I Z A Institute
of Labor Economics
Initiated by Deutsche Post Foundation

DISCUSSION PAPER SERIES

IZA DP No. 16068

The Effect of the Out of Africa Migration on Cultural Diversity

Daniel Crisóstomo Wainstock
Oded Galor
Marc Klemp

APRIL 2023

I Z A Institute
of Labor Economics

Initiated by Deutsche Post Foundation

DISCUSSION PAPER SERIES

IZA DP No. 16068

The Effect of the Out of Africa Migration on Cultural Diversity

Daniel Crisóstomo Wainstock
Brown University

Marc Klemp
University of Copenhagen

Oded Galor
Brown University and IZA

APRIL 2023

ISSN: 2365-9793

IZA – Institute of Labor Economics

Schaumburg-Lippe-Straße 5–9
53113 Bonn, Germany

Phone: +49-228-3894-0
Email: publications@iza.org

www.iza.org

ABSTRACT

The Effect of the Out of Africa Migration on Cultural Diversity

Evidence suggests that the Out of Africa Migration has impacted the degree of intra-population genetic and phenotypic diversity across the globe. This paper provides the first evidence that this migration has shaped cultural diversity. Leveraging a folklore catalogue of 958 oral traditions across the world, we show that ethnic groups further away from East Africa along the migratory routes have lower folkloric diversity. This pattern is consistent with the compression of genetic, phenotypic, and phonemic traits along the Out of Africa migration routes, setting conditions for the emergence and proliferation of differential cultural diversity and economic development across the world.

JEL Classification: O10, Z10
Keywords: diversity, culture, Out of Africa migration, folklore

Corresponding author:
Oded Galor
Department of Economics and Population Studies and Training Center
Brown University
64 Waterman St.
Providence, RI 02912
USA
E-mail: oded galor@brown.edu

The *Out of Africa Migration* is one of the most important chapters of human history as it has shaped the initial conditions for the evolution of human settlements across the world. According to the prevailing hypothesis, this migration was largely characterized by a stepwise expansion, where in each step a subgroup of individuals left their ancestral settlement to establish a new colony farther away, carrying only a subset of the traits of their ancestral settlement. The resulting *Serial Founder Effect* has generated a negative effect of migratory distance from East Africa on genetic, phenotypic, and arguably phonemic diversity[1].

It remains uncertain, however, how deep was the imprint of this process on cultural diversity. Resolving this mystery is critical for the understanding of the extent of the impact of this migratory process on intra-population diversity and its potential persistent effect on the process of economic development over the entire course of human history. In particular, establishing an impact of this migratory process on cultural diversity would contribute to the understanding of the cultural channel for the consequences of the *Out of Africa Migration* on societal outcomes such as interpersonal trust, political preferences, conflict and comparative economic development (Ashraf, Galor, and Klemp 2021).

In view of the impact of the *Serial Founder Effect* on the compression of genetic, phenotypic, and possibly phonemic traits along the Out of Africa migration routes, we hypothesize that this process set the stage for differential emergence and proliferation of cultural diversity across the world, resulting in a negative relationship between cultural diversity and migratory distance from East Africa.

In this paper, we study one important dimension of cultural diversity – *folkloric diversity* – as reflected by the variation in the total number of motifs across oral traditions (i.e., ethnic groups or clusters of ethnic groups along linguistic lines). Leveraging worldwide data on 958 oral traditions (Fig. 1) and 2,564 motifs from the Berezkin's Folklore and Mythology Catalogue (Berezkin 2015; Michalopoulos and Xue 2021), we investigate how the degree of folkloric diversity (Fig. 3) was shaped by this migratory process.

[1] See Atkinson (2011), Betti (2009), Betti et al. (2013), Betti and Manica (2018), Hanihara (2008), Manica et al. (2007), Ramachandran et al. (2005), von Cramon-Taubadel and Lycett (2008).

