Journal of Evolutionary Economics (2019) 29:973–1016
https://doi.org/10.1007/s00191-019-00613-2

REGULAR ARTICLE

Genetic distance and the difference in new firm entry between countries

Maria João Guedes[1] · Nicos Nicolaou[2] · Pankaj C. Patel[3]

Published online: 28 March 2019

Abstract

Does genetic distance between countries explain differences in the level of entrepreneurship between them? Genetic distance, or very long-term divergence in intergenerationally transmitted traits across populations, has been recently tied to a variety of outcomes ranging from differences in economic development to differences in risk preferences between countries. Extending this recent work, we ask whether the genetic distance between countries is associated with differences in new firm entry. Based on a sample of 103 countries and 5253 country-pair observations and controlling for a large variety of factors, we find that genetic distance is positively associated with between country differences in new firm entry. The effects sizes, as expected, are small. In assessing the differences in entrepreneurial activity between country-pairs, policymakers could consider adjusting for genetic distance as an explanation for differences in entrepreneurial activity.

Keywords Genetic distance · Entrepreneurship · Comparative entrepreneurship · Country or areas study · New firm entry

JEL classification L26 · M13

Maria João Guedes, Nicos Nicolaou and Pankaj C. Patel contributed equally to this work.

✉ Maria João Guedes
mjguedes@iseg.ulisboa.pt

Nicos Nicolaou
Nicos.Nicolaou@wbs.ac.uk

Pankaj C. Patel
pankaj.patel@villanova.edu

[1] Advance Research Center, ISEG (School of Economics and Management), Universidade de Lisboa, Rua do Quelhas 6, 1249-078 Lisbon, Portugal

[2] Warwick Business School, University of Warwick, Coventry CV4 7AL, UK

[3] Management & Operations, Villanova University, Bartley Hall Rm 2067, 800 Lancaster Avenue, Villanova, PA 19085, USA

1 Introduction

Researchers have examined a number of factors that influence the differences in rates of entrepreneurship across countries, ranging from tax rates (Hubbard 1998) to availability of capital for new firms (Amit et al. 1998) and from entrepreneurial culture (Krueger et al. 2013) and institutions supporting entrepreneurship (Hébert and Link 1988) to differences in socio-cultural norms (Aldrich and Fiol 1994; Thornton et al. 2011). At the root of social, political, institutional and economic differences between countries that explain the differences in entrepreneurship could be the genetic distance between countries, or "a measure associated with the time elapsed since two populations' last common ancestors," (Spolaore and Wacziarg 2009: 469). A series of studies have shown that, controlling for a variety of factors, genetic distance could explain between-country differences in technology, productivity (Spolaore and Wacziarg 2009), innovation and growth (Gorodnichenko and Roland 2017), risk-taking (Becker et al. 2014), trust (Guiso et al. 2009), income and economic development (Bai and Kung 2011; Spolaore and Wacziarg 2013) and well-being (Burger et al. 2015).

The aforementioned relationships between genetic distance and differences in a variety of between country-level outcomes, suggest accumulating evidence that points to the possibility that the genetic distance between two countries may be associated with differences in the rates of new firm entry between countries. Moving from prior studies in entrepreneurship on individual level biological characteristics and drawing on work linking genetic distance at the country level, we provide the first empirical test of the relationship between genetic distance and differences in entrepreneurship between countries.

Specifically, based on Spolaore and Wacziarg (2009) we examine the association between a country's genetic distance from the world's technological frontier (the US) and new firm entry. The choice of the US as a reference point in measuring genetic distance between two countries is based on Spolaore and Wacziarg (2009), who posit that geographic, cultural and genetic distance to the US is associated with a country's technological development. Their measure, based on Cavalli-Sforza et al. (1994), is the time elapsed since two population groups existed as a single panmimitic population and posits that genetic distance proxies a divergence in traits "biologically and/or culturally" that add barriers to the diffusion of technology (Campbell and Pyun 2017).

We test the relationship between the weighted genetic distance between pairs of 103 countries (5253 country-pair observations) and the differences in the entrepreneurial entry. Based on casewise deletion, in a sample of 820 country-pair observations, the inferences were consistent. Controlling for a wide range of measures including geographical, cultural, religious, linguistic, and historical differences, we find that genetic distance has a statistically significant impact on the differences in the new firm entry between countries. Adding to works of Spolaore and Wacziarg (2009) we include several additional controls – Hofstede's cultural dimensions (inferences were also robust to the inclusion of cultural dimensions from the World Value Survey), Worldwide Governance Indicators, legal origin and additional economic factors. The inclusion of these additional controls further adds to the robustness of our inferences. We find that, for one standard deviation increase in genetic distance between two countries, the difference in the start-up rate ranges from 1.100 to 1.120 firms per 1000 working-age population (those ages 15–64). While a complex set of factors could explain the efficacy of policies promoting entrepreneurship, in the current analysis, the variance explained by genetic distance is small but significant.

It is also important to emphasize that our study is *not* about identifying differences in specific genetic polymorphisms that directly influence entrepreneurship at the individual level.

Genetic polymorphisms are heritable genetic differences among individuals (Ding and Zhang 2010). Our study focuses on neutral genetic distance among countries (Spolaore and Wacziarg 2009) and not on specific genetic traits.

Neutral genetic distance is based on neutral genetic markers that are independent of historic natural selection pressures and hence do not directly influence survival and fitness in the short-term (Spolaore and Wacziarg 2009). As a result, our study is not about certain countries having a higher prevalence of specific genetic traits that directly increase the rates of entrepreneurship. Instead, genetic distance can be interpreted as an overall measure that captures a combination of intergenerationally transmitted characteristics between two countries that could explain the differences in the rates of entrepreneurship.

2 Theoretical background

We emphasize that our study is conducted at the between country-pair distance level of analysis even though we review genetic differences at the individual level as one of the multiple strands in the development of our argument. Traditional entrepreneurship literature shows that a variety of factors at different levels – individual, firm, and institutions – are associated with entrepreneurial activity. Individual-level factors that influence entrepreneurial activity include personality (Baum et al. 2014), cognition (Mitchell et al. 2007), affect (Baron 2008), and passion (Cardon et al. 2009), among others. While individual and firm-related differences significantly affect success in entrepreneurial activity, systematic differences still prevail in entrepreneurial activity across countries (Ács et al. 2014; Blanchflower et al. 2001; Klapper et al. 2010).

Closer to our empirical context, but not directly related to it, is the literature on the role of biological characteristics at the individual level and self-employment outcomes. We briefly discuss this below, however, we do not draw on this as the basis of our hypothesis.

2.1 Biology and self-employment

While institutional and cultural factors explain cross-country differences in entrepreneurship, biological factors associated with entrepreneurship have recently been studied. In providing a brief overview of this literature, we emphasize that genetic predispositions to entrepreneurship are not weighted differently when calculating genetic distance at the country level as the measure of genetic distance focuses on neutral genetic distance between countries and not on specific genetic traits that must be weighed differently to derive the overall genetic distance measure associated with entrepreneurship. Neutral genetic distance is based on neutral genetic markers that are independent of historic natural selection pressures and hence do not directly influence survival and fitness in the short-term (Spolaore and Wacziarg 2009:5). Below we provide a non-exhaustive review of the biology and entrepreneurship literature and refer interested readers to Nofal et al. (2018). The purpose of discussing this literature as a backdrop is to provide an understanding of the role of biology in entrepreneurship-related outcomes in general and not to provide intuition for the proposed hypothesis.

Studies on biology and entrepreneurship have found evidence of heritability for self-employment using studies of twins (Nicolaou et al. 2008) and adoptees (Lindquist et al. 2015). Studies have also shown that testosterone is associated with self-employment (Bönte et al. 2016; Greene et al. 2014). In addition, attention deficit hyperactivity disorder (Antshel 2017; Verheul et al. 2015; Verheul et al. 2016) and dyslexia (Hessels et al. 2014) are more prevalent among the self-employed. Studies have also advocated the use of neuroscientific methods to understand better the decision making patterns of the self-employed (de Holan 2014).

Complementing this rich body of work, we now zoom-out to between country genetic distance, to explain between-country differences in new firm entry.

2.2 Country-level differences in self-employment

A complex combination of social, cultural, institutional, and government factors have been shown to sustain cross-country differences in entrepreneurship over time. Work in comparative economics (Djankov et al. 2002), trade (Foss and Klein 2005) and public policy (Audretsch et al. 2007) have highlighted the role of government policies and country-level characteristics in explaining differences in entrepreneurial activity across countries. Hofstede's cultural dimensions have been associated with cross-country differences in entrepreneurial activities (Hayton and Cacciotti 2013). Cultural values and beliefs socially program individuals to engage in or to avoid entrepreneurship (Thomas and Mueller 2000), and cultural differences related to uncertainty avoidance or materialism explain differences in entrepreneurship rates. In addition to cultural differences, differences in economic conditions could also explain differences in entrepreneurial activities. Differences in policies ranging from bankruptcy laws to credit for private firms also explain differences in entrepreneurial activity among countries (King and Levine 1993).

There is also some related research on differences in psychological traits and differences in rates of entrepreneurship. For example, Rentfrow et al. (2008) used data from over half a million people in the US and found that geographic variation in psychological traits may lead to macro-level differences among regions. In a related paper, Obschonka et al. (2013) found evidence of regional clustering in an entrepreneurship-prone personality profile using data from the US, Germany and the UK.

While differences in entrepreneurial activity among countries could be explained by economic and technological, cultural and institutional differences (Grilo and Thurik 2005), the genetic distance between two countries could have been the harbinger of these differences over time. The need to focus on differences in new firm entry is particularly salient because, despite global integration, there continue to be systematic and persistent differences across countries in entrepreneurship activities. It is plausible that genetic distance between countries could lead to complex social processes over time that drive differences in cultural and political institutions that may, in turn, influence the differences in new firm entry.

Genetic distance between two countries is a result of complex migration patterns over thousands of years, a systematic path-dependent process that resulted from historic patterns of trade and migration. While the difference in entrepreneurship levels may thus not be solely explained by bottom-up micro-level behaviors of individual entrepreneurs, they could also be explained by top-down systematic genetic distances between countries. Indirectly supporting this conjecture is the growing evidence that genetic differences at the country-level influence differences in economic development and risk-taking, technology, and trust. Differences in entrepreneurship-related behaviors may be a result of long-term population stasis – systematic differences in genes among countries – that may explain differences in the level of entrepreneurship.

We propose that increasing genetic distance between two countries is positively associated with differences in entrepreneurial activity between two countries. Due to the lack of a systematic theoretical framework explaining the association between genetic distance and between-country differences in entrepreneurship, instead of providing robust theoretical arguments, the proposed arguments are geared towards building an intuition towards the proposed association.

2.3 Genetic distance and entrepreneurship

The concept of genetic distance is derived from the seminal work of Cavalli-Sforza et al. (1994) and is a measure of the difference in allelic frequencies across populations. A brief examination of the measure of genetic distance begins by considering genes, which are segments of DNA that encode for a certain function. An allele refers to different variants of a particular gene. The measure of genetic distance is essentially a summary of the differences in these alleles across different populations. Most significantly, genetic distance draws on neutral markers and not on selected traits (Spolaore and Wacziarg 2009). As a result, our study is definitely, unequivocally, *not* about specific genetic polymorphisms that directly matter for entrepreneurship.

A small but increasing number of studies have investigated the relationship between genetic distance and differences in economic and non-economic factors between countries. Spolaore and Wacziarg (2009) found that as a country's genetic distance increased, the differences in per capita income also increased. Recently, Spolaore and Wacziarg (2016) found that genetic distance was positively associated with linguistic, cultural and religious distance, and Proto and Oswald (2017) discovered a relationship between genetic distance and differences in well-being.

