INTERPERSONAL DIVERSITY AND SOCIOECONOMIC DISPARITIES ACROSS POPULATIONS: A REPLY TO ROSENBERG AND KANG

QUAMRUL H. ASHRAF, ODED GALOR, AND MARC KLEMP*

Errare humanum est, perseverare diabolicum
- Seneca

Abstract

The exploration of the impact of the prehistoric migration of anatomically modern humans from Africa on comparative economic development has been the focus of a vibrant research agenda in the past decade. This influential literature has attracted the attention of scholars from other disciplines, and in light of existing methodological gaps across fields, it has perhaps unsurprisingly generated some significant misconceptions. In particular, Rosenberg and Kang (2015) suggest that the hump-shaped effect of interpersonal population diversity on population density in the year 1500 is statistically insignificant in an extended sample of genetic diversity that was released more recently. Unfortunately, this assertion is based on elementary statistical errors. In fact, the hump-shaped effect of diversity on population density is even more pronounced in this extended sample of Pemberton et al. (2013), and it is present not only in the year 1500 but over the entire pre-colonial period for which population data are available (i.e., the 10,000BCE to 1500CE timeframe).

JEL classification codes: O11, N10, N30, Z10

Keywords: Comparative development, interpersonal population diversity, the out of Africa hypothesis

*Ashraf: Department of Economics, Williams College (email: Quamrul.H.Ashraf@williams.edu); Galor: Brown University, NBER, CEPR, IZA and CeSifo (email: Oded_Galor@brown.edu); Klemp: University of Copenhagen and CEPR (email: Marc.Klemp@econ.ku.dk).

1 Introduction

The past two decades have witnessed the emergence of a large and influential body of research that has focused on uncovering the roots of comparative economic development across regions, countries, and ethnic groups. A significant portion of this line of inquiry has explored the role of the prehistoric migration of anatomically modern humans from Africa in generating differential development patterns across the globe.

This line of research has explored the persistent effect of the prehistoric exodus of *Homo sapiens* from Africa on comparative economic development across societies from the dawn of civilization to the contemporary era. In particular, this research suggests that greater migratory distances from the cradle of mankind in East Africa to the indigenous settlements of the ancestral populations of nations or ethnic groups diminished their levels of interpersonal diversity and, thereby, generated a persistent hump-shaped influence on development outcomes, reflecting a fundamental trade-off between beneficial and detrimental effects of diversity on productivity at the societal level (Ashraf and Galor, 2013a, 2018; Ashraf, Galor, and Klemp, 2018a). Although diversity may reduce interpersonal trust, cooperation, and social cohesiveness, and can thus adversely affect the productivity of society, complementarity across diverse productive traits may stimulate innovations and gains from specialization, and can thereby contribute to a society's economic performance. Therefore, in the presence of diminishing marginal effects of diversity and homogeneity on productivity, the aggregate economic performance of ethnic groups, countries, or regions that are characterized by intermediate levels of diversity would be expected to be higher than that associated with excessively homogenous or heterogeneous societies.

Consistent with each of the fundamental building blocks of this hypothesis, interpersonal diversity, as proxied by diversity in genetic markers, appears to have contributed to ethnic and cultural heterogeneity (Ashraf and Galor, 2013b), diminished interpersonal trust (Ashraf and Galor 2013a), heterogeneity in political preferences, and the emergence of civil conflicts (Arbatlı, Ashraf, Galor, and Klemp, 2019). Moreover, evidence suggests that interpersonal diversity may foster innovative activity (Ashraf and Galor, 2013a; Cook and Fletcher, 2018), occupational heterogeneity, and gains from specialization (Depetris-Chauvin and Özak, 2018). Further, interpersonal diversity may have shaped the nature of both precolonial and contemporary political institutions. In particular, although diversity may have triggered the development of institutions for mitigating the adverse influence of population diversity on social cohesion, the contribution of diversity to economic inequality and class stratification may have ultimately led to the formation and persistence of extractive and autocratic institutions (Galor and Klemp, 2018).

