning from year 1981 to 2014) by taking the average of all five measures of religiosity available in the survey to represent the overall level of religiosity for each country, $Controls_i$ is a vector of control variables and ε_i the error term. We use a set of geographical controls, which includes island dummy, landlocked dummy, distance to coast, mean elevation and terrain ruggedness in our regression analysis. These are commonly used in empirical studies of growth.

3.2.2 Identification strategies

A potential issue in estimating the influence of religiosity on technology adoption is that religiosity may be endogenous due to simultaneity or omitted variables. As suggested by the modernization theory (see, e.g., Marx, 1973; Inglehart and Baker, 2000), technological development may influence the evolution of cultural beliefs and values, such as religiosity, so causality could run in the reverse direction. Additionally, there may be unobserved factors impacting on technology adoption that are correlated with religiosity at the same time. To address this issue, we deploy two different approaches for identification.

3.2.2.1 Instrumental variable (IV) approach

First, we employ the instrumental variable (IV) approach by using a measure that is able to isolate the exogenous variation in religiosity with respect to technology adoption, i.e., a historical pathogen prevalence index. We draw on the literature of historical pathogen prevalence (see, e.g., Murray and Schaller, 2010; Fincher and Thornhill, 2012), which suggests that variations in the historical level of pathogen prevalence are able to explain observed cross-cultural variations among societies.

The rationale for using pathogen stress as an instrument for religiosity rests on the idea that a large exposure to parasite stress increases the degree of religiosity. In primitive societies plagued with the presence of life-threatening diseases, people believe that disease is a punishment from God and hence tend to behave in a moral way to avoid it. As a result, members of that society are more likely to conceive the existence of a High God who is actively supportive of conventional ethics to ensure the survival of that society. For the same reasoning, societies exposed to high levels of pathogen adversity are more inclined to construct impressive religious structures in order to facilitate the adoption of religious values.

On the contrary, the lower costs of infectious-disease contact in low parasite societies provide people with the flexibility to decide on their level of religious adherence and even adopt secular beliefs, and hence they tend to be less religious on average. Viewed in this light, the intensity of pathogen stress faced by a society should predict the importance of religious participation and commitment among the members of that society.

Additionally, the religious coping theory postulates that people resort to religions as “ways of understanding and dealing with negative life events”, such as in times of

illnesses (see, e.g., Park et al., 1990; Williams et al., 1991; Pargament, 2001). Taken together, heightened levels of religiosity tend to correlate with high historical levels of pathogen prevalence. This is supported by the first-stage F-statistics as reported in Section 3.3.2.

While the preceding discussion establishes the relevance of the IV, we next focus on the second criterion of the IV approach, i.e., exclusion restriction. Under the identifying restriction assumption, the historical level of pathogen stress does not affect the current level of technology adoption directly, other than through shaping religiosity as aforementioned, conditional on the controls included in the regressions; this exclusion restriction is thus an appropriate strategy for identifying the channel of causation.

Nonetheless, the requirement of exclusion restriction, i.e. perfect orthogonality between the instrument and the error, is impossible to validate empirically due to unobservability of the errors, and essentially rely on theoretical arguments (Wooldridge, 2010). In practice, such claims are often subject to debates. In our case, with alternative hypotheses on the channels of pathogen stress abound, for instance, pathogen stress may also work through social networks on technology diffusion as proposed by Fogli and Veldkamp (2012)², the plausibility of our theoretical argument is potentially compromised.

To address this issue, we perform the fractionally re-sampled Anderson–Rubin (FAR) test, which allows valid inference when instrumental variables do not perfectly satisfy the exclusion restriction, following the approach of Berkowitz et al. (2012) and Riquelme et al. (2013). The original Anderson–Rubin (AR) test, which is commonly used to draw inference in cases of weak instruments, requires instrument to be perfectly orthogonal to the structural error term. For the FAR test, the fractionally re-sampling technique based on the jackknife histogram estimator is applied to modify the AR test, to account for violations of the orthogonality condition. Thus, the FAR test is more conservative than the AR test. Our results survive the FAR test, i.e., when allowing for the further assumption that our IV does not perfectly satisfy the exclusion restriction (see Section 3.3.2).

The equation that specifies religiosity with respect to historical pathogen prevalence

²Note that the theoretical claim by Fogli and Veldkamp (2012) on exclusion restriction, is debatable as well, since it cannot be proven empirically.

for the IV approach is as follows:

$$Religiosity_i = \delta + \eta * pathogen_i + Controls_i' * \kappa + \xi_i \quad (3.2)$$

where $pathogen_i$ is the historical prevalence of pathogens for country i from Murray and Schaller (2010) and ξ_i is the residual.

3.2.2.2 Joint estimation with Gaussian copula correction

Second, we apply Gaussian copula correction to jointly estimate the potentially endogenous regressor (i.e., religiosity) and the error term, following the approach of Park and Gupta (2012). In this setup, the error term is assumed to have a normal marginal distribution, while that for the endogenous regressor is calculated using kernel density estimation, i.e., a non-parametric method. This allows more flexibility for specifying the marginal distribution of the regressor concerned, as determined by the data observed. From these one-dimensional marginal distributions, a Gaussian copula method is then deployed to derive the multivariate distribution that can effectively reflect the correlation between the endogenous regressor and the error, thus we can obtain consistent estimates for model parameters and overcome the endogeneity problem.

The model is estimated using maximum likelihood (ML). Park and Gupta (2012) show both theoretically and empirically that the estimators are consistent and asymptotically normal under common mild regularity conditions. For inference, we run 100 bootstraps to compute the standard errors. The bootstrap standard error is the sample standard deviation of the estimates obtained by the ML estimation. Given that inference occurs in two stages of this approach, i.e., the non-parametric estimation of the marginal distribution of the endogenous regressor in the first stage and the ML estimation of the endogenous regressor in the second stage, the usual standard errors calculated based on the information matrix and treat the generated regressors as the given observations, are incorrect in this case (Park and Gupta, 2012). Similarly, percentile confidence intervals derived from the sampling distribution using bootstrapping are adopted as the distribution of the bootstrapped parameters is skewed.

3.2.3 Data

3.2.3.1 Outcome measures

We use the technology adoption index for the year 2000 from Comin et al. (2010), which is constructed using data on technology adoption lag measured by Comin et al. (2008). Specifically, the adoption lag for a major technology is calculated as the number of years elapsed such that country i adopts the said technology with the same intensity in per capita terms as the contemporary technology leader, i.e., the United States does when the technology was available to production. The technologies used in the construction of adoption lag data can be broadly divided into five categories, namely, electricity production, information technologies, communication technologies, transportation technologies and agricultural technologies. Comin et al. (2010) further normalize the adoption lag data by using the number of years since the invention of the technology. The overall adoption index for country i is constructed as one less the average of adoption lags of all technologies for the country, whereby the adoption index for the United States is one by construction. Additionally, sector-level technology adoption indices, namely, on agriculture, communications, transport and industry sectors are available from Comin et al. (2010), and we include them in our analysis.

It is worth noting that the technology adoption index constructed by Comin et al. (2010) focuses on the intensive margin of technology adoption, i.e., on how prevalent the adoption is in a country, which is more relevant to our examination of cross-country variations in technology today. The extensive margin of adoption would be less pertinent to the discussion, owing to greater information exchanges and migrations, as well as decreased transportation costs that enable more rapid technological diffusion across contemporary countries.

3.2.3.2 Main explanatory variable

Our main explanatory variable is the country-level measure of religiosity constructed using data from the World Values Survey (WVS). The WVS is a global research project, which measures and analyzes people's beliefs and values, how these evolve over time and the political and social consequences these may bring about. Since 1981, the WVS survey has been conducted regularly in 98 countries worldwide. We use survey data pertaining to religiosity in our variable construction, from all available survey waves, i.e., for the years of 1981-1984, 1990-1994, 1995-1998, 1999-2004, 2005-2009 and 2010-2014.

Following the approach of Bénabou et al. (2015b), we construct five country-level measures of religiosity, denoted as “R1” to “R5”, respectively, based on the responses to the following survey questions: (1) whether the survey respondent is a religious person, how the respondent perceives (2) the importance of religion, (3) his / her belief in God, and (4) the importance of God, and lastly, (5) the respondent’s church attendance (see Appendix for details on the construction of these measures). Our summary index of religiosity, denoted as “Religiosity”, is constructed by taking the average of all five measures to represent the overall level of religiosity for each country.

3.2.3.3 Instrumental variable

We use a measure on historical prevalence of infectious diseases from Murray and Schaller (2010) as our instrument for religiosity in the IV estimation. This measure is a standardized index of historical pathogen prevalence constructed based on data from the early 1900s, and covers nine diseases including leishmania, schistosoma, trypanosoma, malaria, filaria, leprosy, dengue, typhus and tuberculosis for 230 regions / countries worldwide. Specifically, Murray and Schaller (2010) makes use of historical epidemiological atlases in the estimation of prevalence for each disease, on a four-point scale ranging from zero to three, with zero denotes that there is no case reported in the zone, one denotes that the disease is rarely reported, two indicates that the incidence is sporadically or moderately reported and three for disease that is present at severe or epidemic levels at least once.

Figure 3.1 displays the spatial distribution of religiosity and technology adoption index. Summary statistics of the main variables are presented in Table 3.1. Our data cover a total of 79 countries.

Figure 3.1: Spatial distribution maps

Notes: Maps display the spatial distribution of religiosity and technology adoption index, respectively. Darker shades correspond to larger values.

Source: Religiosity measure is constructed using data from the World Values Survey; the technology adoption index (year 2000) is from Comin et al. (2010).

Table 3.1: Summary statistics

Variable	Observed	Mean	SD	Minimum	Maximum
Technology adoption index	79	0.50	0.20	0.17	1.00
Agriculture sector index	66	0.43	0.23	0.18	1.07
Communications sector index	79	0.66	0.23	0.22	1.00
Transportation sector index	73	0.52	0.19	0.21	1.02
Industry sector index	67	0.54	0.25	0.06	1.03
Religiosity	79	0.70	0.19	0.21	0.95
Historical pathogen stress	78	.035	0.65	-1.31	1.16
Island (dummy)	79	0.06	0.25	0.00	1.00
landlocked (dummy)	79	0.18	0.38	0.00	1.00
Distance to coast (10^3 km)	79	0.37	0.41	0.00	2.21
Elevation	79	166.23	198.65	0.00	1096.50
Terrain ruggedness (index)	79	1.22	0.96	0.02	4.76

Notes: Refer to the Appendix for descriptions of all variables.