A related, but distinct concept to genetic distance is that of genetic diversity. "Genetic distance refers to genetic differences between populations while genetic diversity is defined in terms of heterogeneity within populations" (Spolaore and Wacziarg 2013: 355). Ashraf and Galor (2013) found a hump-shaped relationship between genetic diversity and economic development both before colonial periods and in the modern era. Genetic distance and genetic diversity seem to explain different aspects of genetic effects. Ashraf and Galor (2013) found that genetic diversity has decreasing returns to development, and Spolaore and Wacziarg (2009) found that lower genetic distance increases development. These findings are complementary as Spolaore and Wacziarg (2009) seem to have found support for the left side of hump-shaped effects found by Ashraf and Galor (2013) – thus, lower genetic differences ease the 'transaction costs' of coordination.

While we focus on the differences in new firm registrations between countries predicted by genetic distance, it is possible that genetic distance may also influence a range of country-level differences in the evolution of institutions, inequality, participation from the citizens, among others. We propose that genetic distance between a country-pair is likely to be associated with differences in entrepreneurship between a country-pair. Our intuition is rooted in the logic that genetic distance between country-pairs drives "divergence in the whole set of implicit beliefs, customs, habits, biases, conventions, etc. that are transmitted across generations—biologically and/or culturally—with high persistence" (Spolaore and Wacziarg 2009: 471).

Longer genetic distance has been associated with differential risk attitudes between populations (Becker et al. 2014), which may in turn also explain differences in entrepreneurship. Specifically, Becker et al. (2014) found that countries at shorter genetic distance had lower differences in risk attitudes. In related literature, there is heterogeneity in risk attitudes across countries (Falk et al. 2018), while risk attitudes have been shown to exhibit a genetic predisposition (Kuhnen and Chiao 2009). It is therefore possible that differences in genetic distance between populations may also be associated with differences in entrepreneurial activity.

Again, we emphasize that while we cannot measure the above mechanisms, our logic is based on the premise that, in genealogically distant country pairs, traits and characteristics between the two populations are likely to be distinct (Spolaore and Wacziarg 2009) and that, in turn, may influence the differences in entrepreneurial activities. Overall, the discussion above suggests that genetic distance is likely to be associated with cross-country differences in the rates of entrepreneurship.

> Hypothesis: Longer genetic distance between two countries is positively associated with differences in new firm entries between two countries.

3 Methods

3.1 Data sources

Our sample includes 103 countries and 5253 pairwise observations that result from the matching data on genetic distance (Spolaore and Wacziarg 2009), with data on new firm density from the *Doing Business* report from the WorldBank Group,[1] the United States Census Bureau data on Business Dynamics Statistics, and Federal Reserve (https://research.stlouisfed.org/fred2/).

Because the association of genetic distance could confound with a variety of country-level factors, Spolaore and Wacziarg (2009) used a wide range of controls. To further add robustness to our inferences, in addition to including all controls in Spolaore and Wacziarg (2009), we also included several additional controls related to Hofstede's cultural dimensions (and also tested for additional cultural dimensions from World Value Survey), *Worldwide Governance Indicators,* legal origin and economic characteristics.

[1] The data are available at: http://www.doingbusiness.org/data/exploretopics/entrepreneurship

3.2 Empirical specification

In order to investigate the relationship between the genetic distance between two countries and differences in the new firm entry between the two countries, we explored three specifications (Spolaore and Wacziarg 2009). The first and second specifications consider genetic distance relative to the United States of America, the 'technological frontier' as proposed by Spolaore and Wacziarg (2009), and calculate the genetic distance of every country relative to the US. In the first model, we examine the relationship between a country's genetic distance from the US and its level of entrepreneurial activity, while in the second specification we examine the relationship between a country's genetic distance from the US and the difference in entrepreneurial activity from the US.[2]

The third specification specifies a bilateral model, taking the absolute difference in entrepreneurial activity between pairs of countries and the weighted genetic distance for that pair (Spolaore and Wacziarg 2009). The baseline specification for the third model is:

$$|\log y_i - \log y_j| = \beta_0 + \beta_1 G_{ij}^D + \beta_2{}' X_{ij} + \varepsilon_{ij}$$

where represents the absolute weighted genetic distance between country i and j; y_k is the entrepreneurial activity for country k; X_{ij} is a vector of the control variables; and ε_{ij} is the error term. The dependent variable is the difference in log of the average number of newly registered companies in country i and j; the absolute value reduces the spatial dependence in the dependent variable.

The second specification is of a similar form to the third specification, except that the genetic difference between countries G_{ij}^D is replaced by the difference in relative distances to the US, $G_{ij}^R = |G_{i,US}^R - G_{j,US}^R|$, giving the equation $|\log y_i - \log y_j| = \hat{\beta}_0 + \hat{\beta}_1 G_{ij}^R + \hat{\beta}_2{}' X_{ij} + \hat{\varepsilon}_{ij}$ where any pairs where i or j are the US are excluded.

Because many countries are made up of different ethnic groups that are genetically different, it is important to use a weighted genetic distance measure that adjusts for genetic distance and the share of each ethnic group in a country (Alesina et al. 2003; Spolaore and Wacziarg 2009). Assuming country A contains ethnic groups $i = 1....I$ and country B contains ethnic groups $j = 1....J$; p_{Ai} is the share of ethnic group i in country A and p_{Bj} is the share of ethnic group j in country B; and d_{ij} is the genetic distance between ethnic groups i and j. The weighted measure of genetic distance (F_{ST}) is then given by:

$$F_{ST} = \sum_{i=1}^{I} \sum_{j=1}^{J} \left(p_{Ai} \times p_{Bj} \times d_{ij} \right)$$

[2] We focus on the US for two reasons. First, it is considered the "world technological frontier" (Spolaore and Wacziarg 2009) and second, all previous genetic distance studies used genetic distance relative to the US.

3.3 Estimation procedure

We emphasize that the genetic distance between pairs of countries could be driven by historic migration flows between two pairs of countries. Cultural and linguistic similarities, economic opportunities, among others (Gorodnichenko and Roland 2017; Günther and Jakobsson 2016) could have further influenced historic migration patterns. As such, due to historic bilateral migration patterns, genetic distance could be endogenous, that is, the unobservables related to bilateral migration patterns between country-pairs in the error term of the regression could influence both between country genetic distance and entrepreneurship activity. As such, causation is not implied and correlation is inferred in the testing of the proposed model.

We run two-way clustered standard error regressions based on all three specifications listed in the previous section. This estimation procedure calculates standard errors that account for two dimensions of within-cluster correlation between countries in a pair (Petersen 2009). The two dimensions are country i and country j, thereby allowing us to control for shared unobserved characteristics between country i and country j. Moreover, this estimation procedure provides more conservative estimates by controlling for spatial correlation between two countries (Cameron et al. 2011). Spolaore and Wacziarg (2009) provided an example of how spatial correlation can be present in a pairwise approach. With an illustration of three countries, the authors refered to the case where the observations for the dependent variable, $|\log y_1 - \log y_2|$ and $|\log y_1 - \log y_3|$ are correlated due to the presence of one of the countries (y_1) in both observations. In such a case, using simple least-squares standard errors, would lead to inflated estimates due to spatial (cross-sectional) correlations. Furthermore, we have bilateral variables such as genetic distance and geodesic distance in the right-hand-side of the equation.

3.4 Measures

3.4.1 Differences in new firm entry

Our dependent variable is the difference in the startup rate between two countries. Start-up rate is defined as the number of new limited liability firms per 1000 working-age people (15–64 years old). Based on the specification, we take the absolute difference in the log of the average number of newly registered companies for the period 2008 to 2010.[3] We take the natural log to reduce the influence of skewed rates of differences in entrepreneurship observed for some countries and to increase normality in the distribution of the outcome variable. The data were obtained from the Doing Business report from the WorldBank Group. Because no data were available from this report for the United States of America, we collected data from United States Census Bureau, Business Dynamics Statistics[4] and the Federal Reserve Economic Data - FRED - St. Louis Fed.[5]

[3] In the cases where we did not have data for all of the years, we used the average of the available years.
[4] The data are available at: http://www.census.gov/ces/dataproducts/bds/data_firm.html
[5] The data are available at: http://research.stlouisfed.org/fred2/series/USAWFPNA#.

3.4.2 Genetic distance

The genetic distance variable represents the absolute weighted genetic distance between countries *i* and *j*, representing the genealogical relatedness of two randomly chosen individuals, one from each country (Spolaore and Wacziarg 2009: 485). Higher values are associated with larger differences (Spolaore and Wacziarg 2009).[6] This information was also used to calculate pairwise differences for the alternative specification.

3.4.3 Control variables

We draw on a comprehensive set of controls variables. These could be broadly classified into geographical factors (geodesic distance, latitudinal distance and longitudinal distance), micro-geographical factors (contiguity, landlocked, island and elevation), continent effects, common history variables (linguistic distance, religion distance, colony, common colonizer, current colonial and colonial relationship) and other controls (cultural, governance, institutional and economic factors). We describe the variables in Appendix Table 4.

3.5 Results

In the Appendix, Tables 4 and 5 present the descriptive statistics and correlations, respectively. The list of countries in the sample is included in Table 9. The mean of the start-up rate is 1.710 new firms per thousand working-age population and the mean of genetic distance is 0.10, which is in line with the value in Spolaore and Wacziarg (2009) (in their 9316 pairs of observations the mean was 0.11). Our genetic distance bears a positive correlation of 0.10 ($p < 0.001$) with the start-up rate.

Table 1 presents the unilateral regressions to the technological frontier, the USA. To facilitate the interpretations of the effects sizes, in the last row we list the standardized beta coefficient of the genetic distance variable. The standardized beta refers to how the standard deviation of firm density changes for each standard deviation change in genetic distance. Model 1 includes genetic distance without any controls. The coefficient of genetic distance is positive and significant ($p < 0.01$). Model 2 includes genetic distance and the average number of new firms for the countries in the pair. In Model 3 we also add the control variables related to geographical factors, while Model 4 includes all of the remaining control variables. The coefficient of genetic distance retains its significance in all four models. The results show that genetic distance is positively associated with the difference in the level of new firm entry from the technological frontier. However, when including the controls, the magnitude of effects decreases. For model 1 using standardized beta coefficient in the last row, for 1 s.d. increase in genetic distance there is $e^{0.210} = 1.234$ increase in startups per 1000 working age individuals. Using similar calculations for all the remaining models, for 1 s.d. increase in genetic distance, the difference in the start-up rate to the technological frontier ranges from 1.047 to 1.234 firms per 1000 working-age population (those ages 15–64).

[6] The data are available at: http://sites.tufts.edu/enricospolaore/.