This influential literature has attracted the attention of some scholars from other disciplines, and in light of existing methodological gaps across fields, it has perhaps unsurprisingly generated some significant misconceptions. In particular, Rosenberg and Kang (2015) reexamine only *one* of the range of results established in Ashraf and Galor (2013a), and they suggest that the significant hump-shaped effect of diversity on population density in the year 1500, as established by Ashraf and Galor (2013a) based on the 53 ethnic groups in the Human Genome Diversity Project (HGDP-CEPH), is statistically insignificant in an extended sample based on more than 230 ethnic groups that was more recently assembled by Pemberton et al. (2013). Rosenberg and Kang (2015) conclude that this "challenges the claim for a role of genetic diversity in economic development".

This assertion, however, is a straw-man argument based on elementary statistical errors. The small non-representative sample, consisting of only 39 countries, rather than the 145 countries considered in the original study of Ashraf and Galor (2013a), has no ramification on the validity of the out-of-Africa hypothesis of comparative development. Furthermore, the *global* hump-shaped effect of interpersonal population diversity on population density is in fact *more pronounced* when using the extended sample of Pemberton et al. (2013), and it is present not only in the year 1500 but over the entire pre-colonial period for which population data are available (i.e., the 10,000BCE to 1500CE timeframe).

2 Statistical Errors that Govern Rosenberg and Kang's Findings

Rosenberg and Kang (2015) explore the significance of the hump-shaped association between observed diversity and historical population density in the year 1500 in a small sample of 39 countries, asserting that their analysis suggests that the hypothesized hump-shaped relationship is statistically insignificant. However, their test of the existence of a hump-shaped influence of interpersonal diversity on economic prosperity is fundamentally flawed, suffers from a number of elementary statistical errors, and cannot be used to assess the out-of-Africa hypothesis of comparative development.

2.1 Non-Representative Sample

While valid statistical inferences can be made only in representative samples, Rosenberg and Kang (2015) based their analysis on a small non-representative sample consisting of only 39 countries, a small subset of the 145-country sample considered in the original study of Ashraf and Galor (2013a), based on predicted genetic diversity. This highly limited sample of countries is simply the set of contemporary nations whose territories contain at least one of the ethnic groups in the Pemberton et al. sample, disregarding the distortion that this restriction introduced to the representativeness of the sample.

The arguably non-representative sample of 39 countries precludes statistical inferences and cannot be credibly used to assess the existence of the global relationship between economic development and population diversity. In particular, if a highly significant hump-shaped association is in fact present between diversity and population density in a 145-country sample, one can select 39 observations from this sample that will erroneously suggest that a hump-shaped association does not exists.

In contrast, Ashraf and Galor's (2013a) advances a methodology that permits the projection of population diversity for a globally representative sample of countries, based on the migratory distances of their ancestral populations from East Africa. They established a significant hump-shaped relationship in a globally representative sample of 145 countries, which as discussed in Section 4, remains intact regardless of whether it is based on the HGDP-CEPH data set or on the more recent one by Pemberton et al. (2013).

2.2 Other Elementary Statistical Flaws That Invalidate their Finding

While the choice of the non-representative sample is sufficient to invalidate Rosenberg and Kang's finding, their analysis is afflicted by three additional statistical flaws that would have invalidated their finding even in a representative sample.

2.2.1 Limited 39-Country Sample

The limited 39-country sample used by Rosenberg and Kang diminishes the power of the statistical test and reduces the likelihood of finding a significant statistical association between population diversity and economic outcomes, let alone a non-linear one. Hence, their finding of an insignificant association could merely be an artifact of their limited sample size.

In contrast, in order to overcome sample limitations, Ashraf and Galor (2013a) exploit the pronounced impact of migratory distance from East Africa on observed genetic diversity across ethnic groups in order predict interpersonal diversity for all nations. The use of projected population diversity, associated with a sample of 145 countries, provides augmented statistical power and permit the confirmation of a significant hump-shaped relationship between population diversity and population density in the year 1500. Moreover, since the coefficients of the estimated relationship between migratory distance and observed diversity are virtually identical in the HGDP–CEPH sample versus the extended sample of ethnic groups, the hump-shaped influence of predicted diversity on historical cross-country comparative development remains valid for the extended sample as well.