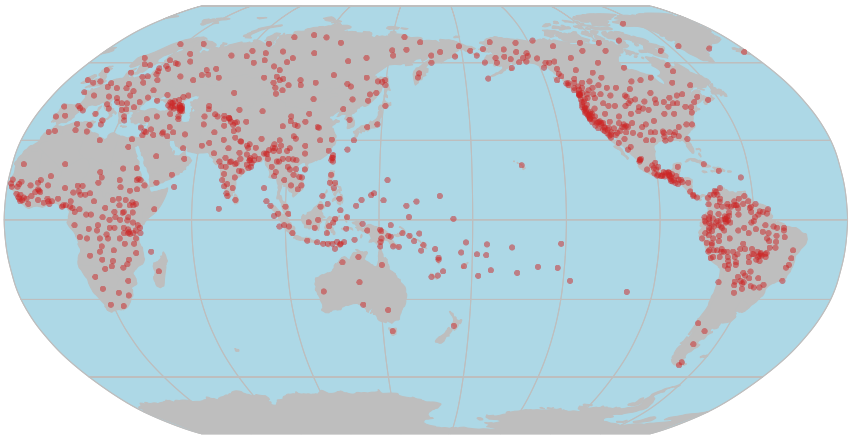


Fig. 1. Locations of Berezkin's Oral Traditions.
Notes: This figure depicts the locations of the 958 oral traditions in Berezkin's Folklore and Mythology Catalogue.

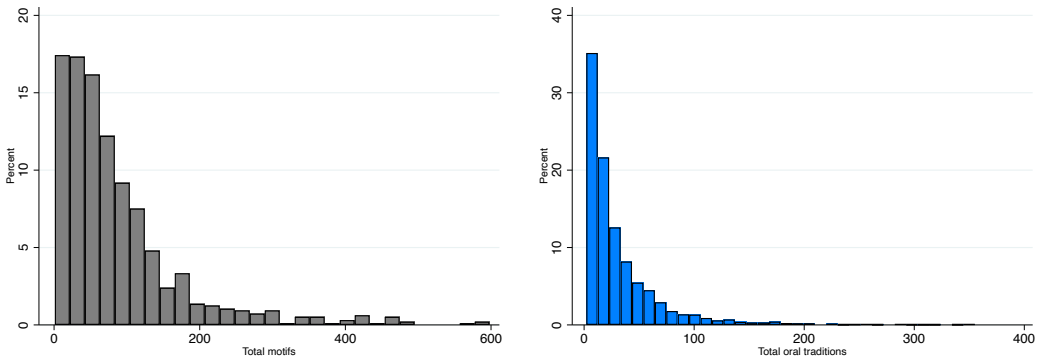
I. Data

Berezkin's Folklore and Mythology Catalogue is the lifetime work of anthropologist and folklorist Yuri Berezkin. He has dedicated his entire career to compiling and classifiying oral traditions for ethnic groups across all continents. His initiative is unique in the sense that it is a large-scale project of documenting folklore which contrasts with small-scale endeavors which focused on specific regions of the world. This feature is essential to our analysis as it gives us consistency in the classification of motifs and breadth of data which allows us to make a general claim about the imprint of the *Out of Africa Migration* on the stories told by indigenous societies across the entire world.

Berezkin has recorded 2,564 motifs across 958 oral traditions after consulting 6,239 books and journal articles from 4,041 authors. The motifs are the fundamental building blocks of oral traditions. They capture significant events, experiences and images, and are shared across two or more oral traditions. An oral tradition is either an individual ethnic group or sometimes a cluster of ethnic groups along linguistic lines where there is not enoguh information on individual groups. The catalogue records whether a motif is present or absent in an oral tradition, and provides no information on the relevance or popularity of that story for a given oral

tradition.

Fig. 1 depicts the locations of these oral traditions. The coverage is specially rich for the Americas both in terms of oral traditions and motifs. This is so because Berezkin began his work in that continent. This feature of the catalogue is very important as it means that, by construction, there exists an uneven coverage of oral traditions and motifs across continents. This motivates why in our statistical model we exploit variation within continents. Fig. 2 provides a histogram of total motifs across oral traditions. The distribution is severely skewed. Half of the groups have fewer than 62 motifs while the groups in top 1% of the distribution have more than 450 motifs. This motivates our choice for a logarithmic transformation of the number of motifs as our dependent variable in our statistical model.



(a) number of motifs per oral traditions **(b)** number of oral traditions per motif

Fig. 2. Berezkin's Folklore and Mythology Catalogue.
Notes: Panel (a) provides a histogram of the number of motifs by oral tradition. Panel (b) provides a histogram of the number of oral traditions which share a given motif.