Table 1 Two-way clustered standard errors unilateral regressions to the technological frontier US

Variables	(1)	(2)	(3)	(4)
		DV = Difference in Start-up rate relative to the US (number of new limited liability firms per 1000 working-age individuals relative to the US)		
Genetic distance	4.2992***	3.4573***	0.9471**	1.0446***
	(20.4190)	(15.5973)	(2.2811)	(4.6599)
Pairwise average number firms		−0.7302***	−1.0567***	−0.6062***
		(−17.4865)	(−38.4014)	(−28.0802)
Geodesic distance			−0.0353***	−0.0542***
			(−11.8465)	(−5.3433)
Latitudinal distance			0.0077***	0.0108***
			(7.4590)	(17.7169)
Longitudinal distance			0.0048***	0.0029***
			(16.2835)	(3.8713)
Contiguous			−0.2704***	−0.1493***
			(−4.9613)	(−3.6132)
Landlocked			0.3447***	0.4137***
			(2.9840)	(2.9071)
Island			0.4992***	0.4719***
			(9.9638)	(11.1794)
Linguistic distance				1.1503
				(0.8618)
Religion distance				0.3945***
				(10.4279)
Constant	0.8877***	1.1573***	0.9691***	−0.3419
	(12.0861)	(12.9470)	(13.3095)	(−0.2675)
Observations (pairwise between USA and country *i*)	102	102	102	100
R-squared	0.0443	0.0763	0.1817	0.1967
Standardized beta of Genetic distance	0.210	0.169	0.046	0.051

The dependent variable is start-up rate. The independent variables include genetic distance and a set of controls. Refer to the text for detailed variable definitions. The STATA code was obtained from Professor Mitchell Petersen's website and is: dependent_variable independent_variables, fcluster(country i) tcluster(country j)

Table 9 (Appendix) presents the list of countries

t-statistics in parentheses. † $p < 0.10$, * $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$

Table 2 presents the results of the two-way clustered standard errors bilateral regressions. Model 1 includes genetic distance and we add control variables in model 2 (just the average number of new firms for the countries in the pair) to model 8 (all controls). The coefficient of genetic distance is positive and significant, indicating that genetic distance is positively associated with cross-country differences in the start-up rate, except for models 3 and 6, with both using geographic controls; perhaps significant collinearity among the closely correlated geographic characteristics related to

Table 2 Two-way clustered standard errors bilateral regressions

	(1) Genetic distance	(2) With average number of firms (log) for the countries	(3) With geographical controls	(4) With religion linguistics and colony	(5) With governance indicators	(6) With geographical controls, religion linguistics and colony and governance indicators	(7) With economic development and law system	(8) Full model
Start-up rate between two pairs of countries *i* and *j* (number of new limited liability firms per 1000 working-age individuals between two pairs of countries *i* and *j*)								
Genetic distance	1.8734**	2.2097**	0.8941	2.1649**	1.6918**	1.0704	2.4634**	1.3244*
	(2.1753)	(2.4104)	(0.9840)	(2.3015)	(2.3341)	(1.1038)	(2.4921)	(1.7884)
Pairwise average number firms		0.4145***	0.5968***	0.5058***	0.3958**	0.6628***	0.9001***	1.0411***
		(2.6236)	(3.8854)	(3.2021)	(2.5654)	(4.0732)	(4.6765)	(6.0658)
Geodesic distance			−0.0055			−0.0013		0.0171
			(−0.2957)			(−0.0575)		(0.9788)
Latitudinal distance			−0.0047**			−0.0037		−0.0018
			(−2.0927)			(−1.5792)		(−0.9396)
Longitudinal distance			0.0006			−0.0004		0.0008
			(0.4282)			(−0.2110)		(0.5646)
Contiguous			−0.2920**			−0.1586		0.1465
			(−2.1848)			(−1.2806)		(1.0198)
Landlocked			0.4960**			0.5053**		0.0437
			(2.1354)			(2.1713)		(0.4418)
Island			0.3556**			0.2978*		0.2547**
			(2.4494)			(1.8658)		(2.5683)
Elevation			−0.0002			−0.0002		−0.0001
			(−1.4186)			(−1.5456)		(−1.4838)
North America Both			−0.0137			0.0315		0.0232
			(−0.3077)			(0.7955)		(0.7245)

Table 2 (continued)

	(1) Genetic distance	(2) With average number of firms (log) for the countries	(3) With geographical controls	(4) With religion linguistics and colony	(5) With governance indicators	(6) With geographical controls, religion linguistics and colony and governance indicators	(7) With economic development and law system	(8) Full model
South America Both			−0.1834***			−0.0553		−0.2081
			(−3.9817)			(−0.3840)		(−1.5357)
Asia Both			0.0123			0.0192		0.1306***
			(0.1862)			(0.2879)		(4.0029)
Africa Both			0.0975**			0.1155**		0.1053
			(2.5038)			(2.5633)		(.)
Europe Both			−0.1939***			−0.1590***		0.0094
			(−4.3131)			(−3.6996)		(0.4998)
Pacific Both			0.0829**			−0.2160**		−0.1502***
			(2.5756)			(−2.5591)		(−3.3939)
Lingustic Distance				0.0013		0.3850		−0.0743
				(0.0015)		(0.3658)		(−0.1469)
Religion distance				0.3918***		0.3116***		0.0346
				(3.5065)		(2.9454)		(0.6215)
Colony				−0.4387**		−0.2425		−0.1745
				(−2.2237)		(−1.3866)		(−1.1129)
Common Colonizer				0.0987		−0.0909		−0.0161
				(0.7229)		(−0.5587)		(−0.1449)
Colonial relationship				0.6095***		0.2796*		0.0969
				(3.3657)		(1.8192)		(0.6331)
Rule of law					−0.4895**			0.1804

Table 2 (continued)

	(1) Genetic distance	(2) With average number of firms (log) for the countries	(3) With geographical controls	(4) With religion linguistics and colony	(5) With governance indicators	(6) With geographical controls, religion linguistics and colony and governance indicators	(7) With economic development and law system	(8) Full model
					(−2.5548)			(1.3010)
Government effectiveness					0.1908			−0.3103*
					(0.6726)			(−1.9529)
Control of corruption					−0.1073			−0.0343
					(−0.6804)			(−0.3390)
Regulatory quality					0.8295***			0.5901***
					(3.4595)			(4.4147)
Political stability					0.4625***			0.1874**
					(4.0755)			(2.3142)
Accountability					−0.0567			−0.0231
					(−0.7562)			(−0.3537)
Economic development							0.0000	−0.0000**
							(0.6291)	(−2.5386)
Efficiency of bankruptcy							0.0089	−0.0211
							(0.4704)	(−1.0319)
Legal origin UK Both							0.0026	−0.1786
							(0.0168)	(−1.3924)
Legal origin French Both							0.1212	0.1332**
							(1.5200)	(2.2337)
Legal origin Socialist Both							−0.1690	−0.0945
							(−1.4075)	(−0.6417)

Table 2 (continued)

	(1) Genetic distance	(2) With average number of firms (log) for the countries	(3) With geographical controls	(4) With religion linguistics and colony	(5) With governance indicators	(6) With geographical controls, religion linguistics and colony and governance indicators	(7) With economic development and law system	(8) Full model
Legal origin German Both							−0.1766	0.0223
							(−1.1073)	(0.3149)
Legal origin Scandinavian Both							−1.1615***	−0.7465***
							(−10.3366)	(−4.0960)
Constant	1.5311***	1.3923***	1.4905***	0.9360	0.5980***	0.7386	0.6175***	0.1700
	(12.3538)	(9.5281)	(6.7701)	(1.1190)	(4.6433)	(0.7591)	(4.8092)	(0.3535)
Observations (pairwise between country I and country j)	5253	5253	5253	5050	5253	5050	2016	2016
R-squared	0.0091	0.0267	0.1037	0.0617	0.2145	0.1215	0.1569	0.3734
Standardized beta of Genetic distance	0.096	0.113	0.046	0.111	0.086	0.055	0.126	0.068

Notes. The dependent variable is start-up rate. The independent variables include genetic distance and a set of controls. Refer to the text for detailed variable definitions. The STATA code was obtained from Professor Mitchell Petersen's website and is: dependent_variable independent_variables, fcluster(country i) tcluster(country j)

t-statistics in parentheses

When running model 1 and 2 just for the smaller sample as in model 3, the results are similar (see Appendix Table 6)

Table 9 (Appendix) presents the list of countries.

† $p < 0.10$, * $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$

geographic location, continent, contiguity with other countries may also influence the extent of genetic distance. In the full model, where we include these and other controls, the coefficient of genetic distance is significant. Based on the estimates, for 1 s.d. increase in genetic distance, the difference in the start-up rate ranges from 1.047 to 1.134 firms per 1000 working-age population (those ages 15–64). For example, for model 8 using the standardized beta coefficient, 1.070 is calculated as $e^{0.068} = 1.070$.

3.6 Robustness checks and extensions

3.6.1 Alternate measure of genetic distance

Although the findings are robust to controlling for an extensive set of control variables, we first test if the findings are robust to an alternative operationalization of genetic distance – N_{ei} genetic distance proposed by Cavalli-Sforza et al. (1994). The estimates from this alternative measure of genetic distance were consistent with the results (Table 3).

3.6.2 Casewise deletion

In our analysis in Tables 2 and 3, the sample size varies across models because we do not use the casewise deletion restriction across all the models. To check that the findings are not an artifact of such a restriction, in Tables 6, 7 and 8 (Appendix), we use casewise deletion across all models and include the Hofstede cultural dimensions. The findings are consistent with the main inferences.

3.6.3 Alternate cultural distance measures

In Table 7 in the Appendix, we used the Hofstede's cultural distance measure. As an additional analysis, based on cultural dimensions in the World Value Survey (WVS), we used the composite measure of cultural distance (model 1, Table 8) and the five individual dimensions of cultural distance (Perceptions of Life; Work-Family; Politics and Society; Religion and Morale; and National Identity; model 2, Table 8). The inferences based on casewise deletion were consistent with the main inferences.

4 Discussion

Based on recent developments in measuring the between-country genetic distance, we tested for its association with differences in new firm entry. We ran unilateral (Table 1) and bilateral (Tables 2 and 3) regressions with country clustering to draw robust inferences. We controlled for a significant set of variables, in addition to those controlled by Spolaore and Wacziarg (2009), to limit the effects of alternate explanations for the identified relationships. The findings, after controlling for factors ranging from cultural factors to historical events such as colonization and from institutional factors to religion, indicated that genetic distance is associated with cross-country differences in the rates of entrepreneurship. The effects sizes are small and, depending

Table 3 Two-way clustered standard errors bilateral regressions – N_{ei} genetic distance

Variables	(1)	(2)	(3)
	DV = Difference in Start-up rate between two pairs of countries i and j (number of new limited liability firms per 1000 working-age individuals between two pairs of countries i and j)		
Genetic distance (N_{ei})	8.2618*	9.7539*	7.0359**
	(1.7705)	(1.9389)	(1.9886)
Pairwise average number firms		0.3983**	1.0444***
		(2.4800)	(6.0747)
Geodesic distance			0.0176
			(1.0212)
Latitudinal distance			−0.0019
			(−1.0060)
Longitudinal distance			0.0007
			(0.5368)
Contiguous			0.1504
			(1.0469)
Landlocked			0.0483
			(0.4988)
Island			0.2547**
			(2.5416)
Elevation			−0.0001
			(−1.4799)
North America Both			0.1673
			(0.7513)
South America Both			−0.4190
			(−1.5291)
Asia Both			0.3842***
			(4.0192)
Africa Both			0.3470
			(0.0000)
Europe Both			0.0444
			(0.4703)
Pacific Both			−0.9025***
			(−3.3669)
Lingustic Distance			0.0110
			(0.0223)
Religion distance			0.0365
			(0.6488)
Colony			−0.1742
			(−1.1099)
Common Colonizer			−0.0062
			(−0.0578)

Table 3 (continued)

Variables	(1)	(2)	(3)
	DV = Difference in Start-up rate between two pairs of countries *i* and *j* (number of new limited liability firms per 1000 working-age individuals between two pairs of countries *i* and *j*)		
Colonial relationship			0.0946
			(0.6110)
Rule of law			0.1844
			(1.3198)
Government effectiveness			−0.3058*
			(−1.9367)
Control of corruption			−0.0370
			(−0.3617)
Regulatory quality			0.5908***
			(4.4028)
Political stability			0.1888**
			(2.3249)
Accountability			−0.0258
			(−0.3941)
Economic development			−0.0000***
			(−2.6714)
Efficiency of bankrupcy			−0.0227
			(−1.1109)
Legal origin UK Both			−0.1837
			(−1.4272)
Legal origin French Both			0.1378**
			(2.3230)
Legal origin Socialist Both			−0.0961
			(−0.6521)
Legal origin German Both			0.0125
			(0.2004)
Legal origin Scandinavian Both			−0.7355***
			(−3.9948)
Constant	1.5797***	1.4536***	0.1015
	(12.1795)	(9.4568)	(0.2119)
Observations (pairwise between country I and country j)	5253	5253	2016
R-squared	0.0056	0.0219	0.3729
Standardized beta of Genetic distance	0.075	0.088	0.064

Notes. The dependent variable is start-up rate. The independent variables include genetic distance and a set of controls. Refer to the text for detailed variable definitions. The Stata code was obtained from Professor Mitchell Petersen's website and is: dependent_variable independent_variables, fcluster(country i) tcluster(country j)

Table 9 (Appendix) presents the list of countries.

t-statistics in parentheses. † $p < 0.10$, * $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$

on the specification, 1 s.d. increase in genetic distance, the difference in the start-up rate ranges from 1.047 to 1.134 firms per 1000 working-age population (those ages 15–64).