2.2.2 Measurement Errors Due to the Erroneous Aggregation of Population Diversity

Rosenberg and Kang's measure of genetic diversity at the country level is constructed based on the unweighted (arithmetic) average of genetic diversity of the ethnic groups within the geographical territory of each modern country. This naïve and superficial aggregation, ignoring the proportional representation of each ethnic group within each country as well as their pairwise genetic distances, introduces significant measurement errors to the independent variable.

As is well-known from statistical theory, measurement errors in the independent variable biases the estimated coefficient towards zero. Hence, the assertion of Rosenberg and Kang (2015) that the estimated coefficients on genetic diversity and its square are insignificantly different from zero, could be plausibly governed by the measurement errors that they introduce to the measure of population diversity, due to this major flaw in their aggregation methodology.

2.2.3 Reverse Causality and Omitted Variable Bias

Observed diversity may reflect past socioeconomic outcomes, such as intraregional social conflicts and migrations that were likely driven by historical differences across societies in economic prosperity. Moreover, omitted geographical characteristics can mask the existence of a hump-shaped relationship. Thus, Rosenberg and Kang's (2015) analysis, which is based on observed genetic diversity and does not include a sufficiently large number of geographical controls, is afflicted by

reverse causality and bias from omitted variables that surely obscure the existence of a hump-shaped relationship between diversity and economic prosperity.

Indeed, even in the context of their non-representative sample of countries, a highly significant hump-shaped association emerges between observed diversity and historical population density once a basic set of potential geographical confounders is accounted for.

In contrast, in order to overcome potential concerns about reverse causality associated with the use of *observed* genetic diversity, Ashraf and Galor (2013a) and Ashraf, Galor and Klemp (2018a) exploit the pronounced impact of migratory distance from East Africa on observed genetic diversity across ethnic groups in order to predict interpersonal diversity for all ethnic groups, thereby applying an instrumental variable approach to confirm the hump shaped impact of population diversity on economic outcomes over the course of human history.

3 An Ethnic-Level Analysis

An analysis at the ethnic-group level provides the first-best setting for evaluating the main predictions of Ashraf and Galor's hypothesis to using Pemberton et al.'s (2013) data on observed diversity in the extended sample of ethnic groups.

Ashraf, Galor, and Klemp (2018a) have empirically examined the influence of diversity on productivity at the ethnic-group level, while accounting for the potentially confounding effects arising from observed heterogeneity in various ethnicity-specific geographical, cultural, and institutional factors, as well as unobserved heterogeneity in country-specific characteristics. Their research establishes that observed diversity in a worldwide sample of ethnic groups (Pemberton et al., 2013), as well as predicted diversity (based on migratory distance from East Africa) in an extended sample, confers a significant hump-shaped influence on economic prosperity, suggesting that the variation in interpersonal diversity across ethnic homelands has contributed to variations in their economic development, as captured by their historical population densities since the Neolithic Revolution, as well as their nighttime luminosities in the contemporary era.

The hump-shaped impact of interpersonal population diversity on historical population densities over the period 10,000BCE – 1500CE is established at the ethnic-group level (across the extended sample of ethnic groups in Pemberton et al., 2013). As reported in Ashraf, Galor, and Klemp (2018a), and as depicted partly in Figure 1, the impact is highly significant for each thousand-year interval over this time period.

Moreover, as established by Ashraf, Galor, and Klemp (2018a) and depicted in Figure 2, the hump-shaped impact of interpersonal population diversity is robust to the use of an alternative measure of contemporary economic development, based on luminosity per capita across the extended sample of ethnic groups in Pemberton et al. (2013).

