

Vu, Trung V.

Working Paper

Are genetic traits associated with riots? The political legacy of prehistorically determined genetic diversity

Suggested Citation: Vu, Trung V. (2021) : Are genetic traits associated with riots? The political legacy of prehistorically determined genetic diversity, ZBW - Leibniz Information Centre for Economics, Kiel, Hamburg

This Version is available at:

<https://hdl.handle.net/10419/234467>

Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

Terms of use:

Documents in EconStor may be saved and copied for your personal and scholarly purposes.

You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.

Are genetic traits associated with riots? The political legacy of prehistorically determined genetic diversity*

Trung V. Vu[†]

Department of Economics, University of Otago, Dunedin, New Zealand

Abstract

This paper establishes that the worldwide distribution of political instability has its deep historical roots in genetic diversity, predetermined over the prehistoric course of the exodus of *Homo sapiens* from East Africa tens of thousands of years ago. It proposes that the relationship between prehistorically determined genetic diversity and contemporary political instability follows a U-shaped pattern. More specifically, genetic diversity at first reduces the persistence of political instability by increasing the opportunity cost of engaging in riots and revolts. However, genetically fragmented societies tend to suffer from interpersonal mistrust and the under-provision of public goods, which plausibly undermine the establishment of politically stable regimes. Using an ancestry-adjusted index of predicted genetic diversity, this paper consistently finds precise estimates that genetic diversity imparts a U-shaped influence on different measures of political instability and the probability of observing the occurrence of riots and revolts across 141 countries. Furthermore, the contribution of genetic diversity to political instability is at least partially mediated through income/productivity levels, the provision of public goods, income inequality and social trust.

Key words: genetic diversity, fractionalization, political instability, riots, conflict.

JEL codes: E02, F50, N30, O11, Z13.

* My gratitude goes to Professors Dorian Owen and Alfred Haug for their constructive comments and kind support from immature ideas to the end of this research. I also thank an Editor, Professor David Stadelmann, and two anonymous reviewers for valuable comments and suggestions. Moreover, the paper benefits from financial support from the University of Otago provided in the form of a Doctoral scholarship. The usual caveat applies. The Supporting Information contains replication files and a supplementary online appendix.

[†] **Correspondence:** Department of Economics, University of Otago, PO Box 56, Dunedin 9054, New Zealand. Email address: trungvu.econs@gmail.com.

One of the most powerful hypotheses in political economy is the notion that social divisions undermine economic progress, not just in extremis, as in the case of a civil war, but also in more normal times (Banerjee et al., 2005, p. 639).

1. Introduction

This paper attempts to uncover one of the deepest roots of political instability through exploiting international differences in the composition of genetic traits, originating from the prehistoric course of the exodus of *Homo sapiens* from East Africa (Ashraf & Galor, 2013, 2018).¹ Before articulating the central hypothesis, this section provides a brief review of an emerging strand of literature exploring causes and consequences of political instability, and justifies the desirability of employing prehistorically determined genetic diversity to explain the global distribution of political unrest.

Political instability or uncertainty is one of the most serious impediments to sustaining long-run economic growth (Azzimonti, 2011; Aisen & Veiga, 2013). It is widely acknowledged that politically stable countries are more likely to enjoy robust growth driven by higher levels of capital per unit of labour, compared with their politically unstable counterparts. This argument rests upon the premise that countries characterized by greater political uncertainty tend to adopt shortsighted policies, thus reducing investment and hindering capital accumulation (Azzimonti, 2011).² Julio and Yook (2012) find that political instability lowers firms' investment, thereby impeding national innovative capacity and economic growth. Other scholars contend that political instability retards economic development through exacerbating public debt (Persson & Svensson, 1989; Alesina & Tabellini, 1990), tax distortions, excessive public spending (Battaglini & Coate, 2008; Yared, 2010), and higher seigniorage (Cukierman et al., 1992; Aisen & Veiga, 2008). Moreover, uncertainty associated with a country's political systems is detrimental to environmental performance (Fredriksson & Svensson, 2003; Fredriksson & Wollscheid, 2014) and financial development (Roe & Siegel, 2011).³ Not surprisingly, a growing body of research has been devoted to understanding factors shaping the persistence and pervasiveness of political instability across the world.

¹ Political instability refers to the probability of a government collapse and/or politically motivated violence, including terrorism.

² Many studies empirically document that social and political unrest exerts a negative influence on economic growth (see, for instance, Alesina & Perotti, 1996; Carmignani, 2003; Jong-A-Pin, 2009; Aisen & Veiga, 2013).

³ Additionally, conflicts and political instability hinder economic development through directly affecting a country's human capital stock. Arbatli et al. (2020) document that more than 16 million deaths have been attributed to civil conflicts across the world since the end of the Second World War.

The existing literature provides several explanations for cross-country differences in political instability. For example, an influential study by Alesina and Perotti (1996) puts forward the idea that income inequality is linked to the prevalence of political instability through inducing social discontent. Furthermore, socio-political unrest is likely to proliferate in the developing world, characterized by lower income per capita, higher poverty rates and poor-quality institutions (Alesina & Perotti, 1996; Fearon & Laitin, 2003; Collier & Hoeffler, 2004; Blattman & Miguel, 2010). More specifically, societies with widespread poverty and low levels of GDP per capita may suffer from the persistence of violent conflicts over scarce resources. This suggests that uncertainty associated with the political environment may be driven by economic performance. Previous studies reveal numerous “proximate” causes of political instability, including, but not limited to, resource wealth (Dutt & Mitra, 2008), the size of population and population density (Goldstone, 2002; Hegre & Sambanis, 2006; Krieger & Meierrieks, 2011; Acemoglu et al., 2020), and trade openness (Martin et al., 2008).

Nevertheless, attempts at empirical estimation of the effects of these factors on political instability face several challenges. A major concern relates to possible endogeneity bias, partially induced by reverse causation running from economic performance to the prevalence of social and political unrest. The causal influence may run in either direction, potentially leading to biased and inconsistent estimates. For instance, an unequal distribution of income and/or power within a country may lead to greater political uncertainty because it induces social dissatisfaction with the government. However, political instability may exacerbate (economic) inequality within a society. The basic intuition, as discussed previously, is that policymakers in politically unstable economies tend to implement sub-optimal economic policies, possibly leading to a more unbalanced distribution of income. This is because a higher probability of losing office constrains the incumbent’s window of opportunity (Campante et al., 2009).⁴ A brief review of the related literature presented earlier also reveals that there exists a two-way relationship between economic performance and political uncertainty. Identification strategies, therefore, critically rest upon generating plausibly exogenous sources of variation in economic performance that help explain the evolution of socio-political unrest. Theoretically, the interrelationship between political instability and its “proximate” causes motivates tracing the deep origins of social and political unrest (Vu, 2021a).

⁴ Campante et al. (2009) suggest that officials in politically unstable countries are more likely to misuse public resources for personal gain with the cost borne by the rest of the society. Hence, high levels of political instability may lead to greater inequality by shortening policymakers’ horizons.

Figures 1 and 2 illustrate several stylized facts that motivate the empirical analysis of the paper. In particular, there exists striking variation in the degree of political instability across the globe. While the African world suffers from pervasive socio-political unrest, European countries are typically endowed with politically stable regimes (Figure 1). Figure 2 depicts the evolution of political instability within selected countries and regions. It reveals that the level of political instability exhibited little variation within a country from 1996 to 2015. This is suggestive of the persistence of this widespread social concern over time. The pattern remains largely unchanged when exploiting time-series data of countries characterized by the highest (Afghanistan and Somalia), lowest (Finland) and medium (Senegal) levels of political instability (Figure 2). Moreover, the regional averages of political instability in Africa and Europe remained relatively stable between 1996 and 2015 (Figure 2). The pervasiveness and persistence of political instability justify the importance of investigating the deep roots of political unrest (Grechyna, 2018; Vu, 2021a).⁵

Against this background, this research, to my knowledge, is the first attempt to explore the extent to which prehistorically determined genetic diversity helps shape the global distribution of political instability. The current paper builds upon the seminal article by Ashraf and Galor (2013), which argues that genetic diversity has persistent and U-shaped effects on income per capita. Indeed, the worldwide distribution of genetic diversity originated from the prehistoric course of the exodus of *Homo sapiens* from East Africa taking place tens of thousands years ago (Ashraf & Galor, 2013, 2018). According to the widely acknowledged “out of Africa” theory, a settlement’s migratory distance from East Africa was inherently linked to lower levels of genetic diversity, which, in turn, translate into long-lasting variation in economic prosperity across the world. Therefore, a major distinguishing feature of this study lies in attempts to understand the deep historical roots of political instability. Importantly, the historical approach of this paper mitigates a key threat to identification stemming from reverse causation. This is because present-day political instability plausibly exerts no direct influence on prehistorically determined genetic diversity. It is worth emphasizing that understanding the deep historical roots of political instability is relevant for obtaining insights into today’s policies. Specifically, effective interventions intended to curtail this widespread social concern

⁵ The persistence of political instability can be driven by a country’s fundamental (fixed) factors, including geographic endowments or historical events (Spolaore & Wacziarg, 2013; Nunn, 2020). A recent empirical analysis by Grechyna (2018) sheds light on the deep roots of political instability, focusing on the effects of geographic proximity on the cross-border diffusion of socio-political unrest. Furthermore, Vu (2021a) investigates the contribution of accumulated statehood experience from 3500BCE to 2000CE to the prevalence of contemporary socio-political unrest outside the European world.

arguably require being reconcilable with the prevailing historical environment (Spolaore & Wacziarg, 2013; Nunn, 2020; Vu, 2021b, 2021a). Surprisingly, little is known about the fundamental drivers of political instability.

My arguments begin with several observations on the cross-country correlation between genetic diversity and political instability. As depicted in Figures 1 and 3, countries characterized by large heterogeneity in the composition of genetic traits tend to experience greater uncertainty associated with the political environment (e.g., Somalia, Ethiopia, Sudan, Central African Republic and Egypt). In addition, socio-political unrest is prevalent in societies that typically experience low levels of genetic diversity (e.g., Bolivia, Peru, Indonesia, and the Philippines). In contrast, stable political systems are commonly established in countries with intermediate levels of genetic diversity (e.g., Finland, New Zealand, Switzerland, Botswana, and Japan). These stylized facts suggest that both high and low degrees of genetic diversity are conducive to political instability. Meanwhile, countries with a medium level of prehistorically determined genetic diversity tend to set up politically stable regimes. This research proposes that genetic diversity, up to a certain threshold, attenuates social and political unrest. However, excessive heterogeneity in the composition of genetic traits is associated with greater political instability. Overall, the relationship between genetic diversity and political instability may exhibit a U-shaped pattern.

Previous studies provide numerous theoretical arguments underpinning the hypothesized U-shaped relationship between genetic diversity and political instability. Conventional wisdom regarding the motivation of engaging social and political unrest suggests that low levels of income per capita (or productivity) are a key determinant of civil conflicts and political instability. Accordingly, an individual is more inclined to participate in violence if the opportunity cost of riots is less than that of peaceful cooperation or negotiation (Collier & Sambanis, 2002; Collier & Hoeffler, 2004; Blattman & Miguel, 2010; Ang & Gupta, 2018). As such, low-income societies are more likely to suffer from political instability because their citizens typically experience a lower opportunity cost of engaging in riots and revolts. Furthermore, people living in less developed economies tend to engage in violent conflicts over scarce resources due to widespread poverty and lack of (institutional) capacity to resolve internal conflicts (Blattman & Miguel, 2010; Grechyna, 2018).

An influential viewpoint in the comparative development literature highlights the beneficial role of genetic diversity in fostering long-term economic performance (Ashraf & Galor, 2013, 2018). Genetic diversity, up to a point, enhances a country's innovative capacity

because it represents a diverse range of individual-level traits, including abilities, skills and cognitive approaches. The availability of a variety of complementary cognitive approaches and human traits is conducive to innovative activities by fostering problem-solving and knowledge creation (Ashraf & Galor, 2013). Moreover, Ashraf and Galor (2018) argue that genetic diversity facilitates economic specialization and the ability to adapt to rapid changes in technologies, leading to higher levels of productivity. Consistent with this viewpoint, Depetris-Chauvin and Özak (2020) document that genetic and linguistic diversity, measured at the ethnic level, created favourable conditions for the division of labour in pre-industrial times. Exploiting data across ethnically homogeneous high schools in the United States, Cook and Fletcher (2018) reveal that high-school genetic diversity imparts a positive influence on the socio-economic outcomes of the school's graduates, such as higher earnings, greater employment prestige and improved educational attainment. Accordingly, students' exposure to genetically diverse environments is associated with the formation of personality traits geared towards creativity and the acceptance of novel ideas (Cook & Fletcher, 2018). Therefore, genetic diversity at first reduces political instability *via* fostering income, productivity and innovation, all of which represent a higher opportunity cost of engaging in riots, revolts and conflicts.