Our study also extends the biosocial model of entrepreneurship (Shane and Nicolaou 2015; White et al. 2007) to genetic distance influencing cross-country differences in entrepreneurship. Although research has confirmed a larger role of 'nurture' relative to 'nature' in influencing the choice to become an entrepreneur, biology is an important, though not a deterministic, factor in entrepreneurship (Nicolaou and Shane 2014).

Factors such as genetic distance could be discounted as 'something one cannot control, so why bother.' In fact, the findings contribute to these rebukes in the following ways. First, as our results can be viewed as evidence of continued long-term effects of barriers across different countries due to migration patterns and institutional differences, significant reductions in entrepreneurial disparities across nations can be achieved by implementing policies that reduce such barriers, such as encouraging cross-country trade, exchanges, the diffusion of entrepreneurial ideas, and openness (Spolaore and Wacziarg 2009). The identified relationship suggests the value of encouraging the diffusion of ideas across countries, which can overcome 'resistance' from genetic distance. Furthermore, without knowing the relative effects of factors such as genetic distance on entrepreneurial activity, estimates of alternate factors driving entrepreneurship would be conflated. Related to studies on twins, as genes explain a significant portion of the likelihood of entrepreneurship, non-inclusion of such factors could lead to conflated estimates of its correlates such as personality. In a similar vein, at the least, controlling for genetic distance in research on country-wise differences in entrepreneurship rates may provide more reliable inferences.

The association between genetic distance and differences in rates of entrepreneurship complements recent work on the association between biology and entrepreneurship. Using samples of identical and fraternal twins, research has shown that genes influence the tendency to become entrepreneurs and recognize entrepreneurial opportunities (Nicolaou et al. 2008; Nicolaou et al. 2009), while a related stream of research has also examined the role of genetically influenced hormones in entrepreneurship (Unger et al. 2015; White et al. 2006). Thus, interest in biology and entrepreneurship has increased significantly in recent years.

The findings also open up new research questions on cross-country differences that explain the differences in the levels of entrepreneurship between countries. If genetic distance drives differences in human behavior, such influences should converge and coalesce to develop distinct cultures and institutions. Path-dependent migratory patterns would lead to the development of complex country related differences. While work on population ecology has called into question the value of entrepreneurial agency, the findings indicate that, while genetic distance is a significant explanatory factor, a significant amount of variance also remains unexplained. This indicates that genetic distance is an important but *not* a definitive explanation of cross-country differences in entrepreneurship.

Future research may identify mediators in the relationship between genetic distance and differences in entrepreneurial activity. Research by personality psychologists on country-level personality traits would be useful in this respect (Schmitt et al. 2007). Explaining the relationship between genetic differences and cultural influence, Bleidorn et al. (2013) find that normative life transitions to adult roles explain personality outcomes (or, social-investment theory), thereby indirectly not finding support for genetic factors explaining personality differences. Future research could assess whether individual-level entrepreneurial personality is based on social investment theory or genetic factors. While we focus on between-country differences in new firm establishments, Obschonka et al. (2013) found that entrepreneurial personality is regionally clustered within the US, Germany, and the UK. Based on these findings, variations in the genetic distance within a country could explain the clustering of entrepreneurial activities. Accordingly, future work can examine whether country-level personality traits mediate the genetic distance-entrepreneurship relationship. Proposing an Entrepreneurial Personality System, Obschonka and Stuetzer (2017) found support for "gravity effect of an intraindividual entrepreneurial Big Five profile on the more malleable psychological factors" (page 203). This complex confluence of genetic, cultural, and individual factors explaining self-employment outcomes is indeed an important area for future research.

4.1 Limitations

Our study has several limitations. First, we explicitly acknowledge that the findings may be confounded by geographic factors. In other words, without identifying an instrumental variable(s) that separates the effects of geographic factors on genetic distance we do not know if genetic distance or geographic factors are influencing the results. In Table 2 (models 3 and 6), when including geographic controls along with genetic distance, the effect of genetic distance is non-significant, perhaps due to partial determination of genetic distance by geographic controls (Geodesic distance, Latitudinal distance, Longitudinal distance, Contiguous, Landlocked, Island, Elevation, North America Both, South America Both, Asia Both, Africa Both, Europe Both, Pacific Both). Future research limiting the collinearity between geographic factors and genetic distance may provide more reliable inferences on the influence of genetic distance on the differences in entrepreneurship rates between countries.

Second, we are unable to explain the macro- or meso-level relationships among the cultural and institutional factors, albeit we control for these factors. These developments are a result of complex historical and social process, the variance of which may not be fully captured in an almost steady state time series available in country-level research. The relationships that are in stasis for a long period of time cannot be fully explained by variance based methodologies, but instead by steady-state econometrics with long-term data that are seldom available for the phenomenon we study here.

Furthermore, the new firm entry data for the US are not available from the World Bank, and therefore we collated this measure for the US from the US Census Bureau. While there is no reason to doubt the data quality from the US Census Bureau, we believe that more uniform data collation from the reporting agencies in different countries may reduce plausible idiosyncrasies in collation procedures across countries.

Third, while the inferences relate to the genetic distance among populations, the study does not explain the regional differences within countries. For example, while genetic distances within continents are likely to be lower than between continents, the within-country variation and the resulting differences in entrepreneurship within a country require further elaboration.

Fourth, a significant amount of R-square remains unexplained, indirectly cautioning that genetic distance *and* the included controls still explain a relatively small amount of variance in entrepreneurship related differences between countries and that micro- and meso-level effects could play a larger role in explaining systematic variations in entrepreneurship. Although our level of analysis is at the country-level, within-country differences in the genetic distance could not be fully ruled out. Such differences could explain differences in entrepreneurial activity within a country. Strong regional inequalities in entrepreneurial activities pointed out in recent literature (Bosma and Schutjens 2011), focus on sub-regional genetic distance and differences in entrepreneurial activities could add further insights on the role of genetic distance in explaining differences in entrepreneurial outcomes.

Fifth, the study focused on 103 countries where data based on pair-wise deletion were available. However, the findings could not be generalized to countries where such data are not available.

Sixth, increasing migration from developing to developed countries in the last two decades could attenuate the effects of genetic distance on between-country differences in startup rates. With an increasing number of immigrants selecting into high-tech entrepreneurship (Hart and Acs 2011; Saxenian 2002), the effects of genetic distance on differences in entrepreneurial activity between developed-developing country pairs could have upward bias wherein entrepreneurial human capital gaps are increased due to higher migration from developing to developed countries. Migration between developed countries may lower the differences in human capital between countries (e.g., migration among EU nations), thereby reducing the differences in entrepreneurial activity between developed-developed country pairs. Conversely, migration may also increase the differences in human capital between countries (e.g., brain drain from developing countries), thereby increasing gap in entrepreneurial activity between developed-developing country pairs. Finally, a complex combination of environmental, cultural, economic, and social factors at multiple levels come into confluence to explain entrepreneurial activity. The individual, country, and between country levels of interactions result in endogenous processes that are difficult to parse out theoretically and empirically. We therefore caution that the inferences in this study are subject to omitted variable bias and the influence of the unobservables, operating at multiple levels of analysis, in the error term of the regression is not fully taken into account.

In closing, the findings must be interpreted with caution. Nevertheless, after including a variety of control variables and specifying alternate regressions, the role of genetic distance in explaining differences in the new firm entry should not be discounted either. The aim of the study is not to suggest that countries with longer or shorter genetic distances are better positioned in their entrepreneurial capabilities. Indeed, genetic distance is a result of the long-term migration process. Nevertheless, controlling for a large number of country-specific effects, the genetic distance may be an important structural predictor in explaining differences in new firm entry between countries.

Funding This study was funded by from Advance Research Center (supported through UID/SOC/04521/2019 project by FCT - Fundação para a Ciência e a Tecnologia, Portugal).

Compliance with ethical standards

Conflict of interest The authors declare that they have no conflict of interest.

Appendix

Table 4 Descriptive statistics (all the variables are for each country-pair)

	Measured at country-pair level	Definition	N	Mean	St. Dev.	Min	Max
1	Difference in Start-up rate	The difference in number newly registered companies with limited liability per 1000 working-age people (those ages 15–64).	5253	1.71	1.37	0	8.40
2	Genetic Distance	The absolute weighted genetic distance between countries *i* and *j*, representing the genealogical relatedness of two randomly chosen individuals, one from each country (Spolaore and Wacziarg 2009: 485).	5253	0.10	0.07	0	0.30
3	Pairwise average number firms	The (log) average number of firms for the countries that are compared pairwise.	5253	0.26	0.44	−1.76	1.29
4	Geodesic distance	The difference in the distance (in kilometers) between the major cities of countries *i* and *j*, divided by 1000 (Mayer and Zignago 2011).[a]	5253	7.30	4.38	0.11	19.54
5	Latitudinal distance	The absolute difference between the latitudes of countries *i* and *j* (Mayer and Zignago 2011).	5253	30.77	23.3	0	108.43
6	Longitudinal distance	The absolute difference between the longitudes of countries *i* and *j* (Mayer and Zignago 2011).	5253	65.04	53.93	0.02	350.02
7	Contiguous	Dummy variable that equals 1 if two countries in a pair are contiguous, that is share a common boundary (Mayer and Zignago 2011).	5253	0.02	0.14	0	1
8	Landlocked	Dummy variable that equals 1 for both landlocked countries (Mayer and Zignago 2011).	5253	0.41	0.49	0	1
9	Island	Dummy variable that equals 1 if either of the countries in a pair is an island. The data were obtained from the World Atlas.	5253	0.32	0.47	0	1
10	Elevation	The absolute difference between the average elevation (meters above sea level) of countries *i* and *j*. The data were obtained from Giuliano et al. (2006).	5253	625.7	642.92	0.02	3187.92
11		Dummy variables that equal 1 if two countries in a pair were on the	5253	0.01	0.09	0	1

Table 4 (continued)