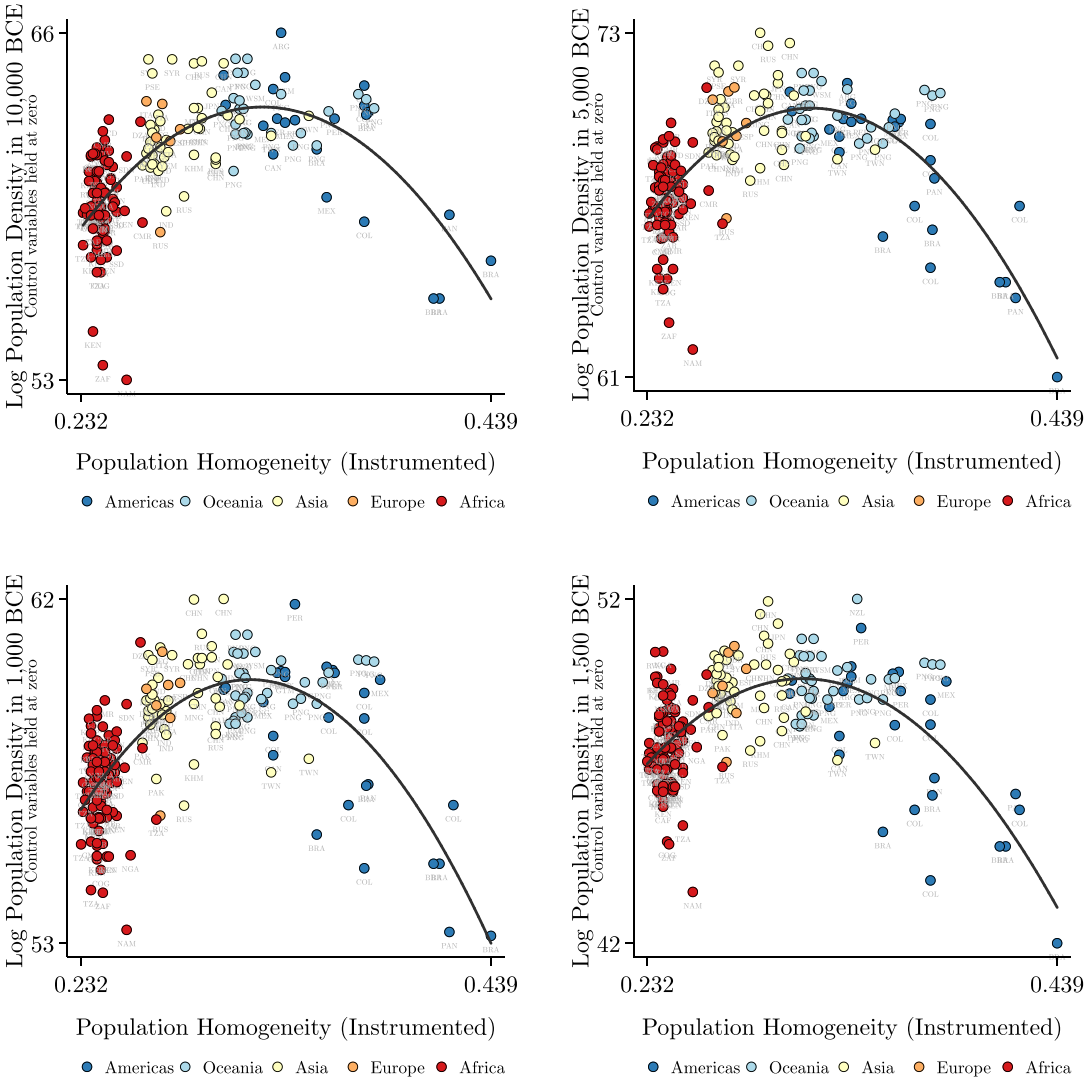

Figure 1. Interpersonal Diversity and Population Density across Ethnic Groups over the Period 10,000BCE – 1500CE

Notes: This figure depicts the hump-shaped influence of interpersonal diversity on population density across ethnic groups in the extended sample of Pemberton et al. (2013). It depicts the impact of observed genetic homogeneity (i.e., 1 minus observed diversity), instrumented by the migratory distance from East Africa, on population density in the years 10,000BCE, 5,000BCE, 1000BCE, and 1500CE, conditional on geographical controls and regional fixed effects. *Source*: Ashraf, Galor, and Klemp (2018a).