3.3 Empirical results

3.3.1 Main regression results

3.3.1.1 Religiosity and the overall level of technology adoption

Table 1.3 presents the ordinary least square (OLS) estimates for the hypothesized correlation between technology adoption and religiosity. Standardized beta coefficients are reported for the ease of interpretation, as all variables are normalized to have a mean of zero and a standard deviation of one. Geographical controls are included in all regressions.

In Columns (1) to (5), we first examine the correlation between the five country-level aggregate WVS survey responses, namely, “R1: religious person”, “R2: importance of religion”, “R3: belief in God”, “R4: importance of God” and “R5: church attendance”, with technology adoption. All measures are negatively correlated with the level of technology adoption. Except for R3, the rest are statistically significant at the 1% level. The significance of the coefficient of R3 is at the 10% level, which is weaker compared with other measures; the magnitude of the coefficient of interest is also slightly smaller than that of the rest, noting that in this case, the number of countries included in the regression is reduced to 68.

Lastly, Column (6) shows the association between technology adoption and our summary index, Religiosity. The influence of Religiosity on technology adoption remains negative and statistically significant at the 1% level, supporting our main hypothesis of the negative correlation between religiosity and technology adoption. The last specification in Column (6) is taken as the baseline estimation in subsequent analyses. Results of the baseline model suggest that an increase of one standard deviation ($SD = 0.19$) in the Religiosity leads to a decrease in the level of technology adoption by 0.564 standard deviations ($SD = 0.20$), holding all other variables constant. Therefore, variations in religiosity are able to explain a significant portion of differences in the level of technology adoption across countries. Figure 3.2 displays the added variable plot of the technology adoption index on religiosity under the baseline specification.

Table 3.2: Main regression results (OLS estimates)

	(1)	(2)	(3)	(4)	(5)	(6)
						Baseline
Dep. Var. = Technology adoption index						
R1: religious person	-0.471*** (-3.763)					
R2: importance of religion		-0.634*** (-6.339)				
R3: belief in God			-0.312* (-1.840)			
R4: importance of God				-0.583*** (-5.326)		
R5: church attendance					-0.606*** (-7.177)	
<i>Religiosity</i>						-0.564*** (-4.656)
Island	0.172* (1.791)	0.174** (2.473)	0.288** (2.540)	0.188** (2.589)	0.216*** (3.631)	0.187** (2.474)
Landlocked	-0.063 (-0.603)	-0.089 (-0.934)	0.036 (0.290)	-0.112 (-1.097)	-0.007 (-0.069)	-0.076 (-0.762)
Distance to coast	-0.283** (-2.415)	-0.383*** (-3.211)	-0.290** (-2.160)	-0.284** (-2.278)	-0.411*** (-3.474)	-0.328*** (-2.777)
Elevation	0.231 (1.435)	0.331** (2.537)	0.274 (1.485)	0.260* (1.849)	0.287** (1.994)	0.272* (1.870)
Terrain ruggedness	-0.025 (-0.197)	-0.014 (-0.134)	0.012 (0.086)	0.021 (0.189)	-0.006 (-0.059)	0.006 (0.057)
R-squared	0.336	0.499	0.228	0.456	0.472	0.432
Observations	78	78	68	79	78	79

Notes: This table reports the correlation between religiosity and technology adoption using the ordinary least square estimation. Standardized beta coefficients of regressions are presented, heteroskedasticity robust standard errors are used and *t* statistics are reported in parentheses. *, ** and *** denote significance at the 10%, 5% and 1% levels, respectively. The intercept estimates are not shown. The baseline controls used under the full specification include island dummy, landlocked dummy, distance to coast, mean elevation and terrain ruggedness.

Figure 3.2: Relationship between technology adoption and religiosity variation

Notes: This figure displays the association between technology adoption and religiosity while partialing out effects of other control variables under the baseline specification.

3.3.1.2 Religiosity and sectoral levels of technology adoption

As aforementioned, indices of technology adoption on sectoral levels for the year 2000 (four sectors) are also available from Comin et al. (2010). These allow us to examine whether religiosity would have heterogeneous effects on technology adoption in different sectors. Table 3.3 presents the OLS estimates on correlations of religiosity and technology adoption in agriculture, communications, transportation and industry sectors, respectively. For all sectors, coefficients of religiosity are negative and statistically significant at the 1% level, consistent with our main hypothesis. Religiosity appears to have a smaller effect on technology adoption in the transportation sector, with an estimated standardized beta coefficient of -0.354 compared with the effects on technology adoption in other sectors, whose estimates range from -0.536 to -0.611 . Figure 3.3 displays the added variable plots of the four sectoral indices of technology adoption on religiosity, using the baseline controls.

Table 3.3: Sector-based regression results (OLS estimates)

	(1)	(2)	(3)	(4)
Dep. Var. = Sector index of	Agriculture	Communications	Transportation	Industry
<i>Religiosity</i>	-0.536*** (-3.968)	-0.562*** (-4.751)	-0.354*** (-2.700)	-0.611*** (-4.655)
Baseline control	Yes	Yes	Yes	Yes
R-squared	0.326	0.428	0.285	0.415
Observations	66	79	73	67

Notes: This table reports the correlation between sector-based indices of technology adoption (2000) and religiosity. The sectors are agriculture, communications, transportation and industry, respectively. Standardized beta coefficients of regressions are presented, heteroskedasticity robust standard errors are used and *t* statistics are reported in parentheses. *, ** and *** denote significance at the 10%, 5% and 1% levels, respectively. The intercept estimates are not shown. The baseline controls include island dummy, landlocked dummy, distance to coast, mean elevation and terrain ruggedness.

Figure 3.3: Relationship between sector technology adoption and religiosity variation

Notes: This figure displays the associations between sector-based indices of technology adoption and religiosity while partialing out effects of other control variables under the baseline specification.

3.3.2 Results of instrumental variable estimations

While the OLS estimates reported in the preceding subsection indicate a negative and statistically significant correlation between religiosity and technology adoption, we further establish the causal relationship between them using the instrumental variable approach. We use the historical pathogen stress, denoted as “Historical Pathogen”, as the instrument to estimate effects of religiosity on technology adoption. Table 3.4 presents the regression results using the two-stage least squares (2SLS) estimations: panel A displays the second-stage regression results, panel B shows the first-stage results, and panel C reports results of diagnostic tests for all the religiosity measures.

First, in panel C, results of heteroskedasticity-robust endogeneity tests reject the null hypothesis that our religiosity measures are exogenous at the 1% significance level in all cases, thus justifying the IV approach. Second, the first-stage F-statistic, which is a criterion commonly used in practice to measure the strength of instruments (Staiger and Stock, 1994), are larger than the threshold value, i.e., ten, in most cases except for R1 and R3. In the case of R1, the F-statistic is 8.534, which is close to ten. For R3, the F-statistic is 2.932. For our summary index, “Religiosity”, the F-statistic is 21.819, thus indicating that historical pathogen stress is a relatively strong instrumental variable in this case. Third, we perform the FAR test as described in Section 3.2.2.1. Results in all regressions reject the null hypothesis that the coefficients of religiosity measures are zero, indicating that our results survive even when assuming that the exclusion restriction is not perfect for the current IV.

The first-stage estimates in panel B show that the historical level of pathogen stress is positively associated with all measures of religiosity at the 1% level of statistical significance except for R3, whose significance level is weaker, i.e., at the 10% level, thus supporting the proposition from the literature of historical pathogen stress.

Lastly, results of the second-stage estimations in panel A are consistent with that of OLS estimations reported earlier, as coefficients of religiosity in all cases are negative and statistically significant at the 1% level. Notably, the magnitudes of IV estimates of religiosity are consistently greater than the that of OLS estimates. This may either imply that the religiosity index in our case is measured with some errors since it is constructed based on WVS survey data *ex post*, or that religiosity influences technology adoption via some unobservable channels.

Additionally, Table 3.B1 in the Appendix reports the regression results on estimating the effects of religiosity technology adoption on the sectoral level using 2SLS, which are

also consistent with the OLS estimates presented in Table 3.3.

3.3.3 Results of joint estimation with Gaussian copula correction

In this section, we report the estimation results using the approach of Gaussian copula correction as described in Section 3.2.2.2.

For identification, the distribution of the endogenous regressor should neither be Bernoulli nor normal. This ensures that the distribution of the endogenous regressor contains adequate information and is different from that of the error term. Our religiosity measures are continuous variables and results of the Shapiro–Wilk normality test reject the null hypothesis that their distributions are normal³, thus satisfying the requirements of identification.

To implement the method, we follow the approach of Gui et al. (2019). First, we use the Epanechnikov kernel function to estimate the marginal distribution of the endogenous regressor. Next, Gaussian copula formula is applied to derive the joint distribution between the endogenous regressor and the error term. Third, we use maximum likelihood estimation by deploying the Nelder-Mead algorithm and using the OLS estimates as the start parameters. Lastly, we use bootstrapping to compute the standard errors (SE) and percentile confidence intervals (CI).

Table 3.5 reports the corresponding ML coefficient estimates for the six religiosity measures using this approach, the Bootstrap SE, and the Bootstrap percentile CIs (95%), conditional on controlling for the baseline specifications. All estimates for the six religiosity measures are negative, with none of the 95% percentile CIs including zero. This lends further support to our main hypothesis that religiosity has a negative effect on technology adoption, even after accounting for the potential correlations between our religiosity measures and the structural error term.

³*p*-values of the said tests for R1, R2, R3, R4, R5, and Religiosity are 0.0004588, 4.093e-05, 9.106e-10, 1.032e-06, 0.01074, and 2.33e-05, respectively.