Nevertheless, Ashraf and Galor (2013, 2018) demonstrate that excessive genetic diversity is linked to interpersonal mistrust and reduced social cohesiveness. Genetically fragmented societies are characterized by large heterogeneity in individual-level values, beliefs, preferences and predispositions in social interactions, which are detrimental to building social networks (Ashraf & Galor, 2018). For this reason, genetic diversity induces mutual mistrust across individuals within populations, thus depressing the level of social capital. In addition, genetic diversity undermines the provision of public goods and redistributive policies by widening the spectrum of preferences for income redistribution (Ashraf & Galor, 2013, 2018; Arbatlı et al., 2020). It is argued that genetically diverse countries may find it hard to reconcile highly heterogeneous preferences across groups and individuals for specific types of public goods that should be provided by the government (Alesina et al., 1999; Alesina & Ferrara, 2005). Political instability largely arises from social dissatisfaction with the government and conflicts between groups and individuals within a society, driven by an unbalanced distribution of wealth and power (Alesina & Perotti, 1996; Østby, 2008; Esteban & Ray, 2011; Ezcurra & Palacios, 2016). Therefore, inefficiencies and distortions in the provision of public goods and the distribution of income, which typically proliferate in highly diverse societies, may provoke the prevalence and persistence of socio-political grievances, leading to greater political instability. Additionally, (violent) conflicts between groups and

individuals are less likely to be resolved by peaceful cooperation or negotiation in genetically fragmented economies due to social mistrust. The basic idea is that social networks and societal cohesiveness facilitate communication and negotiation, which help address collective action problems (Alesina et al., 1999; Grafton & Knowles, 2004; Alesina et al., 2019). These narratives suggest that genetic diversity creates barriers to establishing politically stable regimes *via* hindering the provision of public goods and/or income distribution and inducing interpersonal mistrust.

The aforementioned theoretical arguments suggest that prehistorically determined genetic diversity may exert both beneficial and adverse effects on the stability of political systems. On the one hand, it reduces socio-political unrest by fostering productivity, income per capita and innovative activities. On the other hand, the persistence of riots and revolts is attributable to inefficiencies and distortions in income distribution, the under-provision of public goods and social non-cohesiveness, mainly induced by genetic fragmentation. Following Ashraf and Galor (2013, 2018), this study proposes that countries endowed with a medium level of genetic heterogeneity tend to exploit the advantages of diverse complementary abilities, skills and cognitive approaches to achieve higher levels of productivity. Meanwhile, they are less likely to suffer from societal non-cohesiveness compared with excessively heterogeneous societies. Given the diminishing marginal impacts of diversity and homogeneity on political instability, countries with a medium degree of genetic heterogeneity are able to establish politically stable systems. By contrast, both excessively diverse and homogenous societies are expected to suffer from the persistence and pervasiveness of political instability. This forms a U-shaped relationship between genetic diversity and political instability.

Using data for 141 countries, this paper provides robust evidence of a U-shaped association between genetic diversity and political instability. In line with the main hypothesis, I find that genetic diversity exerts persistent effects on political uncertainty, at least partially, through shaping the opportunity cost of engaging in riots, the provision of public goods and social trust. Overall, the results reveal that the pervasiveness and persistence of political instability are deeply rooted in genetic diversity, predetermined over the prehistoric course of the exodus of human species from East Africa tens of thousands of years ago.

The study proceeds as follows. Section 2 provides a brief review of related literature. Section 3 discusses data and identification methods. Section 4 presents the main findings, followed by a discussion of robustness analyses in Section 5. Section 6 provides additional evidence, and Section 7 concludes.

2. Related literature

This study belongs to a well-established line of research exploring the effects of population diversity, measured at different levels of aggregation, on numerous economic, social and political outcomes (Alesina & Ferrara, 2005; de Soysa & Almås, 2019).

An early contribution by Easterly and Levine (1997) argues that ethnic diversity impedes economic growth via hindering the provision of growth-enhancing public goods. Consistent with this finding, many studies document that ethnolinguistic fractionalization is associated with poor governance, higher corruption, reduced economic growth and distorted public spending (Alesina et al., 1999; Grafton et al., 2002; Alesina et al., 2003; Alesina & Ferrara, 2005; Banerjee et al., 2005; Montalvo & Reynal-Querol, 2005a; Grafton et al., 2007; Desmet et al., 2012; Gören, 2014). The basic idea is that ethnolinguistically fragmented societies typically experience diverse preferences for public goods and redistribution. As such, ethnic diversity lowers redistribution, the quality of institutions and the provision of public goods that require social cohesion and coordination (de Soysa & Almås, 2019). This helps explain substantial variation in economic prosperity throughout the world. However, recent research departs from this viewpoint by postulating that diversity is conducive to several economic outcomes. Alesina et al. (2016), for instance, demonstrate that the diversity of immigrants' birthplaces helps sustain growth. This is attributed to the availability of a diverse range of skills and cognitive abilities that foster long-term economic performance. This viewpoint is also supported by Bove and Elia (2017) and Montalvo and Reynal-Querol (2005a). Using district-level data for Zambia, Gisselquist et al. (2016) indicate that ethnic fractionalization has positive effects on the provision of public goods, which stands in stark contrast to most previous cross-country studies. Hence, the existing literature provides highly mixed findings regarding the contribution of population diversity to economic prosperity.

This paper is also closely related to various attempts at estimating the impacts of ethnic fractionalization on the onset of civil wars. A widely accepted theory posits that ethnolinguistically fractionalized countries tend to suffer from the prevalence of violent conflicts due to highly diverse preferences for public goods and redistribution (Arbatlı et al., 2020). In addition, ethnic diversity induces violent competition over scarce resources, particularly in developing countries (Arbatlı et al., 2020). By contrast, Fearon and Laitin (2003) show that ethnolinguistic diversity exerts no statistically significant influence on civil conflicts. In a similar vein, several scholars demonstrate that ethnic diversity is not a robust predictor of the occurrence of civil wars across the world (Collier & Hoeffler, 2004, 2007). However,

Montalvo and Reynal-Querol (2005b) construct an index of ethnic polarization, and reveal that this measure of population diversity has statistically significant and positive effects on conflicts. By contrast, the contribution of ethnic fractionalization to conflicts is imprecisely estimated at conventionally accepted levels of statistical significance. These findings are supported by a subsequent empirical analysis by Esteban et al. (2012). Therefore, prior research offers mostly inconclusive results regarding the causal influence of ethnolinguistic fractionalization/polarization on the persistence of socio-political unrest.

A commonly used measure of population diversity in previous studies is an index of ethnolinguistic fractionalization or polarization, typically aggregated at the ethnic level. A key concern about using these aggregated metrics relates to the possibility that they fail to capture substantial heterogeneity across individuals within ethnic groups (Ashraf & Galor, 2013, 2018; Cook & Fletcher, 2018; Arbatlı et al., 2020). Specifically, conventional measures of ethnolinguistic diversity are merely based on the proportional representations of different ethnic groups within an economy (Alesina et al., 2003). Hence, they reflect population diversity across ethnic groups, and assume intragroup homogeneity. The use of these aggregated metrics may underestimate the degree of population heterogeneity, possibly accounting for highly mixed findings in previous studies.

Meanwhile, an empirical investigation of the effects of population diversity on political instability, as articulated in Section 1, requires capturing divergence in values, beliefs, predispositions and preferences across individuals. Thus, this paper goes beyond previous studies by employing a measure of prehistorically determined genetic diversity to explain the global distribution of political instability. The index of genetic diversity constructed by Ashraf and Galor (2013) reflects the likelihood of genetic dissimilarities between two randomly selected individuals with respect to a given spectrum of genetic markers. As suggested by Ashraf and Galor (2013, 2018) and Arbatlı et al. (2020), genetic diversity captures the degree of heterogeneity both across and within ethnic groups, thus addressing a major issue related to the adoption of conventional metrics of ethnolinguistic fragmentation.

As discussed earlier, the use of prehistorically determined genetic diversity to explain the worldwide distribution of political instability helps provide a valid basis for statistical inference because reverse causation is largely circumvented. This argument is motivated by the influential work by Ashraf and Galor (2013), in which the well-known theory of human evolution is exploited to explain the global pattern of long-term development. Specifically, archaeological evidence suggests that *Homo sapiens* evolved from East Africa nearly 300,000

years ago before populating the entire globe along their “out of Africa” migration routes approximately 60,000 years ago (Ashraf & Galor, 2013, 2018). The level of genetic heterogeneity within current populations is negatively associated with migratory distance from East Africa, which is based on the “serial founder effect” developed by population geneticists.

In particular, Ashraf and Galor (2018) postulate that the prehistoric course of the exodus of human species from East Africa took place through a series of discrete steps. When a sub-group of human species migrated out of their parental settlement to establish a new society, they carried with them a sub-set of genetic traits of the original population. This process created the serial founder effect, which imparts a persistent influence on variation in genetic traits of current-day societies. It is argued that domestic and international migration flows driven by the persistence and prevalence of riots and revolts may directly affect ethnic diversity. This concern is presumably alleviated by using a prehistorically determined index of genetic diversity. Furthermore, Arbatlı et al. (2020) reveal that the use of conventional measures of ethnic diversity can be subject to measurement errors. This stems from the political economy of ethnicity classifications in a country’s population census. These arguments point to the desirability of adopting prehistorically determined genetic diversity as a plausibly exogenous source of international variation in present-day political instability (Ashraf & Galor, 2013, 2018; Arbatlı et al., 2020).

3. Data and econometric methods

3.1. Model specification

To check for the presence of a U-shaped relationship between genetic diversity and political instability, this paper specifies the following model:

$$PIS_i = \beta_0 + \beta_1 Pdiv_aa_i + \beta_2 Pdiv_aa_sqr_i + \beta_3 X_i + \varepsilon_i \quad [1]$$

in which PIS is a measure of political instability, constructed using the World Bank’s index of Political Stability and Absence of Violence/Terrorism. The main variables of interest include $Pdiv_aa$ and $Pdiv_aa_sqr$, representing the linear and quadratic terms of an ancestry-adjusted index of predicted genetic diversity developed by Ashraf and Galor (2013). The estimated coefficients of these two variables capture the hypothesized non-monotonic relationship between genetic diversity and political instability. Consistent with the main hypothesis, β_1 and β_2 are expected to be negative and positive, respectively. X is a vector of main control variables, including geographic attributes and continent dummies. More specifically, geographic controls include terrain ruggedness, mean levels of elevation, land suitability for

agriculture, distance to the nearest waterways, the fraction of the population at risk of contracting malaria, latitude, average precipitation and temperature, and a binary variable for island countries. Dummy variables for Europe, America, Oceania and Asia are also included in the benchmark model (Africa is the excluded base group). ε denotes the disturbance term. The baseline sample, primarily dictated by the availability of data, covers 141 countries ($i = 1, 2, \dots, 141$). The online Appendix contains a detailed description of variables and data sources. Descriptive statistics for all variables are provided in Appendix Tables A1 to A3.

3.2. Political instability

This study employs the World Bank's index of Political Stability and Absence of Violence/Terrorism to capture cross-country differences in political instability. According to the World Bank's definitions, political instability corresponds to the probability of a government collapse caused by contravention of established conventions (unconventional means) and/or politically motivated violence, including terrorism. Political instability also equates with political turnover/uncertainty when it comes to reflecting major changes in the government (Grechyna, 2018). More broadly, political instability refers to the probability of riots, revolutions and violence, among others.

The index of political instability is derived from one of the World Bank's Worldwide Governance Indicators. The construction of this indicator is based on survey participants' perceptions of the likelihood of a destabilization of the government and politically motivated violence, including terrorism. Higher values correspond to greater political stability and absence of violence. For ease of interpretation, this index is rescaled by computing the difference between the largest value of the whole sample and each country-year observation between 1996 and 2015, following Grechyna (2018). By doing so, I obtain an internationally comparable measure of political instability, with higher values representing greater political uncertainty (Figure 1). The aforementioned discussion suggests that political instability exhibits little within-country variation over time. To estimate the cross-sectional models, I calculate a simple average of political instability across the period 1996-2015. Other measures of political instability are used for robustness checks in Section 5.