4 A Country-Level Analysis

A country-level analysis is the second-best setting for evaluating the robustness of the main predictions of Ashraf and Galor's hypothesis to the use of Pemberton et al.'s (2013) data on observed diversity in the extended sample of ethnic groups.In particular, following Ashraf and Galor's (2013a) methodology, one can generate measures of *predicted* diversity for a globally representative sample of countries, based on the migratory distances of their ancestral population from East Africa. In particular, in order to overcome sample limitations and potential concerns about reverse causality associated with the use of *observed* genetic diversity, Ashraf and Galor (2013a) exploit the pronounced impact of migratory distance from East Africa on observed genetic diversity across ethnic groups in order predict interpersonal diversity for all nations, based on the geographical locations of their ancestral populations in the year 1500, relative to East Africa. Since the coefficients of the estimated relationship between migratory distance and observed diversity are virtually identical in the HGDP–CEPH sample versus the extended sample of ethnic groups, the hump-shaped influence of predicted diversity on historical cross-country comparative development remains unaffected.

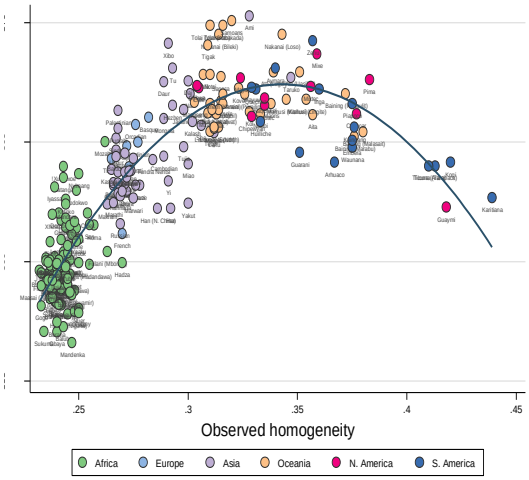

Figure 2. Interpersonal Diversity and Contemporary Comparative Development across Ethnic Groups

Notes: This figure depicts the hump-shaped influence of interpersonal diversity on contemporary economic development across ethnic groups. It depicts the relationship between observed homogeneity (i.e., 1 minus expected heterozygosity) and log average light intensity per capita in the 1992–2013-time horizon, conditional on absolute latitude, soil quality, type of landmass, group size, institutional characteristics (extent of jurisdictional hierarchy and type of class stratification), and regional fixed effects. *Source*: Ashraf, Galor, and Klemp (2018a).

As depicted in Panel A of Figure 3, predicted population diversity has a significant hump-shaped effect on population density in the year 1500 across countries. Moreover, employing an alternative measure of historical development based upon the extent of urbanization, rather than population density in 1500 (Ashraf and Galor, 2011), does not qualitatively alter the hump-shaped influence

of interpersonal diversity on historical development, as depicted in Panel B of Figure 3.[1]

Furthermore, since migratory distance from East Africa has a negative influence on various forms of intragroup phenotypic diversity, *predicted* interpersonal diversity is a valid proxy for diversity in phenotypically and behaviorally expressed traits. In particular, mounting evidence in the fields of physical and cognitive anthropology suggest that an ancient serial founder effect originating in East Africa generated the observed worldwide patterns in various forms of intragroup morphological and cognitive diversity, including interpersonal diversity in skeletal features pertaining to cranial characteristics (Manica et al. 2007; von Cramon-Taubadel and Lycett 2008; Betti et al. 2009), dental attributes (Hanihara 2008), pelvic traits (Betti et al. 2013), and birth canal morphology (Betti and Manica, 2018), as well as intralingual phonemic diversity (Atkinson 2011).