Table 3.4: Regression results using 2SLS estimations

	(1)	(2)	(3)	(4)	(5)	(6)
<i>Panel A: Second-stage results</i>						
Dep. Var. = Technology adoption index						
R1: religious person	-2.070*** (-3.119)					
R2: importance of religion		-1.197*** (-6.202)				
R3: belief in God			-2.858* (-1.835)			
R4: importance of God				-1.309*** (-5.178)		
R5: church attendance					-1.164*** (-6.007)	
<i>Religiosity</i>						-1.350*** (-4.762)
Baseline control	Yes	Yes	Yes	Yes	Yes	Yes
<i>Panel B: First-stage results</i>						
Dep. Var. =	R1	R2	R3	R4	R5	<i>Religiosity</i>
<i>Historical Pathogen</i>	0.347*** (2.921)	0.600*** (6.315)	0.251* (1.712)	0.535*** (5.041)	0.617*** (6.407)	0.519*** (4.671)
Baseline control	Yes	Yes	Yes	Yes	Yes	Yes
Shea's partial R-squared	0.1342	0.3961	0.0626	0.3607	0.4085	0.2897
<i>Panel C: Diagnostic tests</i>						
1. Endogeneity test	0.0000	0.0042	0.0001	0.0024	0.0015	0.0028
Chi-square(1) <i>p</i> -value						
2. First-stage F-statistic	8.534	39.879	2.932	25.411	41.044	21.819
3. FAR test <i>p</i> -value	0.0019	0.0020	0.0089	0.0023	0.0017	0.0022
Observations	77	77	67	78	77	78

Notes: This table reports the correlation between religiosity and technology adoption, using pathogen stress as the instrumental variable for all religiosity measures. Both regression results of first stage and second stage estimations are presented. Standardized beta coefficients of regressions are presented, heteroskedasticity robust standard errors are used and *t* statistics are reported in parentheses. *, ** and *** denote significance at the 10%, 5% and 1% levels, respectively. The intercept estimates are not shown. The baseline controls used include island dummy, landlocked dummy, distance to coast, mean elevation and terrain ruggedness. P-values of heteroskedasticity-robust endogeneity tests, the first-stage F-statistics and p-values of FAR tests are reported for each regression.

Table 3.5: Results of joint estimation using Gaussian copula correction

Estimation	Regressor	Point Estimate	Bootstrap SE	Lower Bootstrap CI (95%)	Upper Bootstrap CI (95%)	Baseline controls	No. of bootstraps
(1)	R1	-0.6369535	0.2679791	-1.2901043	-0.2381514	Yes	100
(2)	R2	-0.9548986	0.1530569	-1.2181481	-0.5980848	Yes	100
(3)	R3	-0.43585964	0.2944751	-1.23009854	-0.07858956	Yes	100
(4)	R4	-0.9927139	0.1841274	-1.1819175	-0.5748748	Yes	100
(5)	R5	-0.8921740	0.2237675	-1.3954796	-0.5554665	Yes	100
(6)	<i>Religiosity</i>	-0.9324796	0.2334609	-1.2356567	-0.3506959	Yes	100

Notes: This table reports results of the joint estimation (MLE) using Gaussian copula correction. Standardized beta coefficients of regressions are presented. Bootstrap standard errors and bootstrap percentile confidence intervals (95%) are reported. The baseline controls used include island dummy, landlocked dummy, distance to coast, mean elevation and terrain ruggedness. AIC and value of log likelihood function are not shown.

3.3.4 Religiosity measures over different periods of time

In this subsection, we construct two measures of religiosity using WVS data from different periods of time, i.e., from year 1981 to 1998 (waves 1 to 3), denoted as “Religiosity 1” and from year 1999 to 2014 (waves 4 to 6), denoted as “Religiosity 2”, and examine their effects on technology adoption. Table 3.6 reports the results of both OLS and IV estimations. Coefficients of both religiosity measures are statistically significant, with the expected negative sign and similar magnitudes for the two periods, though the sample of “Religiosity 1” is reduced to only 45 countries. Similar to the main results, the magnitude of the IV estimates are also larger than that of the OLS. Our findings thus suggest that religiosity has a persistent effect on technology adoption during the period of our study.

Table 3.6: Regression results using religiosity measures from different periods of time

	(1)	(2)	(3)	(4)
Dep. Var. = Technology adoption index	OLS	2SLS (IV=Pathogen)	OLS	2SLS (IV=Pathogen)
<i>Religiosity 1</i> (WVS survey wave 1-3)	-0.453** (-2.611)	-1.781** (-2.201)		
<i>Religiosity 2</i> (WVS survey wave 4-6)			-0.593*** (-4.873)	-1.342*** (-4.674)
Baseline control	Yes	Yes	Yes	Yes
Observations	45	45	73	72

Notes: This table reports the correlation between religiosity and technology adoption using the ordinary least square and IV estimations. Two measures of religiosity were constructed using data from first three and last three survey waves of WVS, respectively. Standardized beta coefficients of regressions are presented, heteroskedasticity robust standard errors are used and *t* statistics are reported in parentheses. *, **, and *** denote significance at the 10%, 5% and 1% levels, respectively. The intercept estimates are not shown. The baseline controls used under the full specification include island dummy, landlocked dummy, distance to coast, mean elevation and terrain ruggedness.

3.4 Robustness to unobservables, culture, potential barriers to diffusion and some other contemporary measures

In this section, we perform several sensitivity tests to further establish the robustness of our main results. Table 3.7 presents the regression results when we control for potential confounding effects of some cultural measures, potential barriers to technology diffusion and other historical and contemporary measures.

3.4.1 Sensitivity to unobservables

We first examine the potential influence of unobservables on our estimation, using the approach of coefficient stability proposed by Altonji et al. (2005) and subsequently developed by Oster (2017). The coefficient of proportionality, denoted as δ , measures the relative ratio of the impact of unobservables to that of observables such that the unobservables would have an equally important impact as observables on the coefficient of the main explanatory variable. Specifically, under the assumption of proportional selection on observables, δ is calculated as follows:

$$\delta = \frac{(\beta^* - \tilde{\beta})(\tilde{R} - \hat{R})}{(R_{max} - \tilde{R})(\tilde{\beta} - \hat{\beta})} \quad (3.3)$$

where $\tilde{\beta}$ and $\tilde{R}$ are the estimated coefficient of interest and the R-squared value associated with the controlled regression respectively, $\hat{\beta}$ and $\hat{R}$ for the uncontrolled regression, and β^* and R_{max} for the hypothetical regression that includes both observed and unobserved controls. Oster (2017) derives a bounded value of $R_{max} = 1.3\tilde{R}$ for this method and argues that a value of δ greater than one (at $\beta^* = 0$) would indicate that results are robust to omitted variable bias. In our case, the estimated δ is 2.29, thus suggesting that the OLS effect in the baseline model is unlikely to be driven by the unobservables.

3.4.2 Religious denominations

First, we examine the influence of different religious denominations on technology adoption. In Column (1) of Table 3.7, we control for the population share of adherents to major religions, which include Protestantism, Islam, Catholicism and Buddhism from McCleary and Barro (2006)⁴. Interestingly, our results show that there is a positive

⁴Including the interaction terms of religiosity and each of religious denominations (after centering the variables) into the regression does not change the main results. However, as both religiosity and the

correlation between the population share of Protestants and technology adoption, while estimates of other religious influences are not statistically significant. This is consistent with findings in the current literature, which suggest that Protestantism increases the literacy rate (Becker and Woessmann, 2009), and in turn may promote the acceptance of new scientific findings and technologies. Additionally, Protestantism helps promulgate moral and cultural values such as individualism, which are conducive to innovation and adoption of new technologies (Ang, 2015).

3.4.3 Controlling for other cultural effects

The effects of culture on technological progress are often discussed in the literature (see, e.g., Mokyr, 2014). A relevant cultural dimension associated with this is individualism, which is defined as “a preference for a loosely-knit social framework in which individuals are expected to take care of only themselves and their immediate families” (Hofstede, 2010). Individualism has been shown to promote innovation by associating social status rewards with innovation in such culture (Shane, 1992, 1993; Gorodnichenko and Roland, 2011), as well as fostering adoption of new technologies (Ang, 2015). We control for the effect of individualism on technology adoption in Column (2). The measure on individualism is constructed based on data from worldwide surveys of IBM employees’ cultural values in 1960s and 2010 (Hofstede, 1984, 2010), which is commonly deployed in cross-cultural psychological studies. In this case, individualism is positively correlated with technology adoption at the 1% level of significance.

3.4.4 Considering social diversity

Next, we consider diversity measures including genetic diversity from Ashraf and Galar (2013), ethnic, linguistic and religious fractionalization from Alesina et al. (2003) in our analysis (in Columns (3) to (6)). These diversity measures are commonly controlled in research that attempts to explain cross-country differences in economic development (see, e.g., Brock and Durlauf, 2001). However, past studies have reported mixed results on the potential influences of population diversity on development outcomes. For instance, while Easterly and Levine (1997) suggest that ethnic and linguistic diversity is negatively correlated to per capita GDP growth, Alesina and Ferrara (2005) suggest that

population share of adherents to major religions are continuous variables, interpreting such interactions is associated with some difficulties (Jaccard et al., 1990; Aiken et al., 1991).

such effects of population diversity are conditional on other societal features. Our regression results show that linguistic fractionalization appears to be negatively correlated with Democ, while other measures are not statistically significant.

3.4.5 Historical level of technology adoption

Comin et al. (2010) suggest that the historical level of technology adoption has an enduring effect on technology adoption today, as the culture of technology adoption tends to evolve gradually. Therefore, we also control for the average level of technology adoption in 1500 AD. In this case, the historical level of technology adoption does not appear to have a significant effect on the current level of adoption.

3.4.6 Robustness to some contemporary measures

This subsection reports regression results that consider the effects of some contemporary measures, which include income levels, education and trade-to-GDP ratio as control variables. Both income (as measured in GDP per capita in the natural logarithm) and education level (as measured in percentage of population with tertiary education) for the year 2000 are positively correlated to technology adoption, consistent with the hypothesis of the modernization theory. Trade-to-GDP ratio is an indicator of the relative importance of trade in a country's economy and thus also reflects the openness of a country, as well as the degree of globalization and information exchange. In this case, trade-to-GDP ratio does not appear to have a significant effect on technology adoption.

3.4.7 Legal system and intellectual property rights protection

Lastly, we consider the potential effects of different origins of legal systems on technology adoption, by using indicator variables of Common law, French civil law, German civil law, Scandinavian law and Socialist law (as the base group) in Column (11). La Porta et al. (2008) find that legal origins can explain a significant portion of differences in systematic legal practices among countries, which in turn matter for economic outcomes. In particular, we also control for the degree of intellectual property (IP) rights protection in Column (12). In this case, our regression results suggest that different legal origins matter for technology adoption, and IP rights protection is positively correlated with technology adoption.