3.3. Prehistorically determined genetic diversity

Genetic diversity refers to the probability that two randomly selected individuals from a relevant population are genetically dissimilar, with respect to a given spectrum of genetic markers (Ashraf & Galor, 2013, 2018). This is widely known as expected heterozygosity

developed by population geneticists (Ashraf & Galor, 2018). The level of genetic diversity of a relevant population is constructed by the average of heterozygosity across multiple loci.

However, historical migration flows might shape the evolution of genetic diversity, leading to potential endogeneity bias. For this reason, I employ an ancestry-adjusted index of predicted genetic diversity ($Pdiv_{aa}$) of Ashraf and Galor (2013), calculated by countries' migratory distance from Addis Ababa. The migratory distance is highly predictive of genetic diversity before the historical expansion of European colonizers in 1500CE. The degree of ancestry-adjusted predicted genetic diversity of today's populations is calculated through a two-step procedure. In the first stage, Ashraf and Galor (2013) measure the level of predicted genetic diversity of ancestral populations in 1500CE based on migratory distance from Addis Ababa. In the second stage, predicted genetic diversity is weighted by the share of the present-day population descended from different places in 1500CE. Data on historical migration flows are taken from the World Migration Matrix developed by Putterman and Weil (2010). This approach accounts for the effects of the European expansion throughout the world starting in 1500CE, which partially shaped the evolution of genetic traits within a population.

3.4. Main control variables

In line with the empirical framework of Ashraf and Galor (2013), Arbatlı et al. (2020) and Depetris-Chauvin and Özak (2020), I allow numerous potential confounding factors to enter the baseline model. This strategy helps address a key concern that my findings are exclusively driven by conventional explanations for the deep roots of political instability.

A major concern relates to the possibility that country-level geographic (fixed) characteristics may simultaneously shape the worldwide distribution of genetic diversity and political instability. If this were the case, my findings would be easily explained away by these exogenous factors. Hence, the main analysis is augmented with numerous geographic variables. Fearon and Laitin (2003) suggest that countries characterized by rugged terrains may suffer from the prevalence of socio-political unrest. This is because terrain ruggedness may provoke riots via providing safe havens for rebels, thus hindering the government's ability to repress their rebellion (Fearon & Laitin, 2003; Arbatlı et al., 2020). Furthermore, Michalopoulos (2012) reveals that rugged terrains are conducive to population diversity through increasing the number of geographically isolated regions. It also follows from the findings of Michalopoulos (2012) that mean elevation and land suitability for agriculture can be linked to uncertainty associated with a country's political systems through shaping population diversity. Hence, the baseline regression accounts for these geographic attributes.

The long-term comparative development literature reveals that other geographic factors, including distance to the nearest waterways, the fraction of population at risk of contracting malaria, latitude, precipitation and temperature, can affect institutional quality, income levels or trade openness. These variables plausibly exert direct or indirect impacts on the prevalence of socio-political unrest through shaping the quality of institutions, productivity or trade-related mechanisms. To mitigate this concern, these additional geographic controls are included in the benchmark model. Following Arbatlı et al. (2020), I include a dummy variable for island countries because these nations might exhibit different patterns of (early) development driven by their geographic isolation. A final consideration is that my findings can be driven by unobserved continent-specific factors. The basic intuition is that countries located in the same region may share common cultures, histories and other geographic features, potentially shaping the evolution of genetic diversity and political instability. This motivates incorporating continent dummies in the main analysis.

It is noteworthy that conventional explanations for international variation in political instability, including income per capita, institutional quality, the provision of public goods, redistributive policies and social trust, are excluded from the main econometric specification for two main reasons. *First*, these “proximate” drivers of political instability are jointly determined by and interrelated with political uncertainty. A key challenge with drawing valid statistical inference on the effects of these variables on political instability stems from reverse causality. Even when I allow them to enter the baseline model specification, the estimated coefficients should not be interpreted as causation. However, the results presented in Section 5 reveal that my findings are not purely driven by these conventional determinants of political instability. Therefore, they are excluded from Eq. [1] to obtain consistent inference.

Second, the main hypothesis implies that genetic diversity may transmit to political instability through shaping cross-country differences in productivity, institutional quality, mistrust and redistributive policies, among others. Including potential channels of transmission in the regression arguably masks much of the reduced-form effects of genetic diversity on political instability. In this regard, Acharya et al. (2016) document that including both the treatment variable (*Pdiv_aa*) and potentially mediating variables (*GDP per capita*) in one single model regression produces biased and inconsistent estimates due to intermediate variable bias. Following the empirical approach developed by Acharya et al. (2016), the current study provides a more thorough analysis of potential channels of influence in Section 6, thus permitting a causal interpretation of the underlying mechanisms.

4. Results

4.1. Main findings

Figure 4 depicts the relationship between genetic diversity and political instability across countries. The unconditional correlation, illustrated together with a nonparametric local polynomial regression line, is suggestive of a U-shaped link between genetic diversity and political instability, consistent with the main hypothesis. While the upward-sloping portion of the proposed U-shaped curve is populated by a larger number of countries, there is also evidence supporting the downward-sloping part. Moreover, the quadratic fit is bounded by the 95% confidence interval associated with the nonparametric regression line (Figure 4). However, these findings by no means permit a causal interpretation because of plausible concerns about omitted variables bias.

Table 1 reports OLS estimates of the effects of an ancestry-adjusted measure of predicted genetic diversity on political instability across 141 countries. Column (1) presents the unconditional estimates obtained from regressing the World Bank's indicator of political instability on genetic diversity and its squared term. Numerous geographic attributes are included in column (2). To account for unobserved heterogeneity across regions, I incorporate continent dummies in column (3). The full specification demonstrated in the last column of Table 1 is selected as the benchmark findings.

In all cases, the estimated coefficients of $Pdiv_{aa}$ are negative and statistically significant at conventionally accepted levels. The main results indicate that genetic diversity up to a certain threshold helps lower political instability, holding everything else constant. This lends credence to the main argument that genetically diverse societies tend to benefit from the availability of a wide range of personal traits, such as skills, abilities and cognitive approaches. The diversity of these individual-level traits is conducive to innovative activities and economic specialization, resulting in higher levels of productivity and income per capita (Ashraf & Galor, 2013, 2018; Depetris-Chauvin & Özak, 2020). This eventually translates into a more stable political environment by increasing the opportunity cost of engaging in riots and revolts. Moreover, higher degrees of productivity or income per capita driven by genetic diversity may shape the worldwide distribution of political unrest through affecting the provision of public goods, income redistribution and the ability to consolidate power. This is because high-income countries are typically endowed with better resources to provide public goods and services, and maintain social order. These factors help lower social dissatisfaction with the government. The results provide evidence of the upward-sloping portion of the proposed U-shaped curve.

The empirical evidence also reveals that highly diverse countries tend to suffer from the persistence of political instability, *ceteris paribus*. In particular, *Pdiv_aa_sqr* enters all the regressions with a positive and precisely estimated coefficient. This lends support to the hypothesized positive association between excessive diversity and political instability. A potential interpretation of these findings is that excessive levels of genetic diversity are associated with the under-provision of public goods and interpersonal mistrust, thus provoking social and political unrest. As argued earlier, high levels of diversity in personal traits, reflecting large heterogeneity in preferences for public goods and redistributive policies, make it more difficult for the government to reconcile such divergence. Moreover, highly diverse societies may suffer from lack of social capital, driven by divergence in values, beliefs, preferences and predispositions across individuals. Therefore, political instability tends to be more prevalent in countries characterized by excessive heterogeneity in the composition of genetic traits. Consistent with the main hypothesis, I find evidence that the relationship between genetic diversity and political instability follows a U-shaped pattern, in which both low and high levels of genetic diversity are conducive to political instability. By contrast, countries characterized by an intermediate level of genetic diversity are among the most politically stable economies in the world.

The main results help reconcile highly mixed findings offered by previous studies exploring the role of genetic diversity in shaping comparative development across the world. Ashraf and Galor (2013) pioneer in this line of inquiry by establishing an inverted U-shaped relationship between genetic diversity and income per capita. Several scholars find empirical support for the upward sloping portion of the inverted U-shaped curve of Ashraf and Galor (2013). Accordingly, genetic diversity is conducive to economic development across countries in the US (Ager & Brueckner, 2018), individual-level socioeconomic outcomes (Cook & Fletcher, 2018) and economic specialization (Depetris-Chauvin & Özak, 2020). Other studies demonstrate that genetic diversity lies at the deep roots of autocratic institutions (Galor & Klemp, 2017) and the onset of civil wars (Arbatlı et al., 2020), leading to persistent underdevelopment. This is consistent with the downward sloping part of the inverted U-shaped curve of Ashraf and Galor (2013). My findings advance the current literature by providing support for the non-monotonic effects of genetic diversity on political instability, consistent with the hypothesis proposed by Ashraf and Galor (2013, 2018).

Additionally, the contribution of genetic diversity to political instability is quantitatively sizeable. The estimated marginal effect of genetic diversity on contemporary political

instability equals $-321.040 + 462.882 \times Pdiv_aa$ (column 3, Table 1). This suggests that a 0.1-unit increase in genetic diversity is associated with a 1.5-unit increase in the World Bank's index of political instability at the mean value of genetic diversity (0.728). However, a 0.1-unit increase in genetic diversity leads to an approximately 3-unit decrease in political instability at the minimum value of genetic diversity (0.628). Furthermore, at the maximum value of genetic diversity (0.774), the marginal effect of a 0.1-unit increase in $Pdiv_aa$ equates to nearly a 3.7-unit increase in the World Bank's index of political instability. Overall, the results are suggestive of the economic and statistical significance of the contribution of prehistorically determined genetic diversity to political instability. The unconditional estimates in column (1) indicate that the predicted degree of political instability reaches a minimum at an optimal level of genetic diversity of 0.704. When geographic attributes are incorporated in the baseline model specification, the optimal value of genetic diversity decreases significantly to 0.696, as shown in column (2). The optimal degree of genetic diversity decreases marginally to 0.694 when exploiting the benchmark estimates with the full specification in column (3).

4.2. Alternative tests for the presence of a U-shaped relationship

The paper provides evidence of a U-shaped relationship between genetic diversity and political instability, based on incorporating the linear and quadratic terms of genetic diversity in the regressions. To provide further support for the main hypothesis, I perform additional checks for the presence of a non-monotonic link between genetic diversity and political instability.

Specifically, I undertake Ramsey's *RESET* test of functional form misspecification under the null hypothesis that the functional form of the model is appropriately specified (Table 1). In all cases, the results indicate that we fail to reject the null hypothesis. I also report Doornik-Hansen's test of normality of the error terms. The null hypothesis of a normally distributed error component is not rejected at conventionally accepted levels of significance (Table 1). These results lend support to the absence of model misspecification of the baseline model, thus providing a valid basis for statistical inference (Owen, 2017, 2018). Additionally, I exclude the squared term of genetic diversity from the regression model. The results, available on request, reveal that the monotonic relationship between $Pdiv_aa$ and PIS is statistically insignificant at conventionally accepted levels. The results of Ramsey's *RESET* test also suggest that the model without $Pdiv_aa_sqr$ does not pass the test of functional form misspecification. These findings provide additional evidence of a non-monotonic association between prehistorically determined genetic diversity and present-day political instability.