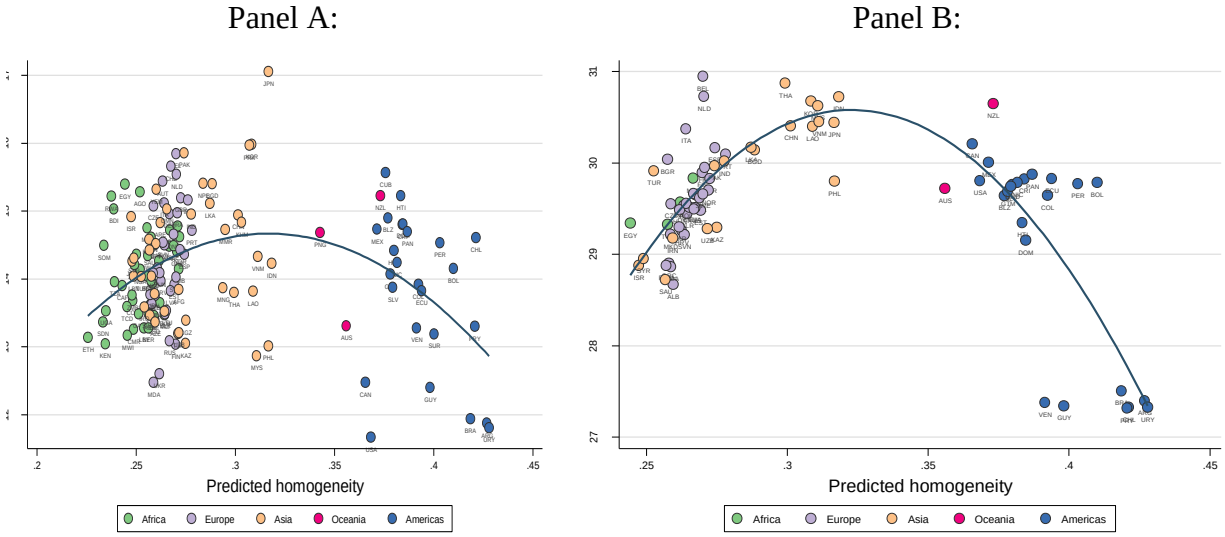

Figure 3. Interpersonal Diversity and Historical Comparative Development across Countries

Notes: This figure depicts the cross-country hump-shaped influence of predicted homogeneity (i.e., 1 minus interpersonal diversity predicted by migratory distance from East Africa) on economic development in the year 1500, as reflected by either log population density (Panel A) or log urbanization rate (Panel B), conditional on the timing of the Neolithic Revolution, land productivity, and continent/regional fixed effects. *Source*: Ashraf and Galor (2013a).

5 Concluding Remarks

Ashraf and Galor (2013a) advanced the hypothesis that migratory distances from the cradle of mankind in East Africa to the indigenous settlements of the ancestral populations of nations or ethnic groups diminished their levels of interpersonal diversity and, thereby, generated a persistent hump-shaped influence on development outcomes, reflecting a fundamental trade-off between beneficial and detrimental effects of diversity on productivity at the societal level. Although di-

[1] It should be noted that the data source for urbanization rates in 1500 is independent of the source for historical population density.

versity may reduce interpersonal trust, cooperation, and social cohesiveness, and can thus adversely affect the productivity of society, complementarity across diverse productive traits may stimulate innovations and gains from specialization, and can thereby contribute to society's economic performance.

The literature on the influence of interpersonal diversity on comparative development across societies has attracted the attention of the scholarly community beyond the discipline of economics and, given methodological divisions, it has perhaps unsurprisingly generated unfounded criticisms. In particular, Rosenberg and Kang (2015) reexamine only *one* of the range of results established in Ashraf and Galor (2013a), suggesting that the hump-shaped effect of diversity on population density in the year 1500 is statistically insignificant in an extended sample of genetic diversity that was released more recently. They conclude that this "challenges the claim for a role of genetic diversity in economic development". However, this assertion is based on elementary statistical errors. In fact, the hump-shaped effect of interpersonal population diversity on population density is even stronger in this extended sample of Pemberton et al. (2013), and it is present not only in the year 1500 but over the entire pre-colonial period for which population data is available (i.e., the 10,000BCE to 1500CE timeframe).