The hypothesized negative relationship between religiosity and technology adoption remains robust in all the above cases. Our main findings are not undermined by

including these controls in the analysis.

3.5 Further analyses

3.5.1 Religiosity and attitude towards science and technology

To better understand the impact of religiosity on technology and the potential channel for such influence, we take a closer look into the influence of religiosity on individuals' attitude and preference towards science and technology by examining responses to the relevant WVS survey questions, as presented in Table 3.8. These questions typically ask whether respondents agree or disagree with views on how science and / or technology influence the world or their daily life. For instance, the first two questions ask whether science and technology change life too fast, or make the world better. Questions (3) and (4) ask respondents' attitude towards religious faith vs. science, i.e., whether we depend too much on science than religious faith and which is right when science and religion are in conflict. Lastly, Question (5) asks respondents to rank the importance of knowing science in daily life and Question (6) asks respondents to assess the direct consequences of science on formation of values and perspective, i.e., whether it is bad that science breaks down ideas of right and wrong. For all outcome variables, we re-scale the responses, i.e., the degree of agreement or disagreement, such that the values fall within the range of 0 to 1 for the ease of comparison. The measure of religiosity at the individual level is constructed in a similar manner as the country-level aggregate measure, using the same set of survey questions in WVS. The control variables include respondents' age, marital status, gender, education and income levels, as well as district-level geographical controls and survey wave dummy variable.

The regression analysis of the above attitudinal variables suggests that religiosity is negatively correlated with positive view and attitude towards science and technology. Individuals who are more religious tend to dislike the changes brought by science and technology, have a propensity for supporting faith when science and religious faith are in conflict, and more likely to devalue the importance of science in daily life. A recent study by McPhetres and Zuckerman (2018) using survey data on contemporary American communities also suggests that higher religiosity is associated with negative attitudes towards science and lower levels of science literacy. Thus, the adverse attitude towards science and technology could be a potential mediator of the observed negative correlation between religiosity and the level of technology adoption at the country

Table 3.7: Robustness checks (OLS estimates)

	(1)	(2)	(3)	(4)	(5)	(6)
Dep. Var. = Technology adoption index						
<i>Religiosity</i>	-0.488*** (-3.819)	-0.244** (-2.272)	-0.563*** (-4.751)	-0.530*** (-4.182)	-0.510*** (-4.025)	-0.528*** (-3.858)
Protestants (%)	0.383*** (4.355)					
Muslims (%)	0.033 (0.239)					
Catholics (%)	0.197 (1.571)					
Buddhists (%)	-0.118 (-1.231)					
Individualism		0.669*** (6.148)				
Predicted genetic diversity			0.101 (1.369)			
Religion fractionalization				0.117 (1.122)		
Language fractionalization					-0.158** (-2.105)	
Ethnic fractionalization						-0.061 (-0.696)
Baseline control	Yes	Yes	Yes	Yes	Yes	Yes
R-squared	0.600	0.711	0.440	0.443	0.448	0.432
Observations	78	66	79	79	78	78
	(7)	(8)	(9)	(10)	(11)	(12)
Dep. Var. = Technology adoption index						
<i>Religiosity</i>	-0.628*** (-4.542)	-0.111* (-1.871)	-0.366*** (-3.297)	-0.558*** (-4.503)	-0.554*** (-4.308)	-0.203* (-1.830)
Technology adoption (1500 AD)	-0.106 (-1.074)					
In GDP per capita		0.842*** (11.176)				
Tertiary education			0.569*** (3.715)			
Trade-to-GDP ratio				-0.001 (-0.011)		
Common law					0.322** (2.380)	
French civil law					0.179 (1.540)	
German civil law					0.225** (2.312)	
Scandinavian law					0.275*** (4.790)	
IP rights protection						0.710*** (7.017)
Baseline control	Yes	Yes	Yes	Yes	Yes	Yes
R-squared	0.472	0.843	0.605	0.427	0.548	0.741
Observations	65	78	71	77	79	76

Notes: This table reports the correlation between religiosity and technology adoption with additional control variables using OLS estimations. Standardized beta coefficients of regressions are presented, heteroskedasticity robust standard errors are used and *t* statistics are reported in parentheses. *, ** and *** denote significance at the 10%, 5% and 1% levels, respectively. The intercept estimates are not shown. The baseline controls used include island dummy, landlocked dummy, distance to coast, mean elevation and terrain ruggedness.

level. Further analysis could be done to disentangle the relationship between religiosity, attitude towards science, science literacy and the level of technology adoption when data are available.

Table 3.8: Attitude and preference towards science and technology (individual-level analysis)

Dep. Var. =	(1)	(2)	(3)	(4)	(5)	(6)
	Science & technology change life too fast: disagree	Science & technology make the world better off: agree	Depend too much on science than faith: disagree	Religion is always right when in conflict with science: disagree	Not important to know science in daily life: disagree	Science is bad as it breaks down ideas of right and wrong: disagree
<i>Religiosity</i>	-0.118*** (-6.894)	-0.024** (-2.247)	-0.128*** (-8.909)	-0.531*** (-40.519)	-0.055*** (-4.713)	-0.142*** (-12.409)
Individual controls	Yes	Yes	Yes	Yes	Yes	Yes
Geographical control	Yes	Yes	Yes	Yes	Yes	Yes
Survey wave dummy	Yes	Yes	Yes	Yes	Yes	Yes
Cluster s.e. (district)	Yes	Yes	Yes	Yes	Yes	Yes
R-squared	0.020	0.046	0.031	0.329	0.032	0.049
No. of districts	537	943	943	703	702	702
No. of countries	40	67	67	52	52	52
No. of observations	45091	108148	105329	58304	61859	59764

Notes: This table reports the correlation between attitude / preference towards science and technology and religiosity using OLS estimations. Standardized beta coefficients of regressions are presented, heteroskedasticity robust standard errors are used and *t* statistics are reported in parentheses. *, ** and *** denote significance at the 10%, 5% and 1% levels, respectively. The intercept estimates are not shown. The individual characteristics controls include age, marital status, gender, education level and income. The geographical controls used are district-level latitude, distance to coast, elevation and terrain roughness. Robust standard errors are clustered at the district level.

3.5.2 Religiosity and other measures of technology

Furthermore, we examine the relationship between country-level religiosity and other measures of technology. Though these measures do not pertain to the level of technology adoption / diffusion, analysis conducted as such helps us gain further insights on how religiosity may impact on other aspects of technological progress. If the results were to be consistent with our main hypothesis, we would observe religiosity to have a negative impact on other aspects of technological progress as well. Table 3.9 presents the regression results.

First, we look at the impact of religiosity on knowledge intensity, which is measured using the Economic Complexity Index (ECI) developed by Hidalgo and Hausmann (2009) in Column (1). ECI is a novel measure that estimates countries' productive capabilities by assessing the diversity and sophistication of their exported goods. Countries that can produce and export more products and more complex products are deemed

to have higher levels of knowledge intensity. Not surprisingly, there is a strong and negative correlation between religiosity and a country's knowledge intensity.

Next, we examine the effect of religiosity on total factor productivity (TFP) relative to the U.S. in the year 2000 in Column (2). In addition, we adopt the number of ISO 9001:2015 certificates for quality management systems issued (per billion PPP \$ GDP) in 2015 as an alternative measure in the analysis in Column (3). The ISO 9000 standards are a family of international standards that provide guidance and tools pertaining to quality management system for companies and organizations. Though it is procedural in nature, the ISO certification reflects an emphasis on high production quality by enterprises and thus serves as a useful indicator of production capability. Religiosity is consistently found to be negatively correlated with these measures of technology.

Lastly, we consider several indicators of knowledge creation in our analysis. In Column (4), we deploy R&D expenditure at the national level, which reflects national efforts and resource allocation to innovation, as our dependent variable. Using data from the World Bank, we regress the average research and development expenditure (as a percentage of GDP) during the years 2005 to 2015 on religiosity. In Column (5), we use the employment in knowledge-intensive services (as percentage of workforce) in 2015 as the dependent variable. Knowledge-intensive services play a vital role in innovation processes by serving as both sources of knowledge creation and facilitators of knowledge transfer in organizations and industries. Thus, it is an important indicator of knowledge production and impact. In Column (6), we use data from the 2017 Global Innovation Index (GII) to measure the impact of religiosity on innovation. GII is an annual ranking on countries' innovation success and capacities published by Cornell University, INSEAD and the World Intellectual Property Organization. It uses both subjective and objective data to assess the innovation performance and contains indicators on different aspects of innovation covering political environment, education, infrastructure and knowledge creation. Likewise, the results suggest that religiosity exhibits a negative and statistically significant impact on knowledge creation at the country level.

Overall, we observe that religiosity is negatively correlated with different aspects of technological progress. These findings are also consistent with the proposed potential mediation channel, i.e., the adverse attitude towards science and technology associated with higher levels of religiosity, as discussed in the preceding subsection.

Table 3.9: Other measures of technology

	(1)	(2)	(3)	(4)	(5)	(6)
Dep. Var. =	Economic complexity index (2000)	TFP (2000)	ISO9001 quality certificates (2015)	R&D expenditure (% of GDP)	Employment in knowledge intensive sector (2015)	Global innovation index (2017)
<i>Religiosity</i>	-0.618*** (-5.568)	-0.550*** (-3.468)	-0.330*** (-3.454)	-0.507*** (-4.820)	-0.612*** (-5.344)	-0.708*** (-9.190)
Baseline control	Yes	Yes	Yes	Yes	Yes	Yes
R-squared	0.426	0.446	0.234	0.308	0.457	0.563
Observations	76	48	74	73	68	74

Notes: This table reports the correlation between religiosity and other measures of technology using OLS estimations. Standardized beta coefficients of regressions are presented, heteroskedasticity robust standard errors are used and *t* statistics are reported in parentheses. *, ** and *** denote significance at the 10%, 5% and 1% levels, respectively. The intercept estimates are not shown. The baseline controls used include island dummy, landlocked dummy, distance to coast, mean elevation and terrain ruggedness.