II. Empirical Strategy

We estimate the following linear regression model:

$$M_t = \alpha + \beta * D_t + \gamma * X_t + \delta * Z_t + \zeta_c + \varepsilon_t$$

where t refers to an oral tradition, M_t is the total number of motifs, D_t is the migratory distance from East Africa, X_t is a vector of baseline controls which include the log number of publications consulted by Berezkin to account for the relation-

ship between the number of motifs and the degree to which an oral tradition was studied by folklorists, and the log year of first publication to account for differences over time in the collection of folklore. Z_t is a vector of additional controls which include absolute latitude, mean caloric suitability, ecological diversity as proxied by the standard deviation in caloric suitability and elevation, an island dummy, and log distance to coast. δ_c are continent fixed-effects. The coefficient of interest is β. According to our hypothesis, we expect $\beta < 0$.

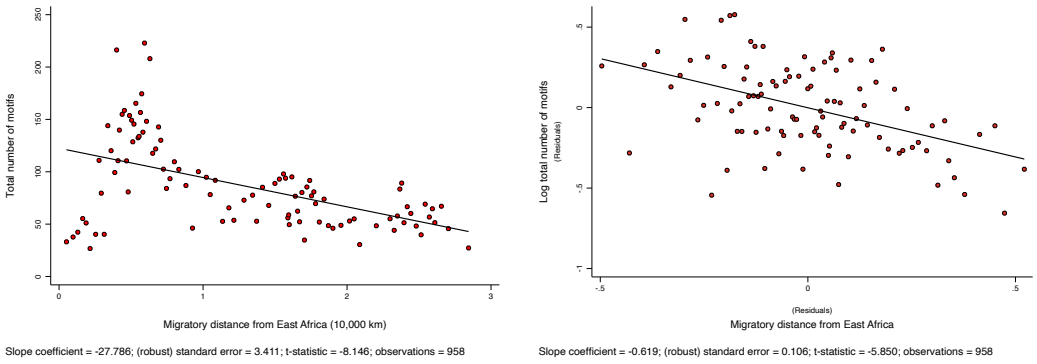
We also employ alternative estimation methods designed to account for count data and dispersion in the dependent variable. In particular, we show the robustness of our results to the use of either log-linear regression or Negative Binomial regressions.

We further test the robutness of our results to the inclusion of potential ethnographic confounders. Namely, the mean size of local communities, the degree of political centralization, the degree of cultural complexity as measured by the presence of high gods, and the engagement of groups with trade. We use the variables from the Ethnographic Atlas. These are respectively variables *v31*, *v33*, and *v34*. To assign a level of either of them to an oral tradition, we rely on the match between Berezkin's Folklore and Mythology Catalogue and the Ethnographic Atlas done by Michalopoulos and Xue (2021). We simply take the average value of either ethnographic variable at the oral tradition level. The Ethnographic Atlas has no information on trade. However, the work of Michalopoulos and Xue (2021) allows us to get a proxy for that: the intensity at which oral traditions mention trade related terms, which they show to strongly correlate with distance to ancient trade routes. It should be noted that these variables are potentially endogenous given the potential the effect of folkloric diversity on them.

III. Results

The raw data indeed suggests that oral traditions further along the migratory routes out of Africa have fewer motifs in their set of narratives (Fig. 3).

Table 1 presents the effect of migratory distance from East Africa on folkloric diversity, measured by the number of motifs, using three estimation methods: lin-



(a) Raw **(b)** Residuals on Residuals

Fig. 3. Total number of motifs and the migratory distance from East Africa across oral traditions.

Notes: This figures presents scatterplots with observations grouped in 100 equally-sized bins. Panel (a) shows the raw association between total number of motifs and migratory distance from East Africa with a fitted line estimated using a linear model. The point estimate is statistically significant at the 1% level. Following our statistical model, we residualize migratory distance from East Africa and the log total of motifs on continent fixed-effects, log year of first publication, and log number of publications consulted by Berezkin. Panel (b) shows the same association between residualized log total of motifs and residualized migratory distance from East Africa. The point estimate is statistically significant at the 1% level.

ear regression (Panel A), log-linear regression (Panel B), and Negative Binomial regression (Panel C). The estimates in Column (1) suggest that indeed the migration out of Africa is a highly significant negative determinant of folkloric diversity on a global scale. Nevertheless, if in fact the migration out of Africa had an effect on folkloric diversity, this should have also operated on a local scale. The confirmation of the indirect effect of the *Serial Founder Effect* on folkloric diversity requires, therefore, that this finding holds within continents (i.e., based on variation in oral traditions within each continent). Accounting for these continental fixed effects, the estimates in Column (2) establish that indeed the migration out of Africa affected folkloric diversity within as well as across continents.