Furthermore, I conduct the U-shaped test developed by Lind and Mehlum (2010). The null hypothesis is that the underlying relationship between genetic diversity and political instability is either monotonic or inversely U-shaped. This is tested against the hypothesis that the proposed relationship follows a U-shaped pattern. This empirical exercise checks whether the estimated relationship is decreasing at low degrees of genetic diversity and increasing at high levels (Lind & Mehlum, 2010). The test is also informative when testing whether the estimated optimal value of genetic diversity falls within the data range (Lind & Mehlum, 2010). Employing the baseline estimates in Table 1, I test whether the joint null hypothesis:

$$H_0 = (\beta_1 + 2\beta_2 \times Pdiv_aa_{min} \geq 0) \cup (\beta_1 + 2\beta_2 \times Pdiv_aa_{max} \leq 0) \quad [2]$$

can be rejected in favour of the joint alternative hypothesis:

$$H_1 = (\beta_1 + 2\beta_2 \times Pdiv_aa_{min} < 0) \cup (\beta_1 + 2\beta_2 \times Pdiv_aa_{max} > 0) \quad [3]$$

where $Pdiv_aa_{min}$ and $Pdiv_aa_{max}$ represent the minimum and maximum values of the genetic diversity index, respectively. Rejecting the null hypothesis lends credence to the existence of a U-shaped relationship. As shown in Table 1, the null hypothesis is rejected at the 1% level of significance, except in column (2). This reveals that the underlying association between genetic diversity and political instability exhibits a U-shaped pattern, rather than an inverted U-shaped or linear relationship.

5. Robustness analyses

5.1. Robustness to endogeneity concerns

The inclusion of several geographic attributes in the benchmark model at least partially mitigates a key concern about omitted variables bias. Nevertheless, it is largely impossible to identify and account for all potentially confounding factors. Hence, the core results can still be driven by selection on unobserved factors, although this potential bias could be minimized by allowing numerous key determinants of political instability to be included in the regression.

To address endogeneity concerns induced by possible failure to control for a relevant factor, I create plausibly exogenous sources of variation in genetic diversity, following the empirical approach developed by Ashraf and Galor (2013, 2018), Depetris-Chauvin and Özak (2020) and Arbatlı et al. (2020). For this purpose, I employ a measure of migratory distance from East Africa of prehistorically indigenous population in each country ($Mdist$) as an instrumental variable for genetic diversity of present-day countries. The extent to which we can achieve causal inference rests upon satisfying two requirements. The first assumption relates to the relevance of the instrumental variables. As discussed previously, this requirement

is justified by the widely accepted theory of “out of Africa” proposed by Ashraf and Galor (2013), which lends support to the highly predictive power of *Mdist* for international differences in genetic diversity (see Section 2).

The exogeneity condition is valid when *Mdist* affects contemporary political instability exclusively through shaping heterogeneity in the composition of genetic traits within an economy. As put forward by Ashraf and Galor (2013), Arbatlı et al. (2020) and Depetris-Chauvin and Özak (2020), migratory distance from East Africa provides a plausibly exogenous source of variation in contemporary economic performance across countries. Using numerous robustness checks, Ashraf and Galor (2013) demonstrate that *Mdist* exerts no statistically significant influence on current economic development once partialling out the contribution of genetic diversity. On this basis, Ashraf and Galor (2013) postulate that geographic distance along the migratory routes of *Homo sapiens* from East Africa is only relevant for economic development through triggering the aforementioned “serial founder effect”. Therefore, the long-term legacy of *Mdist* for contemporary socio-economic development is exclusively driven by genetic heterogeneity. To the extent that *Mdist* has no direct influence on present-day economic performance, the exclusion restriction is unlikely to be violated. For this reason, I employ the linear and quadratic terms of *Mdist* as plausibly exogenous instruments for *Pdiv_aa* and *Pdiv_aa_sqr*, respectively.

Table 2 reports IV-2SLS estimates of the effects of genetic diversity on political instability. The first-stage results reveal that *Mdist* and its squared term (*Mdist_sqr*) are strongly correlated with potentially endogenous variables, including the linear and quadratic terms of genetic diversity. Moreover, the *F*-statistic values of excluded instruments are much bigger than the rule-of-thumb of 10 in all cases, lending support to the relevance of the instrumental variables. Consistent with the main hypothesis, the second-stage estimates demonstrate that the plausibly exogenous component of genetic diversity exerts U-shaped effects on political instability. The causal impacts are also precisely estimated at conventionally accepted levels of significance. However, the magnitude of coefficients turns out to be much larger than that of the baseline estimates.⁶ This suggests that genetic diversity has an

⁶ Recent studies find that the plausibly exogenous component of genetic diversity, generated by migratory distance from East Africa of a country’s indigenous population, has much larger effects on economic performance compared with conventional cross-country OLS estimates (Arbatlı et al., 2020; Depetris-Chauvin & Özak, 2020). This may be attributable to a possible downward bias in the OLS estimates. Nevertheless, the economic and statistical significance of the effects of genetic diversity on political instability remains largely robust to using either OLS or IV-2SLS estimates.

economically and statistically significant effect on the stability of a country's political environment after ruling out potential endogeneity concerns.

5.2. Robustness to additional controls

An additional strategy of identification relies on incorporating an extensive set of control variables in the benchmark regression to alleviate omitted variables bias.

First, I control for numerous factors that have been widely adopted to explain international differences in political instability (Table 3). As demonstrated earlier, the developing world tends to suffer from the prevalence of social and political unrest because people living in low-income economies typically face a lower opportunity cost of engaging in riots. Additionally, underdevelopment, manifested in low levels of income or productivity, entails poor governance and the under-provision of public goods, possibly leading to greater dissatisfaction with the government. I account for this effect by including the log of GDP per capita in column (1). Two different measures of institutional quality, including control of corruption and governance effectiveness, are added to columns (2) and (3). I account for the effects of the provision of public goods by including health and education expenditure as a proportion of total GDP in columns (4) and (5). An unequal distribution of income is a major driver of political instability. This warrants the inclusion of the Gini coefficient of disposable income inequality in column (6). The last column of Table 3 adds social trust to account for the confounding effects of interpersonal mistrust. The U-shaped relationship between genetic diversity and political instability remains precisely estimated at the 1% level of significance, except in column (7) of Table 3.⁷ Hence, the main findings are not easily explained away by conventional explanations for the worldwide distribution of political instability. It is noteworthy that the estimated impacts of these additional controls do not necessarily imply causality due to plausible concerns about reverse causation.

Second, I account for the effects of ethnolinguistic fractionalization on political uncertainty (Table 4). Ethnic fragmentation is a key determinant of violence and civil wars, which may affect the political environment and socioeconomic outcomes. One may well argue that my findings are purely attributed to cross-country differences in ethnic fragmentation. Hence, I replicate the baseline analysis by controlling for commonly used proxies for

⁷ In column (7) of Table 3, the estimated effects of the linear and squared terms of genetic diversity on political instability are statistically significant at the 10% level. However, these results are based on estimating a restricted sample comprising of only 85 countries. Unfortunately, data on social trust are sparse across countries, making it difficult to perform a comparable replication of the core findings.

population diversity, including ethnic and linguistic fractionalization (Alesina et al., 2003) and ethnolinguistic polarization (Desmet et al., 2012). Consistent with the main hypothesis, these variables enter the regression in a quadratic specification. As shown in Table 4, the benchmark estimates retain their signs and significance levels. By contrast, conventional measures of ethnic diversity exert no statistically significant effect on political instability. This is suggestive of the important role of genetic diversity, relative to ethnolinguistic fragmentation, in driving the variation in political instability across countries.

Third, the main regression is augmented with several historical confounders, as shown in columns (1) to (4) of Table 5. A well-established strand of literature documents that early development of societies helps explain substantial variation in economic prosperity across the globe. Some studies highlight the role of accumulated statehood experience (*Statehist*), obtained over thousands of years, in shaping GDP per capita, income distribution and the quality of institutions (Bockstette et al., 2002; Ang, 2013; Borcan et al., 2018; Vu, 2021b).⁸ Others find that the length of time elapsed since the transition to sedentary agriculture (*Agyears*) and the duration of human settlements (*Origtime*) are of importance for contemporary economic performance (Olsson & Hibbs, 2005; Ang, 2013). Acemoglu et al. (2002) argue that population density in 1500 (*Lpd1500*) is highly predictive of today's economic prosperity – the reversal of fortune hypothesis. The baseline results can be biased and inconsistent if these historical confounders matter for the evolution of genetic diversity and socio-political unrest. Moreover, an early start can shape political uncertainty through affecting productivity, income per capita, institutions and income inequality. Following Borcan et al. (2018) and Vu (2021b), I add the quadratic terms of these historical controls to columns (5) to (8) of Table 5.⁹ Nevertheless, my findings withstand accounting for historical confounders.

Finally, Table 6 re-estimates the benchmark model by incorporating other factors driving the global distribution of political instability. The historical event of European colonization, starting around the sixteenth century, imparts a persistent influence on comparative development across the world. Therefore, column (1) includes binary variables for legal origins. This helps account for the identity of colonizers characterized by different patterns (*quality*) of colonization. I also include a dummy for former colonies to account for differences

⁸ Statehood experience, captured by the state history index, reflects a country's experience with state-level institutions obtained over six millennia (Borcan et al., 2018). This is measured in the early presence of states, and states' autonomy and territorial coverage (Bockstette et al., 2002; Borcan et al., 2018).

⁹ Borcan et al. (2018) and Vu (2021b) provide suggestive evidence of a non-linear relationship between statehood experience and numerous economic outcomes, including income per capita, institutions and income inequality.

between colonized and never colonized societies. A recent study by Acemoglu et al. (2020) posits that an increase in population size is positively associated with violent conflicts over scarce resources. I account for this effect by adding population size and density to column (2). Consistent with the findings of Krieger and Meierrieks (2011), I rule out the confounding effects of democratic institutions in column (3). Other scholars put forward that the persistence of socio-political unrest is attributed to trade openness (Martin et al., 2008) and resource endowments (Dutt & Mitra, 2008). Hence, these confounding factors are added to columns (4) and (5). In the last two columns, I check whether the main findings are driven by urbanization and the shares of the population practicing major religions. Nevertheless, the inclusion of a wide range of possible confounders in the regression fails to explain away the U-shaped relationship between *Pdiv_aa* and *PIS*.

5.3. Robustness to using alternative dependent variables

In response to an anonymous reviewer's concern that the results can be exclusively driven by using the World Bank's index of political instability, this study also employs alternative dependent variables. To this end, the main regression is replicated in Table 7 by using six different measures of socio-political risks and political turnover, obtained from the International Country Risk Guide dataset (ICRG), as the outcome variables.¹⁰

The dependent variable in column (1) is an index of government stability, which reflects the extent to which the government can implement its declared programs and its probability of losing office. In column (2), the paper adopts a measure of internal conflicts. This variable represents the pervasiveness of political violence, such as civil wars, terrorism, coups and civil disorder. The outcome variables used in columns (3) and (4) capture the level of tensions within a society induced by religious, ethnic, racial or language divisions. Moreover, genetic diversity may affect socio-economic conditions by inducing social dissatisfaction with the government and constraining its ability to exercise power (e.g., unemployment, consumer confidence and poverty rates). The dependent variable adopted in column (5) reflects variation in broadly defined socio-political instability. In column (6), I use a measure of investment risks, including contract viability/expropriation, profit repatriation and payment delays. In order to estimate the cross-sectional models, data are averaged between 1985 and 2015. Except in column (4), the main results retain their signs and significance levels when using different proxies for socio-political instability (Table 7).

¹⁰ The online Appendix Table A3 contains a correlation matrix of different measures of political instability.

Furthermore, the current paper constructs a measure of political uncertainty, which is adopted as an alternative dependent variable in the last column of Table 7. The construction of this variable exploits information on effective political leaders for a wide range of economies spanning a prolonged period of years, provided by the Archigos dataset (Goemans et al., 2009). More specifically, the Archigos dataset of Goemans et al. (2009) contains information on the years of entry and exit of a country's effective leaders. A transition of leaders is regular if it is driven by prevailing rules, conventions, norms, natural death and voluntary retirement (Goemans et al., 2009). By contrast, to the extent that a political leader is overthrown by contravention of explicit and conventional rules, such as revolts, riots or assassinations, the transition is considered irregular (Goemans et al., 2009). In line with Grechyna (2018), an index of political turnover is computed by the number of irregular changes in a country's effective leaders as a proportion of total irregular and regular changes. Higher values of political turnover represent greater political uncertainty, captured by the prevalence of irregular changes in the national government. Using the Archigos dataset helps address a concern related to the possibility that the main results may be merely driven by international variation in political instability over a limited period from 1996 to 2015. The core findings, however, are insensitive to using a measure of political turnover as an alternative dependent variable.