	Measured at country-pair level	Definition	N	Mean	St. Dev.	Min	Max
	North America Both	same continent (Africa, Asia, Europe, North America, South America and the Pacific). The data was obtained from Mayer and Zignago (2011).					
12	South America Both		5253	0	0.06	0	1
13	Asia Both		5253	0.07	0.25	0	1
14	Africa Both		5253	0.05	0.21	0	1
15	Europe Both		5253	0.09	0.29	0	1
16	Pacific Both		5253	0	0.03	0	1
17	Linguistic distance	Building on Spolaore and Wacziarg (2009) and drawing from Fearon (2003) is equal to: $LD = \sqrt{\frac{(15-\#\text{common nodes})}{15}}$ After counting the common linguistic nodes between subgroups in pairs of countries, the country #common nodes were calculated by taking the average number of common linguistic nodes, weighed by the subgroup population size. Linguistic distance is, thus, the weighted index of linguistic similarity between countries.	5050	0.98	0.05	0.57	1
18	Religion distance	The sum of the average absolute difference in the percentage of each religion, between countries i and j. The main religions were: Christian, Muslim, Buddhist, Hindu, Jewish, and No Religion. The data were obtained from CIA World Factbook.[b]	5253	1.04	0.67	0	2
19	Colony	Dummy variable that equals 1 if the countries in each pair have ever had a colonial history (Mayer and Zignago 2011).	5253	0.02	0.13	0	1
20	Common colonizer	Dummy variable that equals 1 if the countries in each pair had a common colonizer after 1945 (Mayer and Zignago 2011).	5253	0.07	0.26	0	1
21	Colonial relationship	Dummy variable that equals 1 if the countries in each pair have had a colonial relationship after 1945 (Mayer and Zignago 2011).	5253	0.01	0.1	0	1
22	Uncertainty avoidance	The absolute difference in uncertainty avoidance between	1035	29.71	21.39	0	104

Table 4 (continued)

	Measured at country-pair level	Definition	N	Mean	St. Dev.	Min	Max
		countries *i* and *j*. Uncertainty avoidance refers to the degree to which individuals feel uncomfortable with uncertainty and ambiguity (Hofstede 2011).					
23	Individualism vs Collectivism	The absolute difference in individualism between countries *i* and *j*. Individualism relates to the degree to which a country emphasizes individual ("I") or collective achievement ("We") and interpersonal relationships (Hofstede 2011).	1035	28.38	19.61	0	85
24	Power distance	The absolute difference in power distance between countries *i* and *j*. It denotes the degree of equality or inequality between people in a country (Hofstede 2011).	1035	24.97	17.76	0	84
25	Masculinity vs femininity	The absolute difference in masculinity between countries *i* and *j*. It denotes the degree that a country reinforces the traditional masculine role models of achievement, control, and power (Hofstede 2011).	1035	21.62	16.77	0	90
26	Rule of Law	The perception of the extent to which residents have confidence in and abide by the rules of society, and in particular the quality of contract enforcement, property rights, the police, and the courts, as well as the extent of crime and violence	5253	1.18	0.84	0	3.87
27	Control of corruption	It reflects the perceptions of the extent to which public power is exercised for private gain, including both petty and grand forms of corruption, as well as the "capture" of the state by elites and private interests.	5253	1.23	0.92	0	4.06
28	Government effectiveness	It represents the perceptions of the quality of public services, the quality of the civil service and the degree of its independence from political pressures, the quality of policy formulation and implementation, and the credibility of the government's commitment to such policies.	5253	1.13	0.81	0	3.80
29	Voice and accountability	It reflects perceptions of the extent to which a country's citizens are able to participate in selecting their government, as well as	5253	1.08	0.77	0	3.67

Table 4 (continued)

	Measured at country-pair level	Definition	N	Mean	St. Dev.	Min	Max
		freedom of expression, freedom of association, and a free media.					
30	Political stability	It reflects perceptions of the likelihood that the government will be destabilized or overthrown by unconstitutional or violent means, including politically-motivated violence and terrorism.	5253	1.11	0.84	0	4.11
31	Bankruptcy costs	The efficiency of bankruptcy law, using a scale from 0 to 6, where higher scores indicate higher compliance (Mihet 2013). The data were obtained from the World Economic Forum Global Competitiveness Report (2005).	2016	1.99	1.73	0	6.70
32	Economic develop-ment	The average of the 2008 and 2009 GDP per capita, which is the gross domestic product in current U.S. dollars divided by mid-year population. The data were avail-able from Worldbank.[c]	5253	21,365.86	22,002.31	2.47	105,020.10
33	Legal origin UK Both	Dummy variables that equal 1 if the two countries in a pair have the same legal origin (La Porta et al. 1999).[d]	5253	0.09	0.29	0	1
34	Legal origin French Both		5253	0.16	0.36	0	1
35	Legal origin Socialist Both		5253	0.04	0.20	0	1
36	Legal origin German Both		5253	0	0.03	0	1
37	Legal origin Scandina-vian Both		5253	0	0.04	0	1

This table reports the summary statistics and the correlation between the variables of the study

[a] The data are available from Centre d'Etudes Prospectives et d'Informations Internationales (CEPII) at http://www.cepii.fr/CEPII/en/bdd_modele/download.asp?id=6

[b] The data is available at: http://gsociology.icaap.org/data/religion.xls and https://www.cia.gov/library/publications/the-world-factbook/fields/2122.html

[c] The data are available at: http://data.worldbank.org/indicator/NY.GDP.PCAP.CD?page=1

[d] The data are available at: http://faculty.tuck.dartmouth.edu/rafael-laporta/research-publications

Table 5 Pairwise correlations

		1	2	3	4	5	6	7	8	9	10	11	12	13
1	Start-up rate	1												
2	Genetic Distance	0.10***	1											
3	Pairwise average number firms	0.12***	−0.13***	1										
4	Geodesic distance	0.09***	0.33***	0.03**	1									
5	Latitudinal distance	0.04***	0.23***	0.23***	0.49***	1								
6	Longitudinal distance	0.09***	0.15***	0.01	0.85***	0.13***	1							
7	Contiguous	−0.07***	−0.13***	−0.03**	−0.22***	−0.17***	−0.16***	1						
8	Landlocked	0.11***	0.17***	0.20***	−0.10***	−0.01	−0.14***	0.01	1					
9	Island	0.14***	0.09***	0.18***	0.34***	0.09***	0.40***	−0.09***	−0.22***	1				
10	Elevation	0.00	0.09***	−0.13***	0.02	0.00	0.01	−0.04**	0.46***	−0.10***	1			
11	North America Both	−0.00	0.02	−0.02*	−0.11***	−0.08***	−0.09***	0.06***	−0.08***	0.07***	−0.03**	1		
12	South America Both	−0.03*	−0.02	−0.01	−0.06***	−0.03†	−0.06***	0.18***	−0.02	−0.04***	0.01	−0.01	1	
13	Asia Both	0.00	−0.15***	−0.14***	−0.22***	−0.17***	−0.16***	0.08***	0.08***	0.01	0.15***	−0.02†	−0.02	1
14	Africa Both	0.06***	−0.13***	−0.11***	−0.18***	−0.08***	−0.17***	0.08***	0.10***	−0.11***	−0.01	−0.02	−0.01	−0.06***
15	Europe Both	−0.18***	−0.31***	0.19***	−0.43***	−0.31***	−0.31***	0.16***	−0.08***	−0.06***	−0.14***	−0.03*	−0.02	−0.09***
16	Pacific Both	0.03*	−0.01	0.04***	−0.03*	−0.01	0.07***	0.00	−0.03*	0.05***	−0.02	0.00	0	−0.01
17	Linguistic distance	0.08***	0.25***	−0.02	0.16***	0.05***	0.14***	−0.22***	0.08***	0.09***	0.04*	−0.16***	−0.30***	0.06***
18	Religion distance	0.14***	−0.03*	−0.19***	0.06***	−0.11***	0.09***	−0.10***	0.06***	−0.01	0.10***	−0.10**	−0.04***	0.07***
19	Colony	−0.01	−0.04**	0.07***	−0.05***	0.00	−0.05***	0.18***	−0.05***	0.03†	−0.04*	0.00	−0.01	0.01
20	Common colonizer	0.03*	0.05***	−0.04***	−0.06***	−0.12***	−0.02†	0.07***	0.01	0.06***	0.00	−0.02	−0.02	0.1***
21	Colonial relationship	0.03†	0.00	0.07***	−0.04*	0.01	−0.04***	0.08***	−0.03†	0.05***	−0.03†	−0.01	−0.01	0.02
22	Uncertainty avoidance	−0.01	0.15***	0.12***	0.04	−0.04	0.04	−0.10**	−0.11***	0.23***	0.00	0.06†	−0.09**	0.02
23	Individualism vs collectivism	0.15***	0.11***	0.10***	0.14***	0.17***	0.17***	−0.12***	−0.03	0.05	−0.01	0.08*	−0.06†	−0.14***

Table 5 (continued)

		1	2	3	4	5	6	7	8	9	10	11	12	13
24	Power distance	0.08*	0.00	0.03	0.03	0.05	0.05	−0.07*	0.11***	0.03	−0.09**	0.03	−0.09**	−0.01
25	Masculinity vs femininity	−0.11***	−0.02	0.13***	−0.11***	−0.03	−0.07*	−0.06†	0.20***	−0.04	0.00	0.01	−0.02	−0.06†
26	Rule of Law	0.26***	−0.02	0.21***	0.00	0.09***	0.02	−0.09***	0.03*	0.06***	0.07***	0.01	−0.02	−0.02
27	Control of corruption	0.24***	−0.05***	0.3***	0.02	0.10***	0.03*	−0.09***	0.00	0.09***	0.01	−0.01	−0.02†	−0.01
28	Government effectiveness	0.30***	0.03*	0.21***	0.01	0.09***	0.01	−0.09***	0.01	0.07***	0.02†	0.00	−0.04*	−0.02
29	Voice and accountability	0.19***	−0.03*	0.11***	−0.03*	0.06***	−0.03*	−0.09***	0.14***	−0.04**	0.08***	−0.03*	−0.04**	−0.11***
30	Political stability	0.30***	−0.01	−0.08***	0.02	−0.04***	0.05***	−0.08***	0.04**	0.02	0.10***	−0.02†	−0.01	0.08***
31	Bankruptcy costs	−0.02	−0.15***	−0.04*	−0.07***	−0.08***	−0.03	−0.07**	0.12***	−0.08***	0.13***	−0.03	−0.03	0.11***
32	Economic development	0.08***	−0.11***	0.31***	−0.07***	0.10***	−0.07***	−0.07***	0.02	0.01	−0.05***	−0.02	−0.05***	−0.05***
33	Legal origin UK Both	0.08***	0.13***	0.01	0.10***	−0.01	0.12***	0.02	−0.01	0.15***	0.04**	0.01	−0.02	0.06***
34	Legal origin French Both	0.04*	−0.02†	−0.13***	−0.01	−0.06***	−0.03*	0.04*	−0.09***	−0.04**	−0.08***	0.05**	0.15***	−0.07***
35	Legal origin Socialist Both	−0.10***	−0.14***	0.04***	−0.22***	−0.16***	−0.16***	0.17***	0.09***	−0.14***	0.07***	−0.02	−0.01	0.05**
36	Legal origin German Both	−0.02†	−0.02	−0.01	−0.02	−0.03*	0.00	0.11***	0.03*	0.01	0.00	0.00	0	−0.01
37	Legal origin Scandinavian Both	−0.04**	−0.04**	0.05***	−0.06***	−0.05***	−0.04*	0.09***	−0.04*	0.01	−0.02*	0.00	0	−0.01