Moreover, in view of the potential differences in continent-specific factors that may affect the observed folkloric diversity, some of which may be unobserved, the continent-specific fixed effects (e.g., differential sampling of oral traditions and documentation of these narratives, colonialism, slavery, disease environment, and

climatic conditions) which are included in Column (2) account for these potential confounders. The results suggest that the average observed impact of the migration out of Africa is larger within continents compared to the global impact. To further account for differences in the sampling of oral traditions and documentation of narratives within continents, Column (3) accounts for the year of first publication and number of publications consulted to construct the catalogue.

We further establish that this qualitative pattern is unaffected by the geographical features of these oral traditions. In light of the fact that distance from the equator is correlated with declining biodiversity, Column (4) accounts for this potentially confounding factor, adding absolute latitude as a control. Accounting for the range of additional potentially confounding geographic characteristics has no qualitative impact on the point estimate and its statistical significance. In particular, in Column (5), we add ecological diversity as a control to further account for the effect of biodiversity. This feature has been shown to have influenced ethnolinguistic fragmentation (Michalopoulos 2012) and it has been suggested as a potential determinant for the emergence of trade (Bates 1983). Furthermore, since suitability of land for agriculture is associated with the scale and density of communities, Column (6) accounts for this potentially confounding dimension by including the caloric suitability of land as a control (Galor and Ozak 2016). Moreover, in view of the potential influence of the degree of isolation of an ethnic group on the diffusion of stories between groups, Column (7) accounts for this dimension by including a dummy for whether an oral tradition is located in an island. Furthermore, in Column (8), we control for distance to the coast. The rationale is that this control has been an important determinant of the degree of contact among groups, with groups closer to the coast being arguably more in touch with distant cultures. Moreover, as established in Columns (9) and (10), the effect operates even within the modern national boundaries of these ethnic groups.

The estimated effect is sizable. In particular, Column 8 in Panel A suggests that a 10,000 km increase in migratory distance from East Africa to a given ethnic group (which corresponds to the migratory distance between the Inupiat at the Bering Strait and the Kuna in Colombia) is associated with 50 fewer motifs in comparison with the median number of 62 motifs per ethnic group.

It should be noted that given the uncertainty about the origin of humans within the African continent and the massive migration within the continent (e.g., the Bantu migration), migratory distance from East Africa has been shown to be a poor predictors of the *Serial Founder Effect* within the African continent (Ramachandran et al. 2005).

Table 1. Baseline cross-oral tradition regressions on the relationship between log total number of motifs and the migratory distance from East Africa.

	Total Number of Motifs									
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
					A. Linear					
Migratory distance from East Africa										
(in 10,000 kms)	−27.8***	−71.2***	−63.9***	−52.4***	−54.1***	−55.9***	−50.3***	−49.9***	−88.2***	−67.9***
	(3.41)	(11.5)	(9.30)	(9.45)	(9.49)	(9.34)	(9.52)	(9.58)	(22.3)	(17.8)
Adjusted R^2	0.061	0.32	0.54	0.56	0.56	0.57	0.57	0.57	0.46	0.63
					B. Log-linear					
Migratory distance from East Africa										
(in 10,000 kms)	−0.19***	−0.73***	−0.62***	−0.50***	−0.51***	−0.52***	−0.49***	−0.49***	−0.71***	−0.32*
	(0.043)	(0.15)	(0.11)	(0.10)	(0.11)	(0.11)	(0.11)	(0.11)	(0.24)	(0.18)
Adjusted R^2	0.021	0.19	0.58	0.59	0.59	0.59	0.59	0.59	0.26	0.60
					C. Negative Binomial					
Migratory distance from East Africa										
(in 10,000 kms)	−0.35***	−0.81***	−0.64***	−0.54***	−0.56***	−0.57***	−0.52***	−0.52***	−0.78***	−0.46***
	(0.037)	(0.13)	(0.098)	(0.097)	(0.098)	(0.098)	(0.10)	(0.100)	(0.18)	(0.14)
Effect Size	−25.6	−52.3	−42.6	−37.3	−38.4	−38.9	−36.1	−36.0	−50.2	−32.4
Continent FE	no	yes	yes	yes	yes	yes	yes	yes	no	no
Country FE	no	no	no	no	no	no	no	no	yes	yes
Data source controls	no	no	yes	yes	yes	yes	yes	yes	no	yes
Absolute latitude	no	no	no	yes	yes	yes	yes	yes	no	yes
Ecological diversity	no	no	no	no	yes	yes	yes	yes	no	yes
Caloric suitability of land	no	no	no	no	no	yes	yes	yes	no	yes
Island	no	no	no	no	no	no	yes	yes	no	yes
Distance to coast	no	no	no	no	no	no	no	yes	no	yes
Observations	958	958	958	958	957	957	957	957	958	957