5.4. Additional robustness tests

In order to conserve space, the online Appendix contains other sensitivity checks. The results indicate that the main findings are unlikely to be exclusively attributed to (1) spatial dependence of socio-political unrest across countries (Table A4); (2) possible outliers (Table A5); and (3) incorrectly specified functional forms (Table A6). In response to a request of an anonymous reviewer, the main analysis is replicated by (4) removing centralized economies (Table A7); (5) eliminating world regions (Table A8); and (6) using different sub-samples of countries (Table A9). However, the results retain their signs even when the sample size is restricted to specific groups of countries. Due to significant reduction in the feasible degrees of freedom, the statistical precision of the estimates reduces to the 10% level of significance in some cases. In short, the current paper consistently finds suggestive evidence of a U-shaped relationship between prehistorically determined genetic diversity and contemporary political instability across countries.

6. Further evidence

6.1. Genetic diversity and the emergence of socio-political unrest

This sub-section provides additional evidence of the hypothesized U-shaped relationship between prehistorically determined genetic diversity and contemporary political instability by using repeated cross-country data.

In particular, I employ a dichotomous index of the occurrence of riots and revolts across countries between 1960 and 2010, constructed by Acemoglu et al. (2019). This permits an empirical investigation of the role of genetic diversity in triggering the emergence of socio-political unrest over a long period of years. Following Arbatlı et al. (2020) and Ang and Gupta (2018), the empirical framework can be specified as follows.

$$Unrest_{it} = \beta_0 + \beta_1 Pdiv_aa_i + \beta_2 Pdiv_aa_sqr_i + \beta_3 X_i + \beta_4 Unrest_{it-1} + \beta_5 \theta_t + \varepsilon_i$$

where the dependent variable ($Unrest_{it}$) takes a value of one capturing the occurrence of riots and revolts for each country in a given year, and zero otherwise. The lagged dependent variable is incorporated in the regression model to account for temporal spillovers of socio-political unrest ($Unrest_{it-1}$). The model also accounts for unobserved time-specific factors by including year dummies (θ_t). The linear and quadratic terms of genetic diversity, country-level geographic controls and continent dummies are time-invariant. These variables are identical to those included in the baseline model.

Table 8 presents probit estimates of the contribution of prehistorically determined genetic diversity to the occurrence of riots and revolts in repeated cross-country data. Accordingly, $Pdiv_aa$ is negatively associated with the probability of observing the occurrence of socio-political unrest with precisely estimated coefficients. By contrast, $Pdiv_aa_sqr$ exerts positive and statistically significant effects on the likelihood of observing the onset of riots and revolts from 1960 to 2010. These results lend credence to the established U-shaped relationship between genetic diversity and political instability.

6.2. Potential channels of influence

In this sub-section, I explore potential pathways underlying the proposed relationship between genetic diversity and political instability (Table 7). Consistent with the main hypothesis articulated in Section 1, possible mechanisms include income or productivity levels ($lgdppc$), the provision of public goods ($healthexp$ and $eduexp$), income inequality ($gini$) and societal cohesiveness ($trust$). As suggested by Galor and Klemp (2017), prehistorically determined genetic diversity may translate into the persistence of political instability by shaping the quality

of institutions. This channel of influence can be captured by control of corruption (*cce*) or government effectiveness (*gee*). A conventional approach is to augment the main regression with possible mechanisms of transmission. To the extent that the inclusion of the potential pathways attenuates the benchmark estimates toward zero, it is possible to obtain inference on the channels through which genetic diversity transmits to political uncertainty. Nevertheless, Acharya et al. (2016) document that this widely adopted method provides an invalid basis for causal inference due to intermediate variables bias.

Following Acharya et al. (2016), this paper applies a two-step procedure to estimate the average controlled direct effects (ACDE). This approach plausibly permits a causal interpretation of the underlying pathways once accounting for the contribution of mediating variables to political uncertainty.¹¹ In particular, the ACDE estimates reflect the causal impacts of a treatment variable (*Pdiv_aa*) on political unrest holding a potentially mediating factor (*lgdppc*) fixed at a particular level. The first-stage regression requires estimating the effects of potential channels of influence on the outcome variable. Using the first-stage estimates, I transform the dependent variable by removing the effects of a mediator. The second-stage regression involves regressing the demediated dependent variable on the treatment variable, yielding ACDE estimates. If the ACDE estimates turn out to be indistinguishable from zero at conventionally accepted levels of significance, one may argue that the effects of the treatment variable on the dependent variable are exclusively driven by the proposed channel. By contrast, we may contend that the treatment variable transmits to the outcome variable through other channels besides the proposed pathway if the ACDE estimates remain statistically significant at conventionally accepted levels.

It can be observed from Table 9 that the ACDE coefficients are quantitatively larger than the baseline estimates in some cases. Nevertheless, the ACDE estimates remain statistically significant at conventionally accepted levels. There are two implications from these results. First, the benchmark results are not exclusively driven by these channels of transmission. Second, genetic diversity transmits to political unrest through other indirect or direct pathways, besides these proposed mechanisms. However, changes in the magnitude and significance levels of the coefficients reveal that the proposed mechanisms at least partially mediate the U-

¹¹ According to Acharya et al. (2016), there are two main advantages of estimating the ACDE estimates. *First*, this method allows us to assess potential mechanisms underlying the reduced-form relationship between *Pdiv_aa* and *PIS*, which is relevant for formulating policies. *Second*, it reduces potential bias induced by the inclusion of both treatment and intermediate variables in one single regression, thus permitting a causal interpretation. See Vu (2020) for an application of the ACDE estimates to explore potential channels of transmission.

shaped relationship between genetic diversity and political instability. This lends additional support to the central hypothesis of the current study.¹²

7. Concluding remarks

This study improves our understanding of the deep historical roots of political instability across countries. It empirically establishes that prehistorically determined genetic diversity exerts persistent and U-shaped effects on contemporary political instability. The central hypothesis is that genetic diversity, up to a certain point, helps establish politically stable systems by increasing the opportunity cost of engaging in riots and revolts. However, higher degrees of genetic diversity are associated with the persistence and pervasiveness of political instability by creating social non-cohesiveness and undermining the provision of public goods.

The current study proposes that both homogenous and excessively fragmented societies are characterized by the prevalence of socio-political unrest, *ceteris paribus*. By contrast, societies with an intermediate degree of genetic diversity typically enjoy low levels of uncertainty associated with their political regimes, holding everything else constant. Performing a wide range of cross-country analyses for 141 economies, this paper consistently obtains precise estimates indicating that prehistorically determined genetic diversity imparts a U-shaped influence on political instability. Additionally, it documents suggestive evidence of the underlying causal pathways through which genetic diversity translates into the persistence of political instability. Possible mechanisms of transmission include income/productivity levels, the provision of public goods, income inequality and social mistrust.

A potentially misleading interpretation of the results would be that establishing politically stable systems can be facilitated by, for example, forcible changes in the composition of genetic traits within current populations to achieve an optimal degree of diversity. By contrast, policymakers should recognize the long-term political legacy of prehistorically determined genetic diversity as part of policy prescriptions. It is important to

¹² As discussed in Section 1, a growing body of research attempts to investigate the effects of political instability on economic outcomes, including income/productivity levels, capital accumulation, public debt and financial development, among others. As suggested by an anonymous reviewer, the results of the current study may provide a motivation for using prehistorically determined genetic diversity to isolate a plausibly exogenous source of variation in political instability that helps explain cross-country differences in economic performance. The main results lend some credence to the relevance of this possible instrumental variable. However, the application of this potential empirical exercise in future studies requires attention to a major caveat regarding the validity of the exclusion restrictions. A key requirement is that genetic diversity imparts no direct influence on present-day economic development, except through shaping variation in political instability. Empirical attempts at creating an exogenous component of political instability based on its deep historical roots may provide an invalid basis for statistical inference should possible deviation from the exogeneity condition exist.

re-emphasize that the extent to which genetic diversity translates into political instability is partially mediated through numerous “proximate” factors, discussed in Section 6. These intervening mechanisms provide useful suggestions for formulating relevant policies that help attenuate the disadvantages of highly homogenous and fragmented societies.

For example, interpersonal mistrust and the under-provision of public goods, which remain enduring features of genetically diverse economies, can be attenuated by promoting civic engagement and reducing inefficiencies and distortions associated with the provision of public goods. Furthermore, excessively fragmented countries may overcome divergence in political preferences through enacting education policies geared toward fostering diversity tolerance. Given the lack of a diverse range of complementary skills, abilities and cognitive approaches in homogenous societies, addressing political unrest can be fostered by the formation of a culture of open-mindedness and the acceptance of novel ideas. Because history casts a long shadow on the political environment of present-day countries, effective policies arguably lie in their compatibility with the prevailing historical context.

This study exploits cross-country differences in genetic diversity and political instability to estimate numerous cross-sectional models. I attempt to mitigate omitted variables bias by accounting for a series of possibly confounding factors, none of which completely absorbs the statistical precision of the benchmark findings. Nevertheless, plausible concerns about selection on unobservables are far from trivial within a cross-country empirical framework. In this regard, a promising avenue for future research is to check whether the findings can be replicated using variation across regions within a country. Additionally, the adoption of subnational data would complement this research through exploring possible heterogeneity in the relationship between genetic diversity and socio-political unrest across the world. By doing so, we may also obtain a more comprehensive understanding of possible mechanisms underlying the main hypothesis of this paper.

References

- Acemoglu, D., Fergusson, L., & Johnson, S. (2020). Population and Conflict. *Review of Economic Studies*, 87(4), 1565-1604.
- Acemoglu, D., Johnson, S., & Robinson, J. A. (2002). Reversal of Fortune: Geography and Institutions in the Making of the Modern World Income Distribution. *Quarterly Journal of Economics*, 117(4), 1231-1294.
- Acemoglu, D., Naidu, S., Restrepo, P., & Robinson, J. A. (2019). Democracy Does Cause Growth. *Journal of Political Economy*, 127(1), 47-100.
- Acharya, A., Blackwell, M., & Sen, M. (2016). Explaining Causal Findings without Bias: Detecting and Assessing Direct Effects. *American Political Science Review*, 110(3), 512-529.
- Ager, P., & Brueckner, M. (2018). Immigrants' Genes: Genetic Diversity and Economic Development in the United States. *Economic Inquiry*, 56(2), 1149-1164.
- Aisen, A., & Veiga, F. J. (2008). The Political Economy of Seigniorage. *Journal of Development Economics*, 87(1), 29-50.
- Aisen, A., & Veiga, F. J. (2013). How Does Political Instability Affect Economic Growth? *European Journal of Political Economy*, 29, 151-167.
- Alesina, A., Baqir, R., & Easterly, W. (1999). Public Goods and Ethnic Divisions. *Quarterly Journal of Economics*, 114(4), 1243-1284.
- Alesina, A., Devleeschauwer, A., Easterly, W., Kurlat, S., & Wacziarg, R. (2003). Fractionalization. *Journal of Economic Growth*, 8(2), 155-194.
- Alesina, A., & Ferrara, E. L. (2005). Ethnic Diversity and Economic Performance. *Journal of Economic Literature*, 43(3), 762-800.
- Alesina, A., Gennaioli, C., & Lovo, S. (2019). Public Goods and Ethnic Diversity: Evidence from Deforestation in Indonesia. *Economica*, 86(341), 32-66.
- Alesina, A., Harnoss, J., & Rapoport, H. (2016). Birthplace Diversity and Economic Prosperity. *Journal of Economic Growth*, 21(2), 101-138.
- Alesina, A., & Perotti, R. (1996). Income Distribution, Political Instability, and Investment. *European Economic Review*, 40(6), 1203-1228.
- Alesina, A., & Tabellini, G. (1990). A Positive Theory of Fiscal Deficits and Government Debt. *Review of Economic Studies*, 57(3), 403-414.
- Ang, J. B. (2013). Institutions and the Long-Run Impact of Early Development. *Journal of Development Economics*, 105, 1-18.
- Ang, J. B., & Gupta, S. K. (2018). Agricultural Yield and Conflict. *Journal of Environmental Economics and Management*, 92, 397-417.
- Arbath, C. E., Ashraf, Q. H., Galor, O., & Klemp, M. (2020). Diversity and Conflict. *Econometrica*, 88(2), 727-797.
- Ashraf, Q., & Galor, O. (2013). The 'Out of Africa' Hypothesis, Human Genetic Diversity, and Comparative Economic Development. *American Economic Review*, 103(1), 1-46.
- Ashraf, Q., & Galor, O. (2018). The Macrogenoeconomics of Comparative Development. *Journal of Economic Literature*, 56(3), 1119-1155.
- Azzimonti, M. (2011). Barriers to Investment in Polarized Societies. *American Economic Review*, 101(5), 2182-2204.