	14	15	16	17	18	19	20	21	22	23	24	25	26
1													
2													
3													
4													
5													
6													
7													

Table 5 (continued)

	14	15	16	17	18	19	20	21	22	23	24	25	26
8													
9													
10													
11													
12													
13													
14	1												
15	−0.07***	1											
16	−0.01	−0.01	1										
17	−0.22***	−0.07***	−0.01	1									
18	0	−0.23***	−0.01	0.26***	1								
19	−0.02	0.02	0	−0.10***	−0.05***	1							
20	0.21***	−0.07***	0.01	−0.03*	0.05**	−0.04**	1						
21	−0.01	−0.02†	0	0	0	0.74***	−0.03*	1					
22	−0.01	0.03	−0.04	0.24***	−0.03	−0.10**	0.07*	−0.04	1				
23	−0.01	−0.23***	−0.03	0.11***	0.03	0.02	−0.07*	0.13	0.05	1			
24	0.01	−0.11**	−0.02	0.16***	0.02	−0.05	0	0.03	0.09*	0.3***	1		
25	−0.02	0.18***	−0.03	0.05	−0.05	−0.09***	−0.14***	−0.06†	0.03	0.02	0.07*	1	
26	−0.14***	−0.07***	0	0.09***	0.04**	0	−0.10***	0.03†	0.07*	0.37***	0.36***	−0.02	1
27	−0.13***	−0.01	0.02	0.10***	0.04**	−0.01	−0.08***	0.02	0.14***	0.29***	0.34***	0.02	0.89***
28	−0.12***	−0.06***	0.02	0.11***	0.06***	0	−0.05***	0.03*	0.13***	0.35***	0.32***	−0.01	0.89***
29	−0.08***	−0.16***	−0.01	0.16***	0.24***	−0.01	−0.11***	0.03†	−0.01	0.38***	0.34***	−0.03	0.63***
30	−0.05***	−0.17***	−0.02	0.09***	0.22***	−0.04**	0.02	−0.02	−0.06†	0.24***	0.22***	−0.09**	0.51***
31	−0.07**	−0.04†	−0.02	0.12***	0.04	−0.04†	0.02	0.00	0.26***	0.35***	0.38***	−0.02	0.40***
32	−0.19***	0.10***	0.01	0.12***	−0.05***	0.02	−0.10***	0.03*	0.07*	0.36***	0.31***	0.20***	0.60***

Table 5 (continued)

	14	15	16	17	18	19	20	21	22	23	24	25	26
33	0.09***	−0.10***	0.10***	−0.04**	0.06***	0.10***	0.43***	0.12***	−0.11***	0.10**	−0.01	−0.22***	−0.03*
34	0.04*	−0.06***	−0.01	−0.17***	−0.01	0.06***	−0.05***	0.02	−0.21***	−0.06†	−0.20***	−0.11***	−0.04*
35	−0.05***	0.15***	−0.01	−0.04*	−0.04**	0.06***	0.15***	0.09***	−0.08*	0.00	0.01	0.04	−0.06***
36	−0.01	0.05***	0	−0.06***	0	0	−0.01	0.00	−0.04	−0.06	−0.01	−0.03	−0.04*
37	−0.01	0.14***	0	−0.03	−0.03	0.06***	−0.01	0.00	−0.03	−0.07*	−0.05	−0.05	−0.06***

	27	28	29	30	31	32	33	34	35	36
1										
2										
3										
4										
5										
6										
7										
8										
9										
10										
11										
12										
13										
14										
15										
16										
17										

Table 5 (continued)

	27	28	29	30	31	32	33	34	35	36
18										
19										
20										
21										
22										
23										
24										
25										
26										
27	1									
28	0.88***	1								
29	0.56***	0.57***	1							
30	0.45***	0.41***	0.35***	1						
31	0.36***	0.41***	0.45***	−0.01	1					
32	0.66***	0.61***	0.38***	0.23***	0.22***	1				
33	0.01	0.02	−0.10***	0.06***	−0.06*	−0.04*	1			
34	−0.08***	−0.08***	−0.01	−0.01	−0.15***	−0.03*	−0.14***	1		
35	−0.10***	−0.07***	0.03*	−0.08***	0.03	−0.13***	−0.07***	−0.09***	1	
36	−0.03*	−0.04*	−0.03*	−0.03*	−0.04†	−0.01	−0.01	−0.01	−0.01	1
37	−0.05***	−0.05**	−0.06***	−0.05**	−0.05*	−0.01	−0.01	−0.02	−0.01	0.00

This table reports the correlation between the main variables of the study

† $p < 0.10$, * $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$

Table 6 Two-way clustered standard errors bilateral regressions – casewise deletion for each model (same specification as in Table 2)

	(1) Genetic distance	(2) With average number of firms (log) for the countries	(3) With geographical controls	(4) With religion linguistics and colony	(5) With governance indicators	(6) With geographical controls, religion linguistics and colony and governance indicators	(7) With economic development and law system	(8) Full model
Start-up rate between two pairs of countries *i* and *j* (number of new limited liability firms per 1000 working-age individuals between two pairs of countries *i* and *j*)								
Genetic distance	2.5798***	2.4663***	1.3772*	2.3302**	2.0710***	1.3945*	2.4634**	1.3244*
	(3.3914)	(2.6104)	(1.8467)	(2.5664)	(2.6505)	(1.8220)	(2.4921)	(1.7884)
Pairwise average number firms		0.8671***	0.8827***	0.9088***	0.9227***	0.8963***	0.9001***	1.0411***
		(4.5026)	(4.6317)	(4.8432)	(5.3647)	(4.7110)	(4.6765)	(6.0658)
Geodesic distance			0.0031			0.0011		0.0171
			(0.1376)			(0.0497)		(0.9788)
Latitudinal distance			−0.0003			0.0002		−0.0018
			(−0.1032)			(0.0813)		(−0.9396)
Longitudinal distance			0.0014			0.0015		0.0008
			(0.7870)			(0.8462)		(0.5646)
Contiguous			0.0131			0.0610		0.1465
			(0.1063)			(0.5037)		(1.0198)
Landlocked			−0.0334			−0.0343		0.0437
			(−0.3806)			(−0.3719)		(0.4418)
Island			0.3020**			0.2949**		0.2547**
			(2.3796)			(2.3527)		(2.5683)
Elevation			−0.0002**			−0.0002**		−0.0001
			(−2.4680)			(−2.3175)		(−1.4838)
North America Both			0.0017			0.0073		0.0232
			(0.0453)			(0.2074)		(0.7245)

Table 6 (continued)

	(1) Genetic distance	(2) With average number of firms (log) for the countries	(3) With geographical controls	(4) With religion linguistics and colony	(5) With governance indicators	(6) With geographical controls, religion linguistics and colony and governance indicators	(7) With economic development and law system	(8) Full model
Start-up rate between two pairs of countries *i* and *j* (number of new limited liability firms per 1000 working-age individuals between two pairs of countries *i* and *j*)								
South America Both			0.0066			0.0102		−0.2081
			(0.0404)			(0.0643)		(−1.5357)
Asia Both			0.1096**			0.1069**		0.1306***
			(2.5478)			(2.4678)		(4.0029)
Africa Both			0.1221***			0.1149***		0.1053
			(4.9192)			(4.5499)		(0.000)
Europe Both			−0.0373			−0.0319		0.0094
			(−1.2489)			(−1.1305)		(0.4998)
Pacific Both			−0.2493***			−0.2435***		−0.1502***
			(−6.4189)			(−6.3777)		(−3.3939)
Lingustic Distance				0.1872		−0.0347		−0.0743
				(0.2873)		(−0.0626)		(−0.1469)
Religion distance				0.1288		0.0594		0.0346
				(1.6198)		(0.8119)		(0.6215)
Colony				−0.1885		−0.2109		−0.1745
				(−0.9146)		(−0.9600)		(−1.1129)
Common Colonizer				0.0719		−0.0233		−0.0161
				(0.4275)		(−0.1795)		(−0.1449)
Colonial relationship				0.2919		0.1949		0.0969
				(1.4638)		(1.0749)		(0.6331)

Table 6 (continued)

	(1) Genetic distance	(2) With average number of firms (log) for the countries	(3) With geographical controls	(4) With religion linguistics and colony	(5) With governance indicators	(6) With geographical controls, religion linguistics and colony and governance indicators	(7) With economic development and law system	(8) Full model
Start-up rate between two pairs of countries *i* and *j* (number of new limited liability firms per 1000 working-age individuals between two pairs of countries *i* and *j*)								
Rule of law					0.0993			0.1804
					(0.6373)			(1.3010)
Government effectiveness					−0.4476***			−0.3103*
					(−2.7938)			(−1.9529)
Control of corruption					0.0213			−0.0343
					(0.1934)			(−0.3390)
Regulatory quality					0.6795***			0.5901***
					(4.7570)			(4.4147)
Political stability					0.2274***			0.1874**
					(2.5935)			(2.3142)
Accountability					−0.0873			−0.0231
					(−1.4831)			(−0.3537)
Economic development							0.0000	−0.0000**
							(0.6291)	(−2.5386)
Efficiency of bankrutpcy							0.0089	−0.0211
							(0.4704)	(−1.0319)
Legal origin UK Both							0.0026	−0.1786
							(0.0168)	(−1.3924)
Legal origin French Both							0.1212	0.1332**
							(1.5200)	(2.2337)

Table 6 (continued)

	(1) Genetic distance	(2) With average number of firms (log) for the countries	(3) With geographical controls	(4) With religion linguistics and colony	(5) With governance indicators	(6) With geographical controls, religion linguistics and colony and governance indicators	(7) With economic development and law system	(8) Full model
Start-up rate between two pairs of countries *i* and *j* (number of new limited liability firms per 1000 working-age individuals between two pairs of countries *i* and *j*)								
Legal origin Socialist Both							−0.1690	−0.0945
							(−1.4075)	(−0.6417)
Legal origin German Both							−0.1766	0.0223
							(−1.1073)	(0.3149)
Legal origin Scandinavian Both							−1.1615***	−0.7465***
							(−10.3366)	(−4.0960)
Constant	1.0178***	0.6769***	0.6474***	0.3704	0.2523**	0.6068	0.6175***	0.1700
	(12.5741)	(7.5144)	(4.7069)	(0.5992)	(2.2275)	(1.2010)	(4.8092)	(0.3535)
Observations (pairwise between country I and country j)	2016	2016	2016	2016	2016	2016	2016	2016
R-squared	0.0299	0.1457	0.2319	0.1582	0.2975	0.2348	0.1569	0.3734
Standardized beta of Genetic distance	0.125	0.148	0.060	0.145	0.113	0.072	0.165	0.089

Notes. The dependent variable is start-up rate. The independent variables include genetic distance and a set of controls. Refer to the text for detailed variable definitions. The Stata code was obtained from Professor Mitchell Petersen's website and is: dependent_variable independent_variables, fcluster(country i) tcluster(country j)

t-statistics in parentheses. † $p < 0.10$, * $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$

Table 7 Two-way clustered standard errors bilateral regressions – casewise deletion for each model (same specification as in Table 2, now including Cultural dimensions)