3.6 Conclusion

This paper attempts to examine the effects of religiosity on the differentials in the level of technology adoption across countries. Using both the OLS and IV estimations, we establish a robustly negative correlation between them. Heightened levels of religiosity in individuals tend to correlate with more adverse attitude towards science and technology, and consequently are associated with lower levels of technology adoption in society; religiosity also has similar effects on other aspects of technological progress. Our study thus contributes to the ongoing discussions on the enduring influence of religion on economic development in today's world.

Chapter 3 Appendix

This appendix provides a summary of the existing models on the interaction between religion and science/technology, additional regression results as well as details and data sources of all variables employed by the empirical analyses in the current paper.

3.A Religion and its intersection with science/technology

There have been several typologies proposed by historians on the relationship between religion and science (see, e.g., Barbour, 2000; Stenmark, 2004; De Cruz, 2018), which broadly fall under two categories, namely the independence and the contact view. Proponents of the independence view argue that there is no overlap between religion and science, owing to their epistemological differences that lead to exploration of different domains (see, e.g., Gould et al., 2014), i.e., values vs. fact. However, given the vast anecdotal evidence on the intersection of religion and science throughout the course of human history, and in particular, the different explanations offered by the two on topics such as the origin of humans, one may find it difficult to justify for a clear *a priori* demarcation between their domains.

Compared with that, the contact view, which acknowledges the overlap between religion and science is perhaps better received (see, for e.g., De Cruz, 2018). It is worth noting that to some extent, religion could be viewed as a precursor of science, since both seek to satisfy human beings' need to understand the world. In the course of human history, many catholic clerical scientists made significant contribution to the scientific community. For instance, Nicolaus Copernicus published the influential heliocentric model that places the Sun as the center of the universe rather than the Earth in 1543, which pioneered the ensuing Scientific Revolution. George Mendel, who is honored as the father of genetics for his work of discovering the three fundamental laws of inheritance, was an abbot of St. Thomas' Abbey in Moravia.

Several models have since been proposed to characterize the past interactions between religion and science, namely the conflict, the dialogue and the integration models (Barbour, 2000). The conflict model, which assumes that religion and science are in perpetual conflict, has now been increasingly viewed by most historians as shallow interpretation of historical record based on a misreading of events such as the Galileo affair. The dialogue model proposes that religion and science "can be in a graceful duet based on their epistemological overlaps" (De Cruz, 2018, see examples in the

preceding paragraph). The integration model attempts to unify religion and science in various ways but has pale evidence for supporting the purported harmony between them. In sum, there is currently no consensus on a single model that can satisfactorily characterize the interactions between religion and science while doing justice to their complexities.

3.B Additional regression results

Table 3.B1 presents the regression results using the two-stage least squares (2SLS) estimations to examine the effects of religiosity on technology adoption on the sectoral levels.

Table 3.B1: Regression results using 2SLS estimations- Sectoral indices

	(1)	(2)	(3)	(4)
<i>Panel A: Second-stage results</i>				
Dep. Var. =	Agriculture sector index	Communications sector index	Transportation sector index	Industry sector index
<i>Religiosity</i>	-1.407*** (-4.063)	-1.381*** (-4.862)	-0.839*** (-3.855)	-1.555*** (-4.591)
Baseline control	Yes	Yes	Yes	Yes
<i>Panel B: First-stage results</i>				
Dep. Var. = Religiosity				
<i>Pathogen</i>	0.541*** (4.413)	0.519*** (4.671)	0.551*** (5.294)	0.507*** (4.224)
Baseline control	Yes	Yes	Yes	Yes
R-squared	0.3784	0.3555	0.4208	0.3479
First stage F-statistic	19.4769	21.8189	28.0316	17.8383
Observations	66	78	72	67

Notes: This table reports the correlation between religiosity and technology adoption, using pathogen stress as the instrumental variable for religiosity. Both regression results of first stage and second stage estimations are presented. Standardized beta coefficients of regressions are presented, heteroskedasticity robust standard errors are used and *t* statistics are reported in parentheses. *, ** and *** denote significance at the 10%, 5% and 1% levels, respectively. The intercept estimates are not shown. The baseline controls used include island dummy, landlocked dummy, distance to coast, mean elevation and terrain ruggedness.

3.C Variable definitions and data sources

3.C.1 Main dependent variable

Our main dependent variable is the technology adoption index in 2000 AD, which captures the average adoption level of technology in each country for the year 2000, from Comin et al. (2010). The sector-based indices of technology adoption cover the following sectors: agriculture, communications, transportation and industry.

3.C.2 Main explanatory variable

Our main explanatory variable is the country-level measure of religiosity constructed using data from the World Values Survey covering 98 countries. We use data from all available survey waves, i.e., for the year of 1981-1984, 1990-1994, 1995-1998, 1999-2004, 2005-2009 and 2010-2014.

We construct measures of religiosity based on the following five questions. The first question asks whether the survey respondent is a religious person, we denote this measure as “R1: religious person”. The second asks the survey respondent how important religion is in his/her life, we denote this measure as “R2: importance of religion”. The third asks whether the survey respondent believes in God, we denote this measure as “R3: belief in God”. The fourth question asks the survey respondent how important God is in his/her life, we denote this measure as “R4: importance of God”. The last question the survey respondent how often he/she attends religious services, we denote it as “R5: church attendance”. For all measures, we rescale the responses so that they take values that fall within the range of 0 to 1, whereby a larger value corresponds to a higher degree of religiosity. Respective country-level measures of religiosity are constructed by aggregating all survey responses in each country.

Lastly, we also construct a summary index of religiosity by taking the average of all five measures, denoted as “Religiosity”, to represent the overall level of religiosity for each country.

3.C.3 Instrumental variable

Pathogen stress: a standardized historical pathogen prevalence index, which covers nine diseases including leishmania, schistosoma, trypanosoma, malaria, filaria, leprosy, dengue, typhus and tuberculosis, from Murray and Schaller (2010).

3.C.4 Control variables

3.C.4.1 Geographic controls (cross - country analysis)

Island: A dummy variable indicating 1 if a country is an island and 0 otherwise, from the CIA world fact book.

Landlocked: A dummy variable indicating 1 if a country is fully enclosed by land and 0 otherwise, from the CIA world fact book.

Distance to coast: the average distance to the nearest coast (10^3 km), from the Geographically based economic data (G-ECON) project.

Elevation: the average elevation of a country above sea level, from the G-ECON project.

Terrain ruggedness: an index measures terrain irregularities of a country, from Nunn and Puga (2012).

Notes: CIA: Central Intelligence Agency.

3.C.4.2 Controls (individual - level analysis)

Latitude / Distance to coast / Elevation / Terrain roughness at the district level: these are constructed from the grid cell data of the G-ECON project using GIS.

Age / Marital status / Gender / Education / Income: these individual control variables are from the WVS data.

3.C.4.3 Controls used in robustness tests

Protestants (%) / Muslims (%) / Catholics (%) / Buddhists (%): the population share that follows the religion in a given country for the year 2000, from McCleary and Barro (2006).

Individualism: a Hofstede index defined as “a preference for a loosely-knit social framework in which individuals are expected to take care of only themselves and their immediate families”, from Hofstede (2010) and the Hofstede Centre.

Predicted genetic diversity: an index calculated using information on expected heterozygosity at population level and ancient migration distance from East Africa, from Ashraf and Galor (2013). The ancestry adjusted version is constructed using the migration matrix from Putterman and Weil (2010) to make it compatible for contemporary populations.

Ethnic fractionalization: the probability that two randomly selected people come from different ethnic groups, from Alesina et al. (2003).

Language fractionalization: the probability that two randomly selected people speak different ethnic languages, from Alesina et al. (2003).

Religion fractionalization: the probability that two randomly selected people have different religious beliefs, from Alesina et al. (2003).

Technology adoption in 1500 AD: the average level of technology adoption in 1500 AD, from Comin et al. (2010).

GDP per capita in 2000 (log): log of GDP per capita for the year 2000 converted to constant 2005 international dollar using PPP rates, from the World Development Indicators.

Tertiary education: the average of years of tertiary schooling for the population aged 15 and above for the year 2000, from Barro and Lee (2013).

Trade-to-GDP ratio: the sum of exports and imports of goods and services measured as a share of gross domestic product for the year 2000, from the World Bank.

Common law / French civil law / German civil law / Scandinavian law: indicator variables of different legal origins for each country, from La Porta et al. (2008).

Intellectual Property protection: the average of intellectual property protection indices for the year 2006-2014, from the World Economic Forum World Economic Forum (2015).

3.C.5 Attitude and preference towards science and technology

All measures are constructed using data of the WVS. For all measures, we rescale the responses so that they take values that fall within the range of 0 to 1.

1. *Science and technology make our way of life change too fast*: we recode responses such that 0 denotes agree and 1 for disagree.

2. *The world is better off, or worse off, because of science and technology*: we code responses such that 0 denotes worse off and 1 for better off.

3. *We depend too much on science and not enough on faith*: we recode responses such that 0 denotes agree and 1 for disagree.

4. *Whenever science and religion conflict, religion is always right*: 0 denotes agree and 1 for disagree.

5. *It is not important for me to know about science in my daily life*: we recode responses such that 0 denotes agree and 1 for disagree.

6. *One of the bad effects of science is that it breaks down people's ideas of right and wrong*:

we recode responses such that 0 denotes agree and 1 for disagree.

3.C.6 Other measures of technology

Economic Complexity Index (2000): a measure of national knowledge intensity developed by Hidalgo and Hausmann (2009).

TFP in 2000 AD: the estimated total factor productivity relative to the U.S. for year 2000, from the United Nations Industrial Development Organization website: (<https://www.unido.org/data1/wpd/Index.cfm>).

Number of ISO 9001:2015 certificates for quality management systems: the number of ISO 9001:2015 certificates for quality management systems issued (per billion PPP\$ GDP) for the year 2015, from the Global Innovation Index website.

R&D expenditure: the average value of research and development expenditure as percentage of GDP for the year of 1995-2015, from the World Bank.

Employment in knowledge-intensive services: the employment in knowledge-intensive services as percentage of workforce for the year 2015, from the Global Innovation Index website.