- Banerjee, A., Iyer, L., & Somanathan, R. (2005). History, Social Divisions, and Public Goods in Rural India. *Journal of the European Economic Association*, 3(2-3), 639-647.
- Battaglini, M., & Coate, S. (2008). A Dynamic Theory of Public Spending, Taxation, and Debt. *American Economic Review*, 98(1), 201-236.
- Blattman, C., & Miguel, E. (2010). Civil War. *Journal of Economic Literature*, 48(1), 3-57.
- Bockstette, V., Chanda, A., & Putterman, L. (2002). States and Markets: The Advantage of an Early Start. *Journal of Economic Growth*, 7(4), 347-369.
- Borcan, O., Olsson, O., & Putterman, L. (2018). State History and Economic Development: Evidence from Six Millennia. *Journal of Economic Growth*, 23(1), 1-40.
- Bove, V., & Elia, L. (2017). Migration, Diversity, and Economic Growth. *World Development*, 89, 227-239.
- Campante, F. R., Chor, D., & Do, Q.-A. (2009). Instability and the Incentives for Corruption. *Economics & Politics*, 21(1), 42-92.
- Carmignani, F. (2003). Political Instability, Uncertainty and Economics. *Journal of Economic Surveys*, 17(1), 1-54.
- Collier, P., & Hoeffler, A. (2004). Greed and Grievance in Civil War. *Oxford Economic Papers*, 56(4), 563-595.
- Collier, P., & Hoeffler, A. (2007). Civil War. In T. Sandler & K. Hartley (Eds.), *Handbook of Defense Economics* (Vol. 2, pp. 711-739). Amsterdam, The Netherlands: Elsevier.
- Collier, P., & Sambanis, N. (2002). Understanding Civil War: A New Agenda. *Journal of Conflict Resolution*, 46(1), 3-12.
- Cook, C. J., & Fletcher, J. M. (2018). High-School Genetic Diversity and Later-Life Student Outcomes: Micro-Level Evidence from the Wisconsin Longitudinal Study. *Journal of Economic Growth*, 23(3), 307-339.
- Cukierman, A., Edwards, S., & Tabellini, G. (1992). Seigniorage and Political Instability. *American Economic Review*, 82(3), 537-555.
- de Soysa, I., & Almås, S. (2019). Does Ethnolinguistic Diversity Preclude Good Governance? A Comparative Study with Alternative Data, 1990-2015. *Kyklos*, 72(4), 604-636.
- Depetris-Chauvin, E., & Özak, Ö. (2020). The Origins of the Division of Labor in Pre-Industrial Times. *Journal of Economic Growth*, 25(3), 297-340.
- Desmet, K., Ortuño-Ortín, I., & Wacziarg, R. (2012). The Political Economy of Linguistic Cleavages. *Journal of Development Economics*, 97(2), 322-338.
- Dutt, P., & Mitra, D. (2008). Inequality and the Instability of Polity and Policy. *Economic Journal*, 118(531), 1285-1314.
- Easterly, W., & Levine, R. (1997). Africa's Growth Tragedy: Policies and Ethnic Divisions. *Quarterly Journal of Economics*, 112(4), 1203-1250.
- Esteban, J., Mayoral, L., & Ray, D. (2012). Ethnicity and Conflict: An Empirical Study. *American Economic Review*, 102(4), 1310-1342.
- Esteban, J., & Ray, D. (2011). Linking Conflict to Inequality and Polarization. *American Economic Review*, 101(4), 1345-1374.
- Ezcurra, R., & Palacios, D. (2016). Terrorism and Spatial Disparities: Does Interregional Inequality Matter? *European Journal of Political Economy*, 42, 60-74.
- Fearon, J. D., & Laitin, D. D. (2003). Ethnicity, Insurgency, and Civil War. *American Political Science Review*, 97(1), 75-90.

- Fredriksson, P. G., & Svensson, J. (2003). Political Instability, Corruption and Policy Formation: The Case of Environmental Policy. *Journal of Public Economics*, 87(7), 1383-1405.
- Fredriksson, P. G., & Wollscheid, J. R. (2014). Political Institutions, Political Careers and Environmental Policy. *Kyklos*, 67(1), 54-73.
- Galor, O., & Klemp, M. (2017). Roots of Autocracy. NBER Working Paper Series No. 23301.
- Gisselquist, R. M., Leiderer, S., & Niño-Zarazúa, M. (2016). Ethnic Heterogeneity and Public Goods Provision in Zambia: Evidence of a Subnational “Diversity Dividend”. *World Development*, 78, 308-323.
- Goemans, H. E., Gleditsch, K. S., & Chiozza, G. (2009). Introducing Archigos: A Dataset of Political Leaders. *Journal of Peace Research*, 46(2), 269-283.
- Goldstone, J. A. (2002). Population and Security: How Demographic Change Can Lead to Violent Conflict. *Journal of International Affairs*, 56(1), 3-21.
- Gören, E. (2014). How Ethnic Diversity Affects Economic Growth. *World Development*, 59, 275-297.
- Grafton, R. Q., & Knowles, S. (2004). Social Capital and National Environmental Performance: A Cross-Sectional Analysis. *Journal of Environment & Development*, 13(4), 336-370.
- Grafton, R. Q., Knowles, S., & Owen, P. D. (2002). Social Divergence and Productivity: Making a Connection. In A. Sharpe, F. St-Hilaire, & K. Banting (Eds.), *The Review of Economic Performance and Social Progress 2002: Towards a Social Understanding of Productivity* (pp. 203-223). The Institutute for Research on Public Policy, Montreal, and Centre for the Study of Living Standards, Ottawa.
- Grafton, R. Q., Kompas, T., & Owen, P. D. (2007). Bridging the Barriers: Knowledge Connections, Productivity and Capital Accumulation. *Journal of Productivity Analysis*, 28(3), 219-231.
- Grechyna, D. (2018). Shall We Riot Too? The Geographical Neighbor Impact on Political Instability. *Kyklos*, 71(4), 581-612.
- Hegre, H., & Sambanis, N. (2006). Sensitivity Analysis of Empirical Results on Civil War Onset. *Journal of Conflict Resolution*, 50(4), 508-535.
- Jong-A-Pin, R. (2009). On the Measurement of Political Instability and Its Impact on Economic Growth. *European Journal of Political Economy*, 25(1), 15-29.
- Julio, B., & Yook, Y. (2012). Political Uncertainty and Corporate Investment Cycles. *Journal of Finance*, 67(1), 45-83.
- Krieger, T., & Meierrieks, D. (2011). What Causes Terrorism? *Public Choice*, 147(1/2), 3-27.
- Lind, J. T., & Mehlum, H. (2010). With or without U? The Appropriate Test for a U-Shaped Relationship. *Oxford Bulletin of Economics and Statistics*, 72(1), 109-118.
- Martin, P., Mayer, T., & Thoenig, M. (2008). Civil Wars and International Trade. *Journal of the European Economic Association*, 6(2-3), 541-550.
- Michalopoulos, S. (2012). The Origins of Ethnolinguistic Diversity. *American Economic Review*, 102(4), 1508-1539.
- Montalvo, J. G., & Reynal-Querol, M. (2005a). Ethnic Diversity and Economic Development. *Journal of Development Economics*, 76(2), 293-323.

- Montalvo, J. G., & Reynal-Querol, M. (2005b). Ethnic Polarization, Potential Conflict, and Civil Wars. *American Economic Review*, 95(3), 796-816.
- Nunn, N. (2020). The Historical Roots of Economic Development. *Science*, 367, 9986.
- Olsson, O., & Hibbs, D. A. (2005). Biogeography and Long-Run Economic Development. *European Economic Review*, 49(4), 909-938.
- Østby, G. (2008). Polarization, Horizontal Inequalities and Violent Civil Conflict. *Journal of Peace Research*, 45(2), 143-162.
- Owen, P. D. (2017). Evaluating Ingenious Instruments for Fundamental Determinants of Long-Run Economic Growth and Development. *Econometrics*, 5(3), 1-33.
- Owen, P. D. (2018). Replication to Assess Statistical Adequacy. *Economics: The Open-Access, Open-Assessment E-Journal*, 12, 1-16.
- Persson, T., & Svensson, L. E. O. (1989). Why a Stubborn Conservative Would Run a Deficit: Policy with Time-Inconsistent Preferences. *Quarterly Journal of Economics*, 104(2), 325-345.
- Putterman, L., & Weil, D. N. (2010). Post-1500 Population Flows and the Long-Run Determinants of Economic Growth and Inequality. *Quarterly Journal of Economics*, 125(4), 1627-1682.
- Roe, M. J., & Siegel, J. I. (2011). Political Instability: Effects on Financial Development, Roots in the Severity of Economic Inequality. *Journal of Comparative Economics*, 39(3), 279-309.
- Spolaore, E., & Wacziarg, R. (2013). How Deep Are the Roots of Economic Development? *Journal of Economic Literature*, 51(2), 325-369.
- Vu, T. V. (2020). Economic Complexity and Health Outcomes: A Global Perspective. *Social Science & Medicine*, 265, 113480.
- Vu, T. V. (2021a). State History and Political Instability: The Disadvantage of Early State Development. University of Otago Economics Discussion Papers No. 2101.
- Vu, T. V. (2021b). Statehood Experience and Income Inequality: A Historical Perspective. *Economic Modelling*, 94, 415-429.
- Yared, P. (2010). Politicians, Taxes and Debt. *Review of Economic Studies*, 77(2), 806-840.

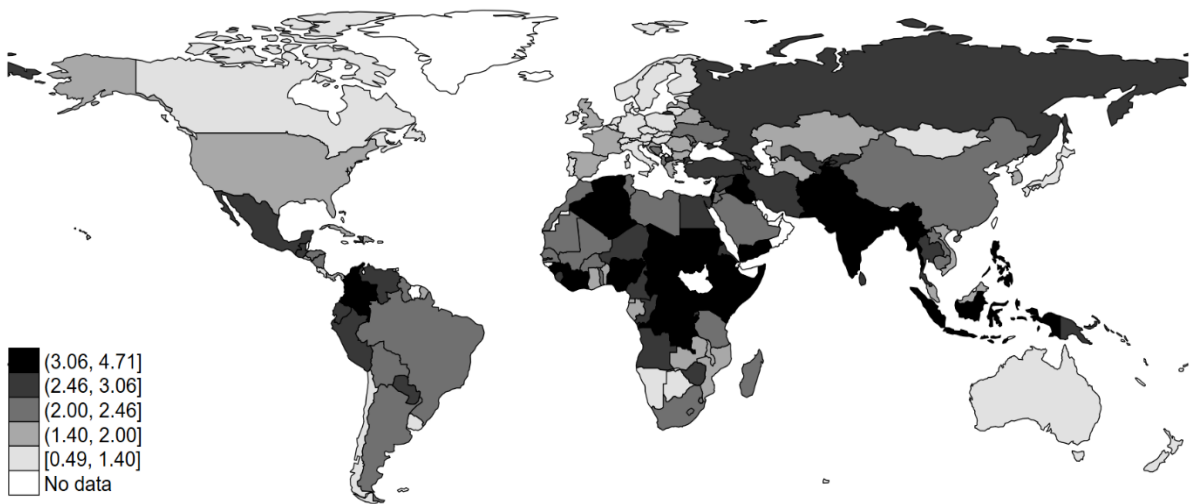


Figure 1. The worldwide distribution of political instability

Notes: This figure depicts international variation in the political instability index, constructed using the World Bank's index of Political Stability and Absence of Violence/Terrorism. Darker areas represent more politically unstable economies.