	(1) Genetic distance	(2) With average number of firms (log) for the countries	(3) With Add geographical controls	(4) With cultural dimensions	(5) With religion linguistics and colony	(6) With governance indicators	(7) With economic development and law system	(8) Full model
	DV = Difference in Start-up rate between two pairs of countries *i* and *j* (number of new limited liability firms per 1000 working-age individuals between two pairs of countries *i* and *j*)							
Genetic distance	3.3627***	3.5443**	2.1847*	3.3402**	3.1447**	3.0311**	3.5712**	2.2254**
	(2.6579)	(2.3250)	(1.9423)	(2.2793)	(2.1239)	(2.2106)	(2.4493)	(2.1954)
Pairwise average number firms		0.7616***	0.7017***	0.7837***	0.7623***	0.7438***	0.8342***	0.9019***
		(3.3052)	(2.9224)	(3.3837)	(3.2960)	(3.6533)	(3.6172)	(4.0509)
Geodesic distance			0.0122					0.0336**
			(0.5890)					(2.0580)
Latitudinal distance			−0.0007					−0.0038**
			(−0.3309)					(−2.0213)
Longitudinal distance			0.0023					0.0010
			(1.4007)					(0.7965)
Contiguous			−0.0381					0.0505
			(−0.2552)					(0.2904)
Landlocked			0.1383					0.1929
			(0.8598)					(1.0639)
Island			0.1627					0.1583*
			(1.2508)					(1.7024)
Elevation			−0.0002					−0.0002
			(−1.6146)					(−1.5979)
North America Both			0.0263					0.0451
			(0.7368)					(1.3264)

Table 7 (continued)

	(1) Genetic distance	(2) With average number of firms (log) for the countries	(3) With Add geographical controls	(4) With cultural dimensions	(5) With religion linguistics and colony	(6) With governance indicators	(7) With economic development and law system	(8) Full model
	DV = Difference in Start-up rate between two pairs of countries *i* and *j* (number of new limited liability firms per 1000 working-age individuals between two pairs of countries *i* and *j*)							
South America Both			0.0925					−0.0932
			(0.6350)					(−0.8006)
Asia Both			0.2718***					0.2733***
			(6.8898)					(14.5709)
Africa Both			0.1739***					0.2726***
			(4.2250)					(6.2945)
Europe Both			0.0086					0.0433*
			(0.3406)					(1.8713)
Pacific Both			−0.1771***					−0.1211**
			(−4.3432)					(−2.3062)
Lingustic Distance					0.9454			0.2783
					(1.0562)			(0.3634)
Religion distance					0.0112			0.0176
					(0.1042)			(0.1966)
Colony					−0.2253			−0.2221*
					(−0.8660)			(−1.6709)
Common Colonizer					−0.1940			−0.6471***
					(−0.9591)			(−3.9568)

Table 7 (continued)

	(1) Genetic distance	(2) With average number of firms (log) for the countries	(3) With Add geographical controls	(4) With cultural dimensions	(5) With religion linguistics and colony	(6) With governance indicators	(7) With economic development and law system	(8) Full model
	DV = Difference in Start-up rate between two pairs of countries *i* and *j* (number of new limited liability firms per 1000 working-age individuals between two pairs of countries *i* and *j*)							
Colonial relationship					0.1678			0.0883
					(0.4398)			(0.4296)
Power distance				0.0079**				0.0048*
				(1.9677)				(1.7397)
Uncertainty avoidance				−0.0009				−0.0036**
				(−0.4819)				(−2.0870)
Individualism vs Collectivism				0.0019				0.0029
				(0.6025)				(1.2374)
Masculinity vs femininity				−0.0039*				−0.0026
				(−1.9264)				(−1.0328)
Rule of law						−0.1042		−0.0134
						(−0.6458)		(−0.0970)
Government effectiveness						−0.6429***		−0.6902***
						(−3.3345)		(−4.0708)
Control of corruption						0.3836**		0.3911***
						(2.4004)		(2.8352)
Regulatory quality						0.7329***		0.6432***
						(3.9934)		(3.9705)
Political stability						0.1552		0.1118
						(1.3858)		(1.3798)

Table 7 (continued)

	(1) Genetic distance	(2) With average number of firms (log) for the countries	(3) With Add geographical controls	(4) With cultural dimensions	(5) With religion linguistics and colony	(6) With governance indicators	(7) With economic development and law system	(8) Full model
	DV = Difference in Start-up rate between two pairs of countries *i* and *j* (number of new limited liability firms per 1000 working-age individuals between two pairs of countries *i* and *j*)							
Accountability						−0.2186***		−0.2280**
						(−2.9784)		(−2.2423)
Economic development							−0.0000*	−0.0000
							(−1.7213)	(−1.5889)
Efficiency of bankruptcy							0.2277**	0.0413
							(2.4950)	(0.3686)
Legal origin UK Both							−0.1394	−0.3178
							(−0.7679)	(−1.6241)
Legal origin French Both							0.0750	0.2039**
							(0.6524)	(2.0164)
Legal origin Socialist Both							0.2951	0.2107
							(0.0000)	(0.0000)
Legal origin German Both							−0.1110	−0.1476
							(−0.7732)	(−1.0106)
Legal origin Scandinavian Both							−1.0181***	−0.8799***
							(−6.4625)	(−3.6877)
Constant	1.0391***	0.6699***	0.4980***	0.5426***	−0.2087	0.3557**	0.4382***	−0.1795
	(9.0250)	(5.1858)	(3.7060)	(3.5340)	(−0.2482)	(2.2366)	(2.6390)	(−0.2575)

Table 7 (continued)

	(1) Genetic distance	(2) With average number of firms (log) for the countries	(3) With Add geographical controls	(4) With cultural dimensions	(5) With religion linguistics and colony	(6) With governance indicators	(7) With economic development and law system	(8) Full model
	DV = Difference in Start-up rate between two pairs of countries *i* and *j* (number of new limited liability firms per 1000 working-age individuals between two pairs of countries *i* and *j*)							
Observations	820	820	820	820	820	820	820	820
R-squared	0.0426	0.1262	0.2096	0.1599	0.1352	0.2718	0.1752	0.3798
Standardized beta of Genetic distance	0.206	0.218	0.134	0.205	0.193	0.186	0.219	0.137

Notes. The dependent variable is start-up rate. The independent variables include genetic distance and a set of controls. Refer to the text for detailed variable definitions. The Stata code was obtained from Professor Mitchell Petersen's website and is: dependent_variable independent_variables, fcluster(country i) tcluster(country j)

t-statistics in parentheses. † $p < 0.10$, * $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$

Table 8 Two-way clustered standard errors bilateral regressions – casewise deletion for each model using World Value Survey (WVS) Cultural distance measure

	(1)	(2)
	DV = Differences in Start-up rate between two pairs of countries *i* and *j* (number of new limited liability firms per 1000 working-age individuals between two pairs of countries *i* and *j*)	
Genetic distance	1.4830*	1.7525*
	(1.6757)	(1.9206)
Pairwise average number firms	0.7487***	0.7842***
	(3.0404)	(3.1659)
Geodesic distance	−0.0028	0.0068
	(−0.1321)	(0.3987)
Latitudinal distance	0.0006	−0.0001
	(0.2501)	(−0.0514)
Longitudinal distance	0.0019	0.0018*
	(1.4241)	(1.7502)
Contiguous	0.2109	0.1903
	(1.3800)	(1.2614)
Landlocked	0.1758	0.1639
	(1.3403)	(1.1590)
Island	0.3924***	0.3287***
	(3.3003)	(2.9357)
Elevation	−0.0001	−0.0001
	(−1.1690)	(−1.1658)
North America Both	−0.1941	−0.1464
	(−1.0442)	(−0.7288)
South America Both	−0.2674	−0.2773
	(−0.8415)	(−0.8108)
Asia Both	0.2308*	0.2181*
	(1.8029)	(1.7947)
Africa Both	0.8380***	0.8787***
	(14.8223)	(11.3016)
Europe Both	0.1027	0.1989
	(0.8030)	(1.4369)
Pacific Both	−0.5743	−0.3699
	(−1.2760)	(−0.8130)
Linguistic Distance	−0.3551	−0.4171
	(−0.4827)	(−0.6547)
Religion distance	−0.0452	−0.0266
	(−0.5992)	(−0.3330)
Colony	−0.1470	−0.1555
	(−0.8243)	(−0.8975)
Common Colonizer	0.0121	0.1121
	(0.0783)	(0.7713)

Table 8 (continued)

	(1)	(2)
	DV = Differences in Start-up rate between two pairs of countries *i* and *j* (number of new limited liability firms per 1000 working-age individuals between two pairs of countries *i* and *j*)	
Colonial relationship	0.3229	0.3351
	(1.3572)	(1.6194)
WVS: Cultural distance Index	0.0050***	
	(2.7429)	
WVS: Perceptions of Life		−0.0062*
		(−1.7537)
WVS: Work		−0.0023
		(−0.3537)
WVS: Family		0.0175
		(1.3040)
WVS: Politics and Society		0.0085
		(1.5922)
WVS: Religion and Morale		0.0212*
		(1.8046)
WVS: National Identity		0.0304
		(1.5204)
Rule of law	0.0825	0.0490
	(0.6893)	(0.4401)
Government effectiveness	−0.2362	−0.2636
	(−1.2555)	(−1.3190)
Control of corruption	−0.0686	−0.0362
	(−0.7186)	(−0.3500)
Regulatory quality	0.5422***	0.5326***
	(3.4421)	(3.3498)
Political stability	0.2570***	0.2216**
	(2.6519)	(2.3937)
Accountability	−0.0818	−0.0308
	(−0.7895)	(−0.3180)
Economic development	−0.0000***	−0.0000***
	(−3.5324)	(−2.8066)
Efficiency of bankruptcy	−0.0142	−0.0008
	(−0.4549)	(−0.0239)
Legal origin UK Both	−0.1147	−0.1548
	(−0.3289)	(−0.4314)
Legal origin French Both	0.0578	0.0118
	(0.8067)	(0.1509)
Legal origin Socialist Both	0.0306	−0.0482
	(0.1689)	(−0.2682)
Legal origin German Both	−0.1309	−0.0775

Table 8 (continued)

	(1)	(2)
	DV = Differences in Start-up rate between two pairs of countries *i* and *j* (number of new limited liability firms per 1000 working-age individuals between two pairs of countries *i* and *j*)	
	(.)	(−1.1293)
Legal origin Scandinavian Both	−0.5942***	−0.6886***
	(−3.0366)	(−3.5035)
Constant	0.7053	0.6928
	(0.9711)	(1.1027)
Observations	1128	1128
R-squared	0.3616	0.3800
Standardized beta of Genetic distance	0.098	0.116

Notes. The dependent variable is start-up rate. The independent variables include genetic distance and a set of controls. Refer to the text for detailed variable definitions. The Stata code was obtained from Professor Mitchell Petersen's website and is: dependent_variable independent_variables, fcluster(country i) tcluster(country j)

t-statistics in parentheses. † $p < 0.10$, * $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$

Table 9 List of countries in the sample

1. Afghanistan	2. El Salvador	3. Kyrgyzstan	4. Rwanda
5. Albania	6. Ethiopia	7. Latvia	8. Senegal
9. Algeria	10. Finland	11. Lesotho	12. Sierra Leone
13. Argentina	14. France	15. Lithuania	16. Singapore
17. Armenia	18. Gabon	19. Luxembourg	20. Slovenia
21. Australia	22. Georgia	23. Malawi	24. South Africa
25. Austria	26. Germany	27. Malta	28. Spain
29. Azerbaijan	30. Ghana	31. Mauritius	32. Sri Lanka
33. Bangladesh	34. Greece	35. Mexico	36. Suriname
37. Belarus	38. Guatemala	39. Moldova	40. Sweden
41. Belgium	42. Haiti	43. Morocco	44. Switzerland
45. Bhutan	46. Hong Kong	47. Namibia	48. Tajikistan
49. Bolivia	50. Hungary	51. Nepal	52. Thailand
53. Botswana	54. Iceland	55. Netherlands	56. Togo
57. Brazil	58. India	59. New Zealand	60. Tonga
61. Bulgaria	62. Indonesia	63. Niger	64. Tunisia
65. Burkina Faso	66. Iraq	67. Nigeria	68. Turkey
69. Canada	70. Ireland	71. Norway	72. U.S.A
73. Chile	74. Israel	75. Oman	76. Uganda
77. Colombia	78. Italy	79. Pakistan	80. Ukraine
81. Costa Rica	82. Jamaica	83. Philippines	84. United Arab Emirates
85. Croatia	86. Japan	87. Poland	88. United Kingdom
89. Czech Republic	90. Jordan	91. Portugal	92. Uruguay
93. Denmark	94. Kazakhstan	95. Qatar	96. Uzbekistan
97. Dominica	98. Kenya	99. Romania	100. Zambia
101. Dominican Republic	102. Kiribati	103. Russian Federation	