Global Innovation Index (GII) for the year 2017: an index based on the annual ranking on countries' innovation success and capacities published by Cornell University, INSEAD and the World Intellectual Property Organization, from the Global Innovation Index website (<https://www.globalinnovationindex.org/>).

Bibliography

- Acemoglu, D., Johnson, S., and Robinson, J. (2005). The rise of europe: Atlantic trade, institutional change, and economic growth. *American economic review*, 95(3):546–579.
- Acemoglu, D., Johnson, S., and Robinson, J. A. (2001). The colonial origins of comparative development: An empirical investigation. *American Economic Review*, 91(5):1369–1401.
- Acemoglu, D., Johnson, S., and Robinson, J. A. (2002). Reversal of fortune: Geography and institutions in the making of the modern world income distribution. *The Quarterly Journal of Economics*, 117(4):1231–1294.
- Acemoglu, D., Naidu, S., Restrepo, P., and Robinson, J. A. (2019). Democracy does cause growth. *Journal of Political Economy*, 127(1):47–100.
- Ahlerup, P. and Olsson, O. (2012). The roots of ethnic diversity. *Journal of Economic Growth*, 17(2):71–102.
- Aiken, L. S., West, S. G., and Reno, R. R. (1991). *Multiple regression: Testing and interpreting interactions*. Sage.
- Alesina, A., Devleeschauwer, A., Easterly, W., Kurlat, S., and Wacziarg, R. (2003). Fractionalization. *Journal of Economic Growth*, 8(2):155–194.
- Alesina, A. and Ferrara, E. L. (2005). Ethnic diversity and economic performance. *Journal of Economic Literature*, 43(3):762–800.
- Alesina, A. and Giuliano, P. (2015). Culture and institutions. *Journal of Economic Literature*, 53(4):898–944.
- Alesina, A., Giuliano, P., and Nunn, N. (2013). On the origins of gender roles: Women and the plough. *The Quarterly Journal of Economics*, 128(2):469–530.
- Altonji, J. G., Elder, T. E., and Taber, C. R. (2005). Selection on observed and unobserved variables: Assessing the effectiveness of Catholic schools. *Journal of Political Economy*, 113(1):151–184.
- Ang, J. B. (2013). Institutions and the long-run impact of early development. *Journal of Development Economics*, 105:1–18.
- Ang, J. B. (2015). Agricultural legacy, individualistic culture, and technology adoption. Technical report, EGC, Nanyang Technological University.
- Ang, J. B. and Fredriksson, P. G. (2017). Wheat agriculture and family ties. *European Economic Review*, 100:236–256.
- Ashraf, Q. and Galor, O. (2013). The ‘Out of Africa’ hypothesis, human genetic diversity, and comparative economic development. *American Economic Review*, 103(1):1–46.
- Barbour, I. G. (2000). When science meets religion: Enemies. *Strangers, or Partners*, 164.
- Barro, R. J. and Lee, J. W. (2013). A new data set of educational attainment in the world, 1950–2010. *Journal of Development Economics*, 104:184–198.
- Barro, R. J. and McCleary, R. M. (2003). Religion and economic growth across countries. *American Sociological Review*, 68(5):760–781.
- Becker, S. O. and Woessmann, L. (2009). Was weber wrong? a human capital theory of protestant economic history. *The Quarterly Journal of Economics*, 124(2):531–596.
- Bénabou, R., Ticchi, D., and Vindigni, A. (2015a). Forbidden fruits: the political

- economy of science, religion, and growth. Technical report, National Bureau of Economic Research.
- Bénabou, R., Ticchi, D., and Vindigni, A. (2015b). Religion and innovation. *American Economic Review*, 105(5):346–51.
- Benjamin, J., Li, L., Patterson, C., Greenberg, B. D., Murphy, D. L., and Hamer, D. H. (1996). Population and familial association between the D4 dopamine receptor gene and measures of novelty seeking. *Nature Genetics*, 12(1):81.
- Bentzen, J. S., Kaarsen, N., and Wingender, A. M. (2017). Irrigation and autocracy. *Journal of the European Economic Association*, 15(1):1–53.
- Berkowitz, D., Caner, M., and Fang, Y. (2012). The validity of instruments revisited. *Journal of Econometrics*, 166(2):255–266.
- Boix, C. (2003). *Democracy and redistribution*. Cambridge University Press.
- Boix, C. and Stokes, S. C. (2003). Endogenous democratization. *World Politics*, 55(4):517–549.
- Bray, F. (1994). *The rice economies: technology and development in Asian societies*. Univ of California Press.
- Brewer, M. B. and Chen, Y.-R. (2007). Where (who) are collectives in collectivism? toward conceptual clarification of individualism and collectivism. *Psychological Review*, 114(1):133.
- Brock, W. A. and Durlauf, S. N. (2001). What have we learned from a decade of empirical research on growth? Growth empirics and reality. *The World Bank Economic Review*, 15(2):229–272.
- Brown, M. (2008). Good religion needs good science. *Posted on the Church of England website September, 12*.
- Bruce, S. (2004). Did Protestantism create democracy? *Democratization*, 11(4):3–20.
- Buck, J. L. et al. (1937). Land utilization in China.
- Campante, F. and Yanagizawa-Drott, D. (2015). Does religion affect economic growth and happiness? Evidence from Ramadan. *The Quarterly Journal of Economics*, 130(2):615–658.
- Chanda, A. and Putterman, L. (2007). Early starts, reversals and catch-up in the process of economic development. *Scandinavian Journal of Economics*, 109(2):387–413.
- Chauhan, B. S., Jabran, K., and Mahajan, G. (2017). *Rice Production Worldwide*, volume 247. Springer.
- Chen, C., Burton, M., Greenberger, E., and Dmitrieva, J. (1999). Population migration and the variation of dopamine D4 receptor (DRD4) allele frequencies around the globe. *Evolution and Human Behavior*, 20(5):309–324.
- Cloninger, C. R. (1994). Temperament and personality. *Current Opinion in Neurobiology*, 4(2):266–273.
- Cloninger, C. R., Przybeck, T. R., Svrakic, D. M., and Wetzel, R. D. (1994). The Temperament and Character Inventory (TCI): A guide to its development and use.
- Coetzer, G. and Gibbison, G. (2016). Mediating influence of time management on the relationship between adult attention deficit and the operational effectiveness of project managers. *Journal of Management Development*, 35(8):970–984.
- Coetzer, G. H. and Trimble, R. (2010). An empirical examination of the relationship between adult attention deficit, cooperative conflict management and efficacy for working in teams. *American Journal of Business*, 25(1):25–34.
- Comin, D., Easterly, W., and Gong, E. (2010). Was the wealth of nations determined in 1000 BC? *American Economic Journal: Macroeconomics*, 2(3):65–97.
- Comin, D. and Hobijn, B. (2010). An exploration of technology diffusion. *American Economic Review*, 100(5):2031–59.
- Comin, D., Hobijn, B., and Rovito, E. (2008). Technology usage lags. *Journal of Economic*

- Growth*, 13(4):237–256.
- Coppedge, M., Gerring, J., Lindberg, S. I., Skaaning, S.-E., Teorell, J., Altman, D., Bernhard, M., Fish, M. S., Glynn, A., Hicken, A., et al. (2015). V-dem [country-year/country-date] dataset v5. *Varieties of Democracy (V-Dem) Project*.
- Cox, P. L., Friedman, B. A., and Tribunella, T. (2011). Relationships among cultural dimensions, national gross domestic product, and environmental sustainability. *Journal of Applied Business and Economics*, 12(6):46–56.
- De Cruz, H. (2018). Religion and science. In Zalta, E. N., editor, *The Stanford Encyclopedia of Philosophy*. Metaphysics Research Lab, Stanford University, fall 2018 edition.
- De Graaf, R., Kessler, R. C., Fayyad, J., ten Have, M., Alonso, J., Angermeyer, M., Borges, G., Demyttenaere, K., Gasquet, I., de Girolamo, G., et al. (2008). The prevalence and effects of adult attention-deficit/hyperactivity disorder (ADHD) on the performance of workers: results from the WHO World Mental Health Survey Initiative. *Occupational and Environmental Medicine*.
- Diamond, J. M. (1997). *Guns, Germs, and Steel: The Fates of Human Societies*. New York: W.W. Norton.
- Dreber, A., Apicella, C. L., Eisenberg, D. T., Garcia, J. R., Zamore, R. S., Lum, J. K., and Campbell, B. (2009). The 7R polymorphism in the dopamine receptor D4 gene (DRD4) is associated with financial risk taking in men. *Evolution and Human Behavior*, 30(2):85–92.
- Dreber, A., Rand, D. G., Wernerfelt, N., Garcia, J. R., Vilar, M. G., Lum, J. K., and Zeckhauser, R. (2011). Dopamine and risk choices in different domains: Findings among serious tournament bridge players. *Journal of Risk and Uncertainty*, 43(1):19–38.
- Easterly, W. (2007). Inequality does cause underdevelopment: Insights from a new instrument. *Journal of Development Economics*, 84(2):755–776.
- Easterly, W. and Levine, R. (1997). Africa’s growth tragedy: policies and ethnic divisions. *The Quarterly Journal of Economics*, 112(4):1203–1250.
- Easterly, W. and Levine, R. (2016). The European origins of economic development. *Journal of Economic Growth*, 21(3):225–257.
- Ebstein, R., Nemanov, L., Klotz, I., Gritsenko, I., and Belmaker, R. (1997). Additional evidence for an association between the dopamine D4 receptor (D4DR) exon III repeat polymorphism and the human personality trait of Novelty Seeking. *Molecular Psychiatry*, 2(6):472.
- Ebstein, R. P., Novick, O., Umansky, R., Priel, B., Osher, Y., Blaine, D., Bennett, E. R., Nemanov, L., Katz, M., and Belmaker, R. H. (1996). Dopamine D4 receptor (D4DR) exon III polymorphism associated with the human personality trait of novelty seeking. *Nature Genetics*, 12(1):78.
- Fagerberg, J., Srholec, M., and Verspagen, B. (2010). Innovation and economic development. In *Handbook of the Economics of Innovation*, volume 2, pages 833–872. Elsevier.
- FAO and IIASA (2000). Global Agro-Ecological Zones (GAEZ).
- Fincher, C. L. and Thornhill, R. (2012). Parasite-stress promotes in-group assortative sociality: The cases of strong family ties and heightened religiosity. *Behavioral and Brain Sciences*, 35(2):61–79.
- Fincher, C. L., Thornhill, R., Murray, D. R., and Schaller, M. (2008). Pathogen prevalence predicts human cross-cultural variability in individualism/collectivism. *Proceedings of the Royal Society of London B: Biological Sciences*, 275(1640):1279–1285.
- Fischer, G., van Velthuisen, H. T., and Nachtergaele, F. O. (2000). Global agro-ecological zones assessment: methodology and results.
- Fogli, A. and Veldkamp, L. (2012). Germs, social networks and growth. Technical report,