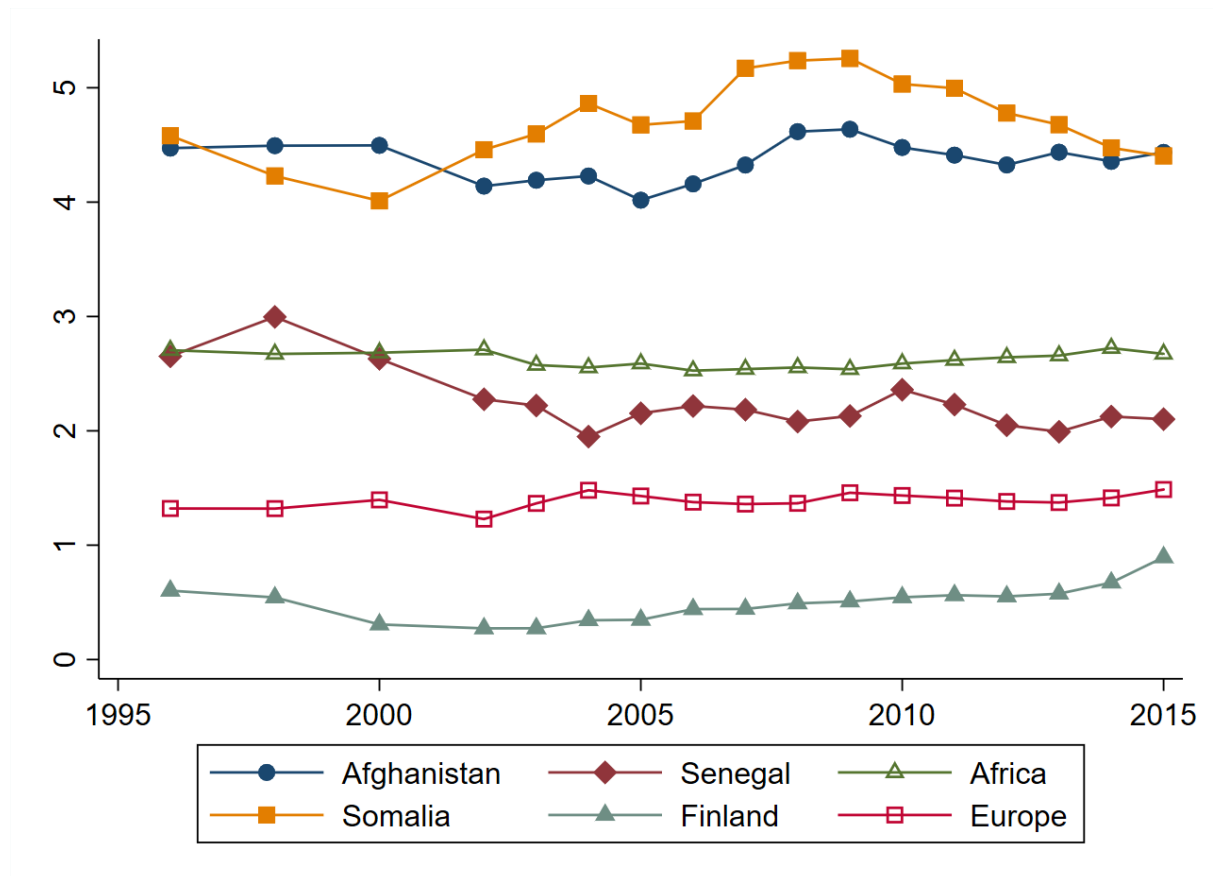


Figure 2. The evolution of political instability within selected countries and regions

Notes: This figure depicts variation in the political instability index within several countries and regions between 1996 and 2015. I present data for countries with the highest, lowest and medium levels of political instability. Additionally, I illustrate regional averages of political instability for African and European countries. The time series data for other countries and regions, not reported for brevity, exhibit broadly similar patterns. See also the notes to Figure 1.

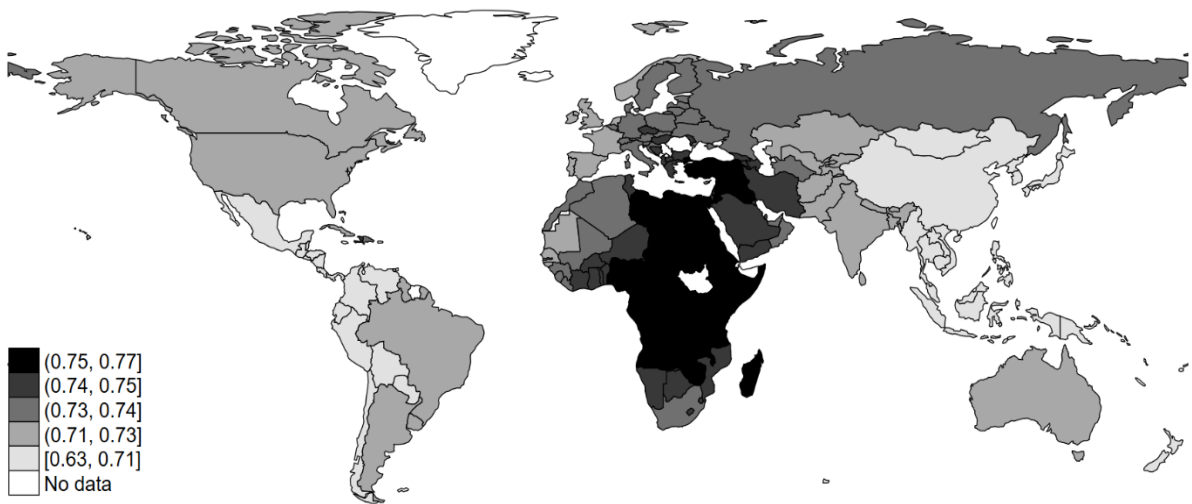


Figure 3. The worldwide distribution of genetic diversity

Notes: This figure depicts international variation in prehistorically determined genetic diversity, measured by an ancestry-adjusted index of predicted genetic diversity of Ashraf and Galor (2013). Darker areas represent countries with greater heterogeneity in the composition of genetic traits.

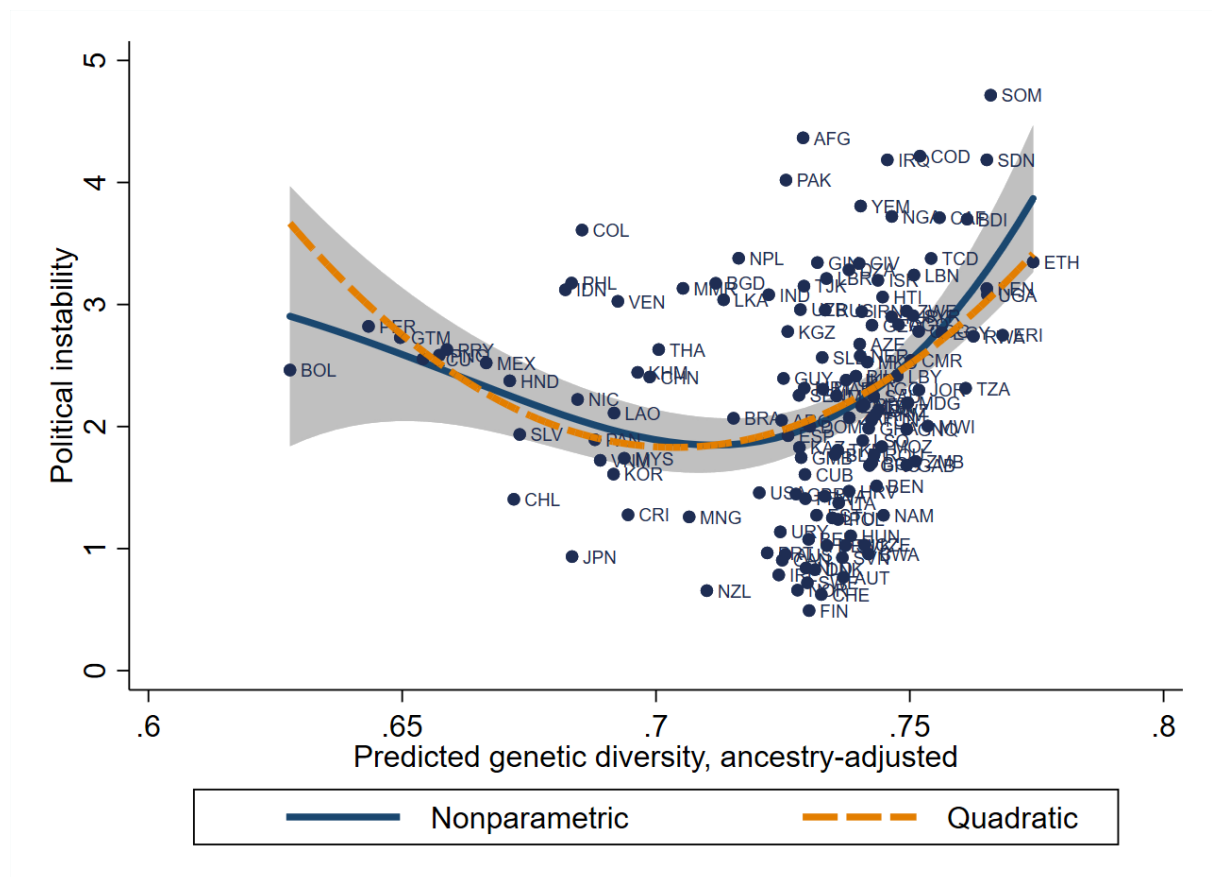


Figure 4. The relationship between genetic diversity and political instability

Notes: This figure is a scatter plot of the unconditional correlation between prehistorically determined genetic diversity and present-day political instability across 141 countries. It also depicts a nonparametric regression line, based on kernel-weighted local polynomial smoothing. The shaded area is the 95% confidence interval of the nonparametric estimates. See also the notes to Figures 1 and 3.

Table 1. The effects of genetic diversity on political instability, main results

	(1)	(2)	(3)
Dep_var: <i>PIS</i>	Unconditional estimates	Incorporating geographic controls	Incorporating continent dummies (Full specification)
<i>Pdiv_aa</i>	-448.935*** [102.472]	-210.204** [90.716]	-321.040*** [94.484]
<i>Pdiv_aa_sqr</i>	318.947*** [72.264]	151.105** [64.433]	231.441*** [67.050]
Terrain ruggedness		0.023 [0.103]	0.018 [0.098]
Mean elevation		0.498** [0.228]	0.290 [0.238]
Mean land suitability for agriculture		-0.106 [0.289]	-0.117 [0.297]
Distance to the nearest waterways		0.055*** [0.019]	0.044* [0.023]
Malaria (% of population)		-0.176 [0.231]	-0.025 [0.234]
Latitude		1.113*** [0.320]	0.661* [0.369]
Precipitation		0.013 [0.017]	0.014 [0.016]
Temperature		0.086*** [0.011]	0.068*** [0.012]
Island dummy		-0.086 [0.267]	-0.251 [0.247]
Optimal <i>Pdiv_aa</i>	0.704	0.696	0.694
Continent dummies	No	No	Yes
Observations (# of countries)	141	141	141
R-squared	0.136	0.516	0.572
RESET	0.016	0.832	0.998
Normality	0.039	0.226	0.347
U-shaped test	0.000	0.022	0.002

Notes: This table reports OLS estimates of the effects of an ancestry-adjusted measure of predicted genetic diversity (*Pdiv_aa*) on political instability (*PIS*). An intercept, omitted for brevity, is included in all regressions. *RESET* reports *p*-values of Ramsey's test for functional form misspecification under the null hypothesis that functional form of the model is appropriately specified. *Normality* reports *p*-values of Doornik-Hansen's test for the normal distribution of the error terms under the null hypothesis that they are normally distributed. U-shaped test reports *p*-values of the test for a U-shaped relationship developed by Lind and Mehlum (2010). The null hypothesis is that the relationship is monotonic or inversely U-shaped. Robust standard errors in parentheses. *** *p*<0.01, ** *p*<0.05, * *p*<0.1.

Table 2. Robustness to endogeneity concerns, IV-2SLS estimates

	Second-stage estimates	First-stage estimates	
	(1)	(2)	(3)
Dep_var	<i>PIS</i>	<i>Pdiv_aa</i>	<i>Pdiv_aa_sqr</i>
<i>Pdiv_aa</i>	-1,006.288* [514.812]		
<i>Pdiv_aa_sqr</i>	705.311** [354.946]		
<i>Mdist</i>		-0.010*** [0.001]	-0.015*** [0.001]
<i>Mdist_sqr</i>		0.0002*** [0.0001]	0.0004*** [0.0001]
Geographic controls	Yes	Yes	Yes
Continent dummies	Yes	Yes	Yes
First-stage <i>F</i> -statistic		74.34	80.19
Observations	141	141	141
R-squared	0.298	0.794	0.806
U-shaped test	0.042		

Notes: This table reports IV-2SLS estimates of the effects of *Pdiv_aa* on *PIS*. *Mdist* and *Mdist_sqr* are the linear and quadratic terms of migratory distance of a country's indigenous population from East Africa. The values of the first-stage *F*-statistic of excluded instruments provide evidence of the relevance of the instrumental variables. Robust standard errors in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. See also the notes to Table 1.