References

Ács ZJ, Autio E, Szerb L (2014) National systems of entrepreneurship: measurement issues and policy implications. Res Policy 43:476–494
Aldrich HE, Fiol MC (1994) Fools rush in? The institutional context of industry creation. Acad Manag Rev 19:645–670
Alesina A, Devleeschauwer A, Easterly W, Kurlat S, Wacziarg R (2003) Fractionalization. J Econ Growth 8: 155–194
Amit R, Brander J, Zott C (1998) Why do venture capital firms exist? Theory and Canadian evidence. J Bus Ventur 13:441–466
Antshel K (2017) Attention deficit/hyperactivity disorder (ADHD) and entrepreneurship. The Academy of Management Perspectives:amp 2016.0144
Ashraf Q, Galor O (2013) The "out of Africa" hypothesis, human genetic diversity, and comparative economic development. Am Econ Rev 103:1–46
Audretsch DB, Grilo I, Thurik AR (2007) Handbook of research on entrepreneurship policy. Edward Elgar Publishing, Cheltenham
Bai Y, Kung JK-s (2011) Genetic distance and income difference: evidence from changes in China's cross-strait relations. Econ Lett 110:255–258
Baron RA (2008) The role of affect in the entrepreneurial process. Acad Manag Rev 33(2):328–340
Baum JR, Frese M, Baron RA (2014) Born to be an entrepreneur? Revisiting the personality approach to entrepreneurship. In: The psychology of entrepreneurship. Psychology Press, pp 73–98
Becker A, Dohmen T, Enke B, Falk A (2014) The ancient origins of the cross-country heterogeneity in risk preferences Stanford Institute for Theoretical Economics (SITE), Department of Economics, Stanford, August 7
Blanchflower DG, Oswald A, Stutzer A (2001) Latent entrepreneurship across nations. Eur Econ Rev 45:680–691
Bleidorn W, Klimstra TA, Denissen JJ, Rentfrow PJ, Potter J, Gosling SD (2013) Personality maturation around the world: a cross-cultural examination of social-investment theory. Psychol Sci 24: 2530–2540
Bönte W, Procher VD, Urbig D (2016) Biology and selection into entrepreneurship—the relevance of prenatal testosterone exposure. Entrep Theory Pract 40:1121–1148
Bosma N, Schutjens V (2011) Understanding regional variation in entrepreneurial activity and entrepreneurial attitude in Europe. Ann Reg Sci 47:711–742
Burger M, Veenhoven R, Kakar L, Commandeur H (2015) Genetic distance and differences in happiness across nations: some preliminary evidence. J Happiness Well-Being 3:142–158
Cameron AC, Gelbach JB, Miller DL (2011) Robust inference with multiway clustering. J Bus Econ Stat 29: 238–249
Campbell DL, Pyun JH (2017) The diffusion of development: along genetic or geographic lines? J Int Dev 29: 198–210
Cardon MS, Wincent J, Singh J, Drnovsek M (2009) The nature and experience of entrepreneurial passion. Acad Manag Rev 34:511–532
Cavalli-Sforza LL, Cavalli-Sforza L, Menozzi P, Piazza A (1994) The history and geography of human genes. Princeton University Press, Princeton
de Holan PM (2014) It's all in your head why we need neuroentrepreneurship. J Manag Inq 23:93–97
Ding X, Zhang Q (2010) 4.02-enzyme regulation comprehensive toxicology (2nd edn) p 9–29, CA McQueen (ed), Oxford: Elsevier
Djankov S, La Porta R, Lopez-de-Silanes F, Shleifer A (2002) The regulation of entry. Q J Econ 117:1–37
Falk A, Becker A, Dohmen T, Enke B, Huffman D, Sunde U (2018) Global evidence on economic preferences. Q J Econ 133(4):1645–1692
Foss NJ, Klein PG (2005) Entrepreneurship and the economic theory of the firm: any gains from trade? In: Handbook of entrepreneurship research. Springer, pp 55–80
Giuliano P, Spilimbergo A, Tonon G (2006) Genetic, cultural and geographical distances. Cent Econ Pol Res
Gorodnichenko Y, Roland G (2017) Culture, institutions, and the wealth of nations. Rev Econ Stat 99:402–416
Greene FJ, Han L, Martin S, Zhang S, Wittert G (2014) Testosterone is associated with self-employment among Australian men. Econ Hum Biol 13:76–84
Grilo I, Thurik R (2005) Latent and actual entrepreneurship in Europe and the US: some recent developments. Int Entrep Manag J 1:441–459

Guiso L, Sapienza P, Zingales L (2009) Cultural biases in economic exchange? Q J Econ 124:1095–1131
Günther T, Jakobsson M (2016) Genes mirror migrations and cultures in prehistoric Europe—a population genomic perspective. Curr Opin Genet Dev 41:115–123
Hart DM, Acs ZJ (2011) High-tech immigrant entrepreneurship in the United States. Econ Dev Q 25:116–129
Hayton JC, Cacciotti G (2013) Is there an entrepreneurial culture? A review of empirical research. Entrep Reg Dev 25:708–731
Hébert RF, Link AN (1988) The entrepreneur: mainstream views & radical critiques. Praeger Publishers, New York
Hessels J, Rietveld CA, van der Zwan P (2014) Unraveling two myths about entrepreneurs. Econ Lett 122: 435–438
Hofstede G (2011) Dimensionalizing cultures: the Hofstede model in context. Online readings in psychology and culture 2:8
Hubbard R (1998) The golden goose? Understanding (and taxing) the saving of entrepreneurs. Adv Entrep Innov Econ Growth 10:43–69
King RG, Levine R (1993) Finance, entrepreneurship and growth. J Monet Econ 32:513–542
Klapper L, Amit R, Guillen M, Quesada J (2010) In: Lerner J, Schoar A (eds) Entrepreneurship and firm formation across countries, international differences in entrepreneurship. University of Chicago Press, Chicago
Krueger N, Liñán F, Nabi G (2013) Cultural values and entrepreneurship. Taylor & Francis, London
Kuhnen CM, Chiao JY (2009) Genetic determinants of financial risk taking. PLoS One 4(2):e4362
La Porta R, Lopez-de-Silanes F, Shleifer A, Vishny R (1999) The quality of government. J Law Econ Org 15: 222–279
Lindquist MJ, Sol J, Van Praag M (2015) Why do entrepreneurial parents have entrepreneurial children? J Labor Econ 33:269–296
Mayer T, Zignago S (2011) Notes on CEPII distances measures: The GeoDist database. Working Papers 2011–25, CEPII Research Center
Mihet R (2013) Effects of culture on firm risk-taking: a cross-country and cross-industry analysis. J Cult Econ 37:109–151
Mitchell RK, Busenitz LW, Bird B, Marie Gaglio C, McMullen JS, Morse EA, Smith JB (2007) The central question in entrepreneurial cognition research 2007. Entrep Theory Pract 31:1–27
Nicolaou N, Shane S (2014) Biology, neuroscience, and entrepreneurship. J Manag Inq 23:98–100
Nicolaou N, Shane S, Cherkas L, Hunkin J, Spector TD (2008) Is the tendency to engage in entrepreneurship genetic? Manag Sci 54:167–179
Nicolaou N, Shane S, Cherkas L, Spector TD (2009) Opportunity recognition and the tendency to be an entrepreneur: a bivariate genetics perspective. Organ Behav Hum Decis Process 110:108–117
Nofal AM, Nicolaou N, Symeonidou N, Shane S (2018) Biology and management: a review, critique, and research agenda. J Manag 44:7–31
Obschonka M, Stuetzer M (2017) Integrating psychological approaches to entrepreneurship: the entrepreneurial personality system (EPS). Small Bus Econ 49:203–231
Obschonka M, Schmitt-Rodermund E, Silbereisen RK, Gosling SD, Potter J (2013) The regional distribution and correlates of an entrepreneurship-prone personality profile in the United States, Germany, and the United Kingdom: a socioecological perspective. J Pers Soc Psychol 105:104
Petersen MA (2009) Estimating standard errors in finance panel data sets: comparing approaches. Rev Financ Stud 22:435–480
Proto E, Oswald AJ (2017) National happiness and genetic distance: a cautious exploration. Econ J 127:2127–2152
Rentfrow PJ, Gosling SD, Potter J (2008) A theory of the emergence, persistence, and expression of geographic variation in psychological characteristics. Perspect Psychol Sci 3:339–369
Saxenian A (2002) Silicon Valley's new immigrant high-growth entrepreneurs. Econ Dev Q 16:20–31
Schmitt DP, Allik J, McCrae RR, Benet-Martínez V (2007) The geographic distribution of big five personality traits: patterns and profiles of human self-description across 56 nations. J Cross-Cult Psychol 38:173–212
Shane S, Nicolaou N (2015) The biological basis of entrepreneurship. In: Colarelli SM, Arvey RD (eds.) The biological foundations of organizational behavior. University of Chicago Press, Chicago, p 71–89
Schwab K (2005) World economic forum, global competitiveness report (2005-2006). Geneva: WEF
Spolaore E, Wacziarg R (2009) The diffusion of development. Q J Econ 124:469–529
Spolaore E, Wacziarg R (2013) How deep are the roots of economic development? J Econ Lit 51:325–369
Spolaore E, Wacziarg R (2016) Ancestry, language and culture. In: The palgrave handbook of economics and language. Palgrave Macmillan, London, pp. 174–211

Thomas AS, Mueller SL (2000) A case for comparative entrepreneurship: assessing the relevance of culture. J Int Bus Stud 31:287–301

Thornton PH, Ribeiro-Soriano D, Urbano D (2011) Socio-cultural factors and entrepreneurial activity: an overview. Int Small Bus J 29:105–118

Unger JM, Rauch A, Weis SE, Frese M (2015) Biology (prenatal testosterone), psychology (achievement need) and entrepreneurial impact. J Bus Ventur Insights 4:1–5

Verheul I, Block J, Burmeister-Lamp K, Thurik R, Tiemeier H, Turturea R (2015) ADHD-like behavior and entrepreneurial intentions. Small Bus Econ 45:85–101

Verheul I, Rietdijk W, Block J, Franken I, Larsson H, Thurik R (2016) The association between attention-deficit/hyperactivity (ADHD) symptoms and self-employment. Eur J Epidemiol 31:793–801

White RE, Thornhill S, Hampson E (2006) Entrepreneurs and evolutionary biology: the relationship between testosterone and new venture creation. Organ Behav Hum Decis Process 100:21–34

White RE, Thornhill S, Hampson E (2007) A biosocial model of entrepreneurship: the combined effects of nurture and nature. J Organ Behav 28:451–466

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.