- National Bureau of Economic Research.
- Fouka, V. and Schlaepfer, A. (2017). Agricultural returns to labor and the origins of work ethics.
- Friedman, B. M. (2006). The moral consequences of economic growth. *Society*, 43(2):15–22.
- Galor, O. and Klemp, M. (2017). Roots of autocracy. Technical report, National Bureau of Economic Research.
- Galor, O. and Michalopoulos, S. (2012). Evolution and the growth process: Natural selection of entrepreneurial traits. *Journal of Economic Theory*, 147(2):759–780.
- Galor, O. and Moav, O. (2002). Natural selection and the origin of economic growth. *The Quarterly Journal of Economics*, 117(4):1133–1191.
- Geroski, P. A. (2000). Models of technology diffusion. *Research policy*, 29(4-5):603–625.
- Gilmore, E., Gleditsch, N. P., Lujala, P., and Rød, J. K. (2005). Conflict diamonds: A new dataset. *Conflict Management and Peace Science*, 22(3):257–272.
- Giuliano, P. and Nunn, N. (2013). The Transmission of Democracy: From the Village to the Nation-State. *American Economic Review*, 103(3):86–92.
- Glaeser, E. L., La Porta, R., Lopez-de Silanes, F., and Shleifer, A. (2004). Do institutions cause growth? *Journal of economic Growth*, 9(3):271–303.
- Goncalo, J. A. and Staw, B. M. (2006). Individualism–collectivism and group creativity. *Organizational Behavior and Human Decision Processes*, 100(1):96–109.
- Gören, E. (2016). The biogeographic origins of novelty-seeking traits. *Evolution and Human Behavior*, 37(6):456–469.
- Gören, E. (2017). The persistent effects of novelty-seeking traits on comparative economic development. *Journal of Development Economics*, 126:112–126.
- Gorodnichenko, Y. and Roland, G. (2011). Individualism, innovation, and long-run growth. *Proceedings of the National Academy of Sciences*, 108(Supplement 4):21316–21319.
- Gorodnichenko, Y. and Roland, G. (2015). Culture, institutions and democratization. Technical report, National Bureau of Economic Research.
- Gorodnichenko, Y. and Roland, G. (2017). Culture, institutions, and the wealth of nations. *Review of Economics and Statistics*, 99(3):402–416.
- Gould, S. J. et al. (2014). Nonoverlapping magisteria. *Filozoficzne Aspekty Genezy*, (11):7–21.
- Gründler, K. and Krieger, T. (2016). Democracy and growth: Evidence from a machine learning indicator. *European Journal of Political Economy*, 45:85–107.
- Gui, R., Meierer, M., and Algesheimer, R. (2019). Rendo: A package to address endogeneity without external instrumental variables.
- Guiso, L., Sapienza, P., and Zingales, L. (2003). People’s opium? religion and economic attitudes. *Journal of Monetary Economics*, 50(1):225–282.
- Guiso, L., Sapienza, P., and Zingales, L. (2006). Does culture affect economic outcomes? *Journal of Economic Perspectives*, 20(2):23–48.
- Gurr, T. R., Marshall, M. G., and Jaggers, K. (2010). Polity iv project: Political regime characteristics and transitions, 1800–2009. *Center for International Development and Conflict Management at the University of Maryland College Park*.
- Halbesleben, J. R., Wheeler, A. R., and Shanine, K. K. (2013). The moderating role of attention-deficit/hyperactivity disorder in the work engagement–performance process. *Journal of Occupational Health Psychology*, 18(2):132.
- Hartmann, D., Guevara, M. R., Jara-Figueroa, C., Aristarán, M., and Hidalgo, C. A. (2017). Linking economic complexity, institutions, and income inequality. *World Development*, 93:75–93.
- Hausmann, R., Hidalgo, C. A., Bustos, S., Coscia, M., Simoes, A., and Yildirim, M. A.

- (2014). *The atlas of economic complexity: Mapping paths to prosperity*. Mit Press.
- Hibbs, D. A. and Olsson, O. (2004). Geography, biogeography, and why some countries are rich and others are poor. *Proceedings of the National Academy of Sciences*, 101(10):3715–3720.
- Hidalgo, C. A. and Hausmann, R. (2009). The building blocks of economic complexity. *Proceedings of the National Academy of Sciences*, 106(26):10570–10575.
- Hofstede, G. (1984). *Culture's consequences: International differences in work-related values*, volume 5. Sage.
- Hofstede, G., H. G. J. . M. M. (2010). *Cultures and organizations: Software of the mind*. New York, NY: McGraw Hill.
- Holsinger, K. E. and Weir, B. S. (2009). Genetics in geographically structured populations: defining, estimating and interpreting F_{ST} . *Nature Reviews Genetics*, 10(9):639.
- Iannaccone, L. R. (1998). Introduction to the economics of religion. *Journal of economic literature*, 36(3):1465–1495.
- Index, G. I. (2017). Innovation feeding the world. INCEAD.–2017.–463 p.–Access mode: <https://www.globalinnovationindex.org>.
- Inglehart, R. and Baker, W. E. (2000). Modernization, cultural change, and the persistence of traditional values. *American Sociological Review*, pages 19–51.
- Jaccard, J., Wan, C. K., and Turrisi, R. (1990). The detection and interpretation of interaction effects between continuous variables in multiple regression. *Multivariate behavioral research*, 25(4):467–478.
- Jensen, N. and Wantchekon, L. (2004). Resource wealth and political regimes in Africa. *Comparative Political Studies*, 37(7):816–841.
- Kateb, G. (1992). *The inner ocean: Individualism and democratic culture*. Cornell University Press.
- Kaufmann, D., Kraay, A., and Mastruzzi, M. (2011). The worldwide governance indicators: methodology and analytical issues. *Hague Journal on the Rule of Law*, 3(2):220–246.
- Kessler, R., Lane, M., Stang, P., and Van Brunt, D. (2009). The prevalence and workplace costs of adult attention deficit hyperactivity disorder in a large manufacturing firm. *Psychological Medicine*, 39(1):137–147.
- Kessler, R. C., Adler, L., Ames, M., Barkley, R. A., Birnbaum, H., Greenberg, P., Johnston, J. A., Spencer, T., and Üstün, T. B. (2005). The prevalence and effects of adult attention deficit/hyperactivity disorder on work performance in a nationally representative sample of workers. *Journal of Occupational and Environmental Medicine*, 47(6):565–572.
- Khush, G. S. (1997). Origin, dispersal, cultivation and variation of rice. *Plant Molecular Biology*, 35(1-2):25–34.
- Kuhnen, C. M. and Chiao, J. Y. (2009). Genetic determinants of financial risk taking. *PloS One*, 4(2):e4362.
- Küpper, T., Haavik, J., Drexler, H., Ramos-Quiroga, J. A., Wermelskirchen, D., Prutz, C., and Schauble, B. (2012). The negative impact of attention-deficit/hyperactivity disorder on occupational health in adults and adolescents. *International Archives of Occupational and Environmental Health*, 85(8):837–847.
- La Porta, R., Lopez-de Silanes, F., and Shleifer, A. (2008). The economic consequences of legal origins. *Journal of Economic Literature*, 46(2):285–332.
- La Porta, R., Lopez-de Silanes, F., Shleifer, A., and Vishny, R. (1999). The quality of government. *The Journal of Law, Economics, and Organization*, 15(1):222–279.
- Lange, M., Mahoney, J., and Vom Hau, M. (2006). Colonialism and development: a comparative analysis of Spanish and British colonies. *American Journal of Sociology*, 111(5):1412–1462.
- Leung, P. W., Lee, C., Hung, S., Ho, T., Tang, C., Kwong, S., Leung, S., Yuen, S., Lieh-Mak,

- F., Oosterlaan, J., et al. (2005). Dopamine receptor D4 (DRD4) gene in Han Chinese children with attention-deficit/hyperactivity disorder (ADHD): Increased prevalence of the 2-repeat allele. *American Journal of Medical Genetics Part B: Neuropsychiatric Genetics*, 133(1):54–56.
- Leung, P. W.-l., Chan, J. K. Y., Chen, L. H., Lee, C. C., Hung, S. F., Ho, T. P., Tang, C. P., Moyzis, R. K., and Swanson, J. M. (2017). Family-based association study of DRD4 gene in methylphenidate-responded Attention Deficit/Hyperactivity Disorder. *PloS One*, 12(3):e0173748.
- Lind, J. T. and Mehlum, H. (2010). With or without U? the Appropriate Test for a U-Shaped Relationship. *Oxford Bulletin of Economics and Statistics*, 72(1):109–118.
- Lipset, S. M. (1959). Some social requisites of democracy: Economic development and political legitimacy. *American Political Science Review*, 53(1):69–105.
- Madsen, J. B., Raschky, P. A., and Skali, A. (2015). Does democracy drive income in the world, 1500–2000? *European Economic Review*, 78:175–195.
- Mahdavy, H. and Cook, M. (1970). The patterns and problems of economic development in rentier states: The case of Iran. *Life*, 1000(1).
- Markus, H. R. and Kitayama, S. (1991). Culture and the self: Implications for cognition, emotion, and motivation. *Psychological Review*, 98(2):224.
- Markus, H. R. and Kitayama, S. (1994). A collective fear of the collective: Implications for selves and theories of selves. *Personality and Social Psychology Bulletin*, 20(5):568–579.
- Marshall, M. G., Gurr, T., and Jaggers, K. (2014). Polity IV project. Political regime characteristics and transitions, 1800–2013. Dataset users' manual. Center for Systemic Peace.
- Marx, K. (1973). Foundations of the critique of political economy.
- Marx, K. (1977). *CAPITAL EN 3 VOL. ED SOCIALES*.
- Matthews, L. J. and Butler, P. M. (2011). Novelty-seeking DRD4 polymorphisms are associated with human migration distance out-of-Africa after controlling for neutral population gene structure. *American Journal of Physical Anthropology*, 145(3):382–389.
- Matza, L. S., Paramore, C., and Prasad, M. (2005). A review of the economic burden of ADHD. *Cost Effectiveness and Resource Allocation*, 3(1):5.
- Mayshar, J., Moav, O., Neeman, Z., and Pascali, L. (2015). Cereals, appropriability and hierarchy.
- McCleary, R. M. and Barro, R. J. (2006). Religion and economy. *Journal of Economic Perspectives*, 20(2):49–72.
- McPhetres, J. and Zuckerman, M. (2018). Religiosity predicts negative attitudes towards science and lower levels of science literacy. *PloS One*, 13(11):e0207125.
- Mokyr, J. (2005a). The intellectual origins of modern economic growth. *The Journal of Economic History*, 65(2):285–351.
- Mokyr, J. (2005b). Long-term economic growth and the history of technology. In *Handbook of Economic Growth*, volume 1, pages 1113–1180. Elsevier.
- Mokyr, J. (2014). Culture, institutions, and modern growth. *Institutions, Property Rights, and Economic Growth: The Legacy of Douglass North*, pages 151–191.
- Munck, G. L. and Verkuilen, J. (2002). Conceptualizing and measuring democracy: Evaluating alternative indices. *Comparative Political Studies*, 35(1):5–34.
- Murray, D. R. and Schaller, M. (2010). Historical prevalence of infectious diseases within 230 geopolitical regions: A tool for investigating origins of culture. *Journal of Cross-Cultural Psychology*, 41(1):99–108.
- Muthayya, S., Sugimoto, J. D., Montgomery, S., and Maberly, G. F. (2014). An overview of global rice production, supply, trade, and consumption. *Annals of the New York Academy of Sciences*, 1324(1):7–14.