Table 3. Robustness to controlling for “proximate” causes of political instability

Dep_var: <i>PIS</i>	(1)	(2)	(3)	(4)	(5)	(6)	(7)
<i>Pdiv_aa</i>	-288.164*** [86.631]	-214.492*** [75.481]	-234.673*** [81.107]	-252.616*** [80.350]	-262.677*** [90.635]	-333.975*** [97.683]	-226.485* [127.033]
<i>Pdiv_aa_sqr</i>	208.440*** [61.559]	155.160*** [53.872]	169.010*** [57.741]	183.449*** [57.097]	190.144*** [64.117]	241.022*** [69.293]	161.657* [90.364]
GDP per capita, logged	-0.041 [0.037]						
Control of corruption		-0.499*** [0.070]					
Government effectiveness			-0.481*** [0.081]				
Health expenditure (% of GDP)				-0.052* [0.031]			
Education expenditure (% of GDP)					-0.110*** [0.039]		
Income inequality						-0.249 [1.411]	
Social trust							-1.680** [0.662]
Geographic controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Continent dummies	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	133	139	139	139	129	139	85
R-squared	0.579	0.700	0.687	0.587	0.619	0.576	0.655
RESET	0.661	0.015	0.277	0.931	0.680	0.955	0.109
Normality	0.751	0.804	0.423	0.615	0.458	0.384	0.705
U-shaped test	0.004	0.008	0.006	0.007	0.012	0.002	0.048

Notes: This table replicates the main analysis by controlling for the “proximate” determinants of *PIS*. Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1. See also the notes to Table 1.

Table 4. Robustness to controlling for conventional measures of ethnic fragmentation

Dep_var: <i>PIS</i>	(1)	(2)	(3)	(4)
	<i>Efrac</i>	<i>Des_pol</i>	<i>Lfrac</i>	<i>Full controls</i>
<i>Pdiv_aa</i>	-306.986*** [100.022]	-316.571*** [98.383]	-324.285*** [98.271]	-302.253*** [107.215]
<i>Pdiv_aa_sqr</i>	221.185*** [70.779]	228.560*** [69.769]	233.997*** [69.746]	218.066*** [75.650]
<i>Efrac</i>	0.075 [0.861]			-0.107 [1.255]
<i>Efrac_sqr</i>	0.650 [1.055]			1.117 [1.537]
<i>Des_pol</i>		0.541 [0.764]		-0.008 [0.993]
<i>Des_pol_sqr</i>		-0.425 [0.938]		-0.283 [1.055]
<i>Lfrac</i>			0.248 [0.781]	1.061 [1.258]
<i>Lfrac_sqr</i>			-0.026 [0.883]	-1.306 [1.428]
Geographic controls	Yes	Yes	Yes	Yes
Continent dummies	Yes	Yes	Yes	Yes
Observations	140	141	137	136
R-squared	0.584	0.575	0.577	0.591
RESET	0.992	0.999	0.972	0.865
Normality	0.303	0.359	0.369	0.360
U-shaped test	0.006	0.004	0.003	0.013

Notes: This table replicates the main analysis by controlling for conventional measures of ethnic fragmentation, including *Efrac* (ethnic fractionalization), *Des_pol* (ethnolinguistic polarization) and *Lfrac* (linguistic fractionalization). Consistent with the main hypothesis, I allow these variables to enter the benchmark model in a quadratic form. The results, not reported for brevity, remain insensitive to excluding the quadratic terms of these additional controls. Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1. See also the notes to Table 1.

Table 5. Robustness to controlling for the long-term legacy of early development

Dep_var: <i>PIS</i>	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
<i>Pdiv_aa</i>	-264.689*** [99.441]	-267.578*** [100.218]	-322.706*** [93.935]	-303.668*** [94.804]	-269.523*** [98.425]	-265.268*** [98.429]	-329.165*** [96.586]	-251.634*** [94.481]
<i>Pdiv_aa_sqr</i>	191.830*** [70.480]	193.000*** [71.398]	233.302*** [66.612]	219.430*** [67.187]	194.678*** [69.860]	192.218*** [70.358]	237.975*** [68.513]	183.148*** [66.863]
<i>Statehist</i>	0.813* [0.482]				-0.349 [1.626]			
<i>Agyears</i>		0.125*** [0.045]				0.227* [0.123]		
<i>Origtime</i>			-0.522** [0.209]				-0.142 [0.540]	
<i>Lpd1500</i>				0.065 [0.062]				0.144** [0.070]
<i>Statehist_sqr</i>					1.782 [2.375]			
<i>Agyears_sqr</i>						-0.009 [0.011]		
<i>Origtime_sqr</i>							-0.271 [0.331]	
<i>Lpd1500_sqr</i>								-0.038** [0.018]
Geographic controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Continent dummies	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	141	139	141	141	141	139	141	141
R-squared	0.580	0.602	0.584	0.576	0.582	0.605	0.585	0.589
RESET	0.900	0.641	0.952	0.898	0.981	0.630	0.943	0.938
Normality	0.248	0.293	0.201	0.305	0.216	0.235	0.206	0.423
U-shaped test	0.018	0.011	0.003	0.005	0.013	0.012	0.003	0.024

Notes: In columns (1) to (4), I replicate the main analysis by controlling for the long-term effects of early development, using an index of state history (*Statehist*), the length of time elapsed since the Neolithic revolution (*Agyears*), the duration of human settlements (*Origtime*) and the log of population density in 1500CE (*Lpd1500*). In columns (5) to (8), I incorporate the squared term of these additional controls in the regression. Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1. See also the notes to Table 1.

Table 6. Robustness to controlling for other confounding factors

Dep_var: <i>PIS</i>	(1)	(2)	(3)	(4)	(5)	(6)	(7)
<i>Pdiv_aa</i>	-287.558*** [84.504]	-321.822*** [91.027]	-296.729*** [94.961]	-271.476*** [93.357]	-322.275*** [96.142]	-382.532*** [110.151]	-329.801*** [94.108]
<i>Pdiv_aa_sqr</i>	208.796*** [59.920]	232.637*** [64.723]	213.817*** [67.476]	194.624*** [66.413]	232.143*** [68.220]	273.645*** [77.948]	236.523*** [66.725]
<i>Legal tradition</i>	Yes						
<i>Colony</i>	Yes						
<i>Population</i>		Yes					
<i>Democracy</i>			Yes				
<i>Resource</i>				Yes			
<i>Trade</i>					Yes		
<i>Urbanization</i>						Yes	
<i>Religion</i>							Yes
Geographic controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Continent dummies	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	138	140	139	141	140	117	140
R-squared	0.586	0.600	0.585	0.596	0.579	0.595	0.617
RESET	0.936	0.584	0.907	0.639	0.963	0.994	0.828
Normality	0.533	0.196	0.208	0.304	0.169	0.133	0.439
U-shaped test	0.005	0.002	0.004	0.005	0.002	0.001	0.001

Notes: This table replicates the main analysis by controlling for numerous confounding factors. *Legal tradition* is binary variables for common law and mixed law, with civil law being excluded as the base group. *Colony* is a dummy variable for countries with colonial history. *Population* denotes the size of population and population density. *Democracy* is the Polity IV index of democratic institutions. *Resource* represents four variables of resource endowments, including oil, gas, mineral and forest rents as a proportion of total GDP. *Trade* is the sum of exports and imports as a proportion of total GDP. *Urbanization* is the fraction of urban population. *Religion* captures three variables of the proportions of the population practicing major religions, including Catholics, Muslims and Protestants. The estimated coefficients of these additional controls are not reported to conserve space. Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1. See also the notes to Table 1.

Table 7. Robustness to using alternative dependent variables

Dep_var	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	<i>Govinst</i>	<i>Iconflict</i>	<i>Etensions</i>	<i>Rtensions</i>	<i>Sconditions</i>	<i>Irisks</i>	<i>Polturnover</i>
<i>Pdiv_aa</i>	-308.968** [145.526]	-777.315*** [225.262]	-542.687*** [145.013]	-99.429 [134.948]	-595.149*** [190.607]	-364.450* [216.853]	-66.578** [29.791]
<i>Pdiv_aa_sqr</i>	218.621** [103.980]	555.767*** [160.403]	386.563*** [103.103]	76.571 [96.671]	432.883*** [136.227]	265.961* [154.476]	46.417** [21.398]
Geographic controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Continent dummies	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	120	120	120	120	120	120	119
R-squared	0.301	0.600	0.401	0.563	0.569	0.473	0.423
RESET	0.097	0.674	0.629	0.0001	0.021	0.481	0.007
Normality	0.061	0.086	0.012	0.141	0.316	0.000	0.995
U-shaped test	0.033	0.001	0.0003	0.408	0.006	0.098	0.065

Notes: This table replicates the main analysis by using alternative dependent variables, including government instability (*Govinst*), internal conflict (*Iconflict*), ethnic tensions (*Etensions*), religious tensions (*Rtensions*), socio-economic conditions (*Sconditions*), investment risks (*Irisks*) and an index of political turnover (*Polturnover*). Higher values of these outcome variables reflect greater political instability. Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1. See also the notes to Table 1.

Table 8. Genetic diversity and the occurrence of unrest in repeated cross-country data

Dep_var: <i>Unrest</i>	(1)	(2)	(3)	(4)
<i>Pdiv_aa</i>	-102.210***	-87.209***	-118.368***	-122.772***
	[26.321]	[32.443]	[33.351]	[33.991]
<i>Pdiv_aa_sqr</i>	71.041***	61.102***	83.640***	86.806***
	[18.562]	[23.046]	[23.691]	[24.142]
Controls for temporal spillovers	Yes	Yes	Yes	Yes
Geographic controls	No	Yes	Yes	Yes
Continent dummies	No	No	Yes	Yes
Year dummies	No	No	No	Yes
Observations	4,436	4,436	4,436	4,436
# of countries	105	105	105	105
Pseudo <i>R</i> -squared	0.141	0.147	0.151	0.172

Notes: This table reports Probit estimates of the effects of *Pdiv_aa* on the incidence of socio-political unrest (*Unrest*), using repeated cross-country data from 1960 to 2010. *Unrest* is a binary variable capturing the onset of riots and revolts constructed by Acemoglu et al. (2019). The lagged dependent variable is incorporated in all regressions to account for the path dependence and temporal spillovers of socio-political unrest. The regression is also augmented with year dummies to account for unobserved year-specific factors. The set of geographic controls and continent dummies is similar to those included in Table 1. Robust standard errors in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. See also the notes to Table 1.

Table 9. Evidence on potential channels of transmission, ACDE estimates

Mediating variables	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	<i>lgdppc</i>	<i>cce</i>	<i>gee</i>	<i>healthexp</i>	<i>eduexp</i>	<i>gini</i>	<i>trust</i>
<i>Pdiv_aa</i>	-340.67***	-260.52**	-331.79***	-342.29***	-317.45***	-335.55***	-324.90**
[95% bootstrapped CI]	[-554.61, -126.74]	[-495.12, -25.93]	[-551.09, -112.51]	[-568.80, -115.80]	[-531.19, -103.70]	[-554.45, -116.66]	[-636.84, -12.96]
<i>Pdiv_aa_sqr</i>	245.62***	188.28**	239.05***	246.11***	228.86***	241.71***	233.19**
[95% bootstrapped CI]	[94.23, 397.02]	[21.07, 355.49]	[83.89, 394.21]	[86.15, 406.06]	[77.65, 380.08]	[86.83, 396.58]	[12.16, 454.22]
Geographic controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Continent dummies	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Notes: This table reports the average controlled direct effects (ACDE) of *Pdiv_aa* on *PIS* when accounting for the impacts of potentially mediating factors. The results are produced using a two-stage regression procedure developed by Acharya et al. (2016). Potentially mediating variables include the log of GDP per capita (*lgdppc*), control of corruption (*cce*), government effectiveness (*gee*), health (*healthexp*) and education (*eduexp*) expenditure as a proportion of total GDP, the Gini coefficient of inequality of disposable income (*gini*) and social trust (*trust*). These mediating variables are similar to those included in Table 3. Following Acharya et al. (2016), all regressions are performed via a bootstrapped method using 1000 replications. Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1. See also the notes to Table 1.