REFERENCES

Arbatlı, Cemal Eren, Quamrul H. Ashraf, Oded Galor, and Marc Klemp. 2019. "Diversity and Conflict." *Econometrica* (forthcoming).

Ashraf, Quamrul, and Oded Galor. 2011. "Dynamics and Stagnation in the Malthusian Epoch." *American Economic Review*, 101(5): 2003–2041.

Ashraf, Quamrul, and Oded Galor. 2013a. "The 'Out of Africa' Hypothesis, Human Genetic Diversity, and Comparative Economic Development." *American Economic Review*, 103(1): 1–46.

Ashraf, Quamrul, and Oded Galor. 2013b. "Genetic Diversity and the Origins of Cultural Fragmentation," *American Economic Review*, 103(3): 528–533.

Ashraf, Quamrul H., and Oded Galor. 2018. "The Macrogenoeconomics of Comparative Development," *Journal of Economic Literature*, 56(3): 1119–1155.

Ashraf, Quamrul H., Oded Galor, and Marc Klemp. 2018a. "Population Diversity and Differential Paths of Long-Run Development since the Neolithic Revolution." Unpublished.

Ashraf, Quamrul H., Oded Galor, and Marc Klemp. 2018b. "The Out of Africa Hypothesis of Comparative Economic Development: Common Misperceptions." Working Paper No. 2018-13, Brown University.

Atkinson, Quentin D. 2011. "Phonemic Diversity Supports a Serial Founder Effect Model of Language Expansion from Africa." *Science*, 332(6027): 346–349.

Betti, Lia, François Balloux, William Amos, Tsunehiko Hanihara, and Andrea Manica. 2009. "Distance from Africa, Not Climate, Explains Within-Population Phenotypic Diversity in Humans." *Proceedings of the Royal Society B: Biological Sciences*, 276(1658): 809–814.

Betti, Lia, Noreen von Cramon-Taubadel, Andrea Manica, and Stephen J. Lycett. 2013. "Global Geometric Morphometric Analyses of the Human Pelvis Reveal Substantial Neutral Population History Effects, Even across Sexes." *PLoS ONE*, 8(2): e55909.

Betti, Lia, and Andrea Manica. 2018. "Human Variation in the Shape of the Birth Canal is Significant and Geographically Structured." *Proceedings of the Royal Society B: Biological Sciences*, 285(1889), DOI: 10.1098/rspb.2018.1807.

Cook, C. Justin, and Jason M. Fletcher. 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.

Depetris-Chauvin, Emilio, and Ömer Özak. 2018. "The Origins of the Division of Labor in Pre-Modern Times." IZA Discussion Paper 11554, Institute for the Study of Labor (IZA).

Galor, Oded, and Marc Klemp. 2018. "The Roots of Autocracy." National Bureau of Economic Research Working Paper 23301.

Hanihara, Tsunehiko. 2008. "Morphological Variation of Major Human Populations Based on Nonmetric Dental Traits." *American Journal of Physical Anthropology*, 136(2): 169–182.

Manica, Andrea, William Amos, François Balloux, and Tsunehiko Hanihara. 2007. "The Effect of Ancient Population Bottlenecks on Human Phenotypic Variation." *Nature*, 448(7151): 346–348.

Pemberton, Trevor J., Michael DeGiorgio, and Noah A. Rosenberg. 2013. "Population Structure in a Comprehensive Genomic Data Set on Human Microsatellite Variation." *G3: Genes, Genomes, Genetics*, 3(5): 891–907.

Prugnolle, Franck, Andrea Manica, and François Balloux. 2005. "Geography Predicts Neutral Genetic Diversity of Human Populations." *Current Biology*, 15(5): R159–R160.

Rosenberg, Noah A., and Jonathan T. L. Kang. 2015. "Genetic Diversity and Societally Important Disparities." *Genetics*, 201(1): 1–12.