- Nordhaus, W., Azam, Q., Corderi, D., Hood, K., Victor, N. M., Mohammed, M., Miltner, A., and Weiss, J. (2006). The G-Econ database on gridded output: Methods and data. *Yale University, New Haven*, 6.
- North, D. C. (1991). Institutions. *Journal of Economic Perspectives*, 5(1):97–112.
- Nunn, N. (2009). The importance of history for economic development. *Annu. Rev. Econ.*, 1(1):65–92.
- Nunn, N. and Puga, D. (2012). Ruggedness: The blessing of bad geography in Africa. *Review of Economics and Statistics*, 94(1):20–36.
- Nunn, N. and Qian, N. (2011). The potato's contribution to population and urbanization: evidence from a historical experiment. *The Quarterly Journal of Economics*, 126(2):593–650.
- Olsson, O. (2009). On the democratic legacy of colonialism. *Journal of Comparative Economics*, 37(4):534–551.
- Olsson, O. and Hibbs Jr, D. A. (2005). Biogeography and long-run economic development. *European Economic Review*, 49(4):909–938.
- Oster, E. (2017). Unobservable selection and coefficient stability: Theory and evidence. *Journal of Business & Economic Statistics*, pages 1–18.
- Pargament, K. I. (2001). *The psychology of religion and coping: Theory, research, practice*. Guilford Press.
- Park, C., Cohen, L. H., and Herb, L. (1990). Intrinsic religiousness and religious coping as life stress moderators for Catholics versus Protestants. *Journal of Personality and Social Psychology*, 59(3):562.
- Park, S. and Gupta, S. (2012). Handling endogenous regressors by joint estimation using copulas. *Marketing Science*, 31(4):567–586.
- Paul, P. J. (1996). Message to the pontifical academy of sciences on evolution. *ORIGINS-WASHINGTON-*, 26:349–349.
- Pio, P. et al. (1950). *Humani generis*. Universidad.
- Putterman, L. (2004). State Antiquity Index (Statehist) Dataset, version 3.1. *Brown University*.
- Putterman, L. (2008). Agriculture, diffusion and development: Ripple effects of the neolithic revolution. *Economica*, 75(300):729–748.
- Putterman, L. and Trainor, C. A. (2006). Agricultural transition year country data set. *Brown University*, pages 729–48.
- Putterman, L. and Weil, D. N. (2010). Post-1500 population flows and the long-run determinants of economic growth and inequality. *The Quarterly Journal of Economics*, 125(4):1627–1682.
- Reist, C., Ozdemir, V., Wang, E., Hashemzadeh, M., Mee, S., and Moyzis, R. (2007). Novelty seeking and the dopamine D4 receptor gene (DRD4) revisited in Asians: Haplotype characterization and relevance of the 2-repeat allele. *American Journal of Medical Genetics Part B: Neuropsychiatric Genetics*, 144(4):453–457.
- Riquelme, A., Berkowitz, D., and Caner, M. (2013). Valid tests when instrumental variables do not perfectly satisfy the exclusion restriction. *The Stata Journal*, 13(3):528–546.
- Ross, M. L. (2001). Does oil hinder democracy? *World Politics*, 53(3):325–361.
- Russell, B. (1930). *Has religion made useful contributions to civilization?* Rationalist Press Association, Limited.
- Schaller, M. and Murray, D. R. (2008). Pathogens, personality, and culture: disease prevalence predicts worldwide variability in sociosexuality, extraversion, and openness to experience. *Journal of Personality and Social Psychology*, 95(1):212.
- Schweizer, T. S. (2006). The psychology of novelty-seeking, creativity and innovation: neurocognitive aspects within a work-psychological perspective. *Creativity and*

- Innovation Management*, 15(2):164–172.
- Shane, S. (1992). Why do some societies invent more than others? *Journal of Business Venturing*, 7(1):29–46.
- Shane, S. (1993). Cultural influences on national rates of innovation. *Journal of Business Venturing*, 8(1):59–73.
- Spolaore, E. and Wacziarg, R. (2009). The diffusion of development. *The Quarterly Journal of Economics*, 124(2):469–529.
- Staiger, D. O. and Stock, J. H. (1994). Instrumental variables regression with weak instruments.
- Stenmark, M. (2004). *How to relate science and religion: A multidimensional model*. Wm. B. Eerdmans Publishing.
- Stojkoski, V., Utkovski, Z., and Kocarev, L. (2016). The impact of services on economic complexity: service sophistication as route for economic growth. *PloS One*, 11(8):e0161633.
- Strobel, A., Wehr, A., Michel, A., and Brocke, B. (1999). Association between the dopamine D4 receptor (DRD4) exon III polymorphism and measures of novelty seeking in a German population. *Molecular Psychiatry*, 4(4):378.
- Swanson, J., Oosterlaan, J., Murias, M., Schuck, S., Flodman, P., Spence, M. A., Wasdell, M., Ding, Y., Chi, H.-C., Smith, M., et al. (2000). Attention deficit/hyperactivity disorder children with a 7-repeat allele of the dopamine receptor D4 gene have extreme behavior but normal performance on critical neuropsychological tests of attention. *Proceedings of the National Academy of Sciences*, 97(9):4754–4759.
- Sweeney, M. and McCouch, S. (2007). The complex history of the domestication of rice. *Annals of Botany*, 100(5):951–957.
- Talhelm, T., Zhang, X., Oishi, S., Shimin, C., Duan, D., Lan, X., and Kitayama, S. (2014). Large-scale psychological differences within China explained by rice versus wheat agriculture. *Science*, 344(6184):603–608.
- Taylor, M. Z. and Wilson, S. (2012). Does culture still matter?: The effects of individualism on national innovation rates. *Journal of Business Venturing*, 27(2):234–247.
- Teorell, J., Dahlberg, S., Holmberg, S., Rothstein, B., Hartmann, F., and Svensson, R. (2015). The quality of government standard dataset, version jan15. *University of Gothenburg: The Quality of Government Institute*.
- Thornhill, R. and Fincher, C. L. (2014). *The parasite-stress theory of values and sociality: Infectious disease, history and human values worldwide*. Springer.
- Tomitaka, M., Tomitaka, S.-i., Otuka, Y., Kim, K., Matuki, H., Sakamoto, K., and Tanaka, A. (1999). Association between novelty seeking and dopamine receptor D4 (DRD4) exon III polymorphism in Japanese subjects. *American Journal of Medical Genetics*, 88(5):469–471.
- Treier, S. and Jackman, S. (2008). Democracy as a latent variable. *American Journal of Political Science*, 52(1):201–217.
- Triandis, H. C. (2018). *Individualism and Collectivism*. Routledge.
- Vandello, J. A. and Cohen, D. (1999). Patterns of individualism and collectivism across the United States. *Journal of Personality and Social Psychology*, 77(2):279.
- Wang, E., Ding, Y.-C., Flodman, P., Kidd, J., Kidd, K., Grady, D., Ryder, O., Spence, M., Swanson, J., and Moyzis, R. (2004). The genetic architecture of selection at the human dopamine receptor D4 (DRD4) gene locus. *The American Journal of Human Genetics*, 74(5):931–944.
- Wantchekon, L. (2002). Why do resource dependent countries have authoritarian governments? *Journal of African Finance and Economic Development*, 2(1):57–77.
- WDI (2012). *World development indicators 2012 (CD ROM)*. World Bank Publications.
- Weber, M. (1930). *The Protestant ethic and the spirit of capitalism*. Routledge.

- Williams, D. R., Larson, D. B., Buckler, R. E., Heckmann, R. C., and Pyle, C. M. (1991). Religion and psychological distress in a community sample. *Social Science & Medicine*, 32(11):1257–1262.
- Williams, J. and Taylor, E. (2006). The evolution of hyperactivity, impulsivity and cognitive diversity. *Journal of the Royal Society Interface*, 3(8):399–413.
- Woodberry, R. D. (2012). The missionary roots of liberal democracy. *American Political Science Review*, 106(2):244–274.
- Wooldridge, J. M. (2010). *Econometric analysis of cross section and panel data*. MIT press.
- World Bank Group (2014). *World Development Indicators 2014*. World Bank Publications.
- World Economic Forum (2015). The global competitiveness report 2014-2015. Technical report, The World Economic Forum.
- World Values Survey Association and others (2015). World Values Survey 1981–2014 Longitudinal Aggregate v. 20150418.