Notes: The table establishes a significant negative effect of migratory distance from East Africa on folkloric diversity. Panel A presents the results from linear regressions. Panel B presents the results from log-linear regressions, in which the outcome variable is the log total number of motifs. Panel C presents the results from Negative Binomial regressions. Effect Size refers to the difference in the estimated marginal mean between the mean distance in the sample and a location 10,000 kms further away from East Africa, with the control variables held at their mean value. Heteroskedasticity robust standard errors are reported in parantheses. *** Significant at the 1 percent level. ** Significant at the 5 percent level. * Significant at the 10 percent level.

Table 2 shows that accounting for a range of potentially confounding cultural and institutional characteristics has no qualitative impact on the point estimate and its statistical significance. In particular, in light of the fact that state experience has arguably shaped cultural diversity, in Columns (1) to (3), we restrict the analysis to the set of oral traditions for which we observe the level of political centralization, and account for this potentially confounding dimension as a control. Furthermore, since the presence of knowledgeable and punitive gods engaged in human affairs has been shown to be a fundamental characteristic to understand the evolution of human sociality (Purzycki et al. 2016), in Columns (4) to (6), we restrict the analysis to the set of oral traditions for which we observe the presence of high gods, and account for this potentially confounding factor as a control. Moreover, as economic development is plausibly a driver of the production of stories, in Columns (7) to (9), we restrict the analysis to the set of oral traditions for which we observe the mean size of local communities, and account for this potentially confounding feature as a control. Furthermore, in view of the potential influence of the degree of exposure to markets of an ethnic group on the diffusion of stories between groups, in Columns (10) to (12), we account for this dimension by including the share of motifs on trade as a control.

Table 2. Robustness to Ethnographic Characteristics.

	Log Total Number of Motifs											
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Migratory distance from East Africa												
(in 10,000 kms)	−0.76***		−0.76***	−0.76***		−0.77***	−0.67***		−0.68***	−0.62***		−0.48***
	(0.13)		(0.13)	(0.14)		(0.14)	(0.16)		(0.16)	(0.11)		(0.11)
Political centralization		0.063*	0.064*									
		(0.036)	(0.034)									
High gods					0.040	0.047						
					(0.039)	(0.036)						
Mean size of local communities								0.0099	0.024			
								(0.042)	(0.040)			
Share of motifs on trade											0.20***	0.19***
											(0.034)	(0.035)
Observations	579	579	579	460	460	460	397	397	397	958	958	958
Adjusted R^2	0.58	0.56	0.58	0.59	0.56	0.59	0.60	0.59	0.60	0.58	0.59	0.60

Notes: This table establishes the robustness of the results in Table 1 to the inclusion of control variables for political centralization, belief in high gods, size of local communities, and the engagement with trade. In Columns (1) to (3), we restrict the analysis to the set of oral traditions for which we observe the level of political centralization. In Columns (4) to (6), we restrict the analysis to the set of oral traditions for which we observe the type of high gods. In Columns (7) to (9), we restrict the analysis to the set of oral traditions for which we observe the mean size of local communities. All specifications include continent fixed-effects as well as data source controls (i.e., log year of first publication and log number of publications consulted by Berezkin). Heteroskedasticity robust standard errors are reported in parantheses. *** Significant at the 1 percent level. ** Significant at the 5 percent level. * Significant at the 10 percent level.

IV. Concluding Remarks

This paper provides the first evidence about the impact of the *Out of Africa Migration* on cultural diversity. In particular, it shows that ethnic groups further from East Africa along the migratory routes have fewer motifs in their set of narratives. This result is consistent with the compression of genetic, phenotypic, and possibly phonemic traits along the out of Africa migration routes, setting the stage for the differential emergence and proliferation of cultural diversity across the world. This finding underscores the importance of the exploration of the long-lasting effect of the migration of humans out of Africa on cultural and institutional characteristics, as well as comparative development across countries and ethnic groups, highlighting a unique source of unexplored variations in cultural diversity across regions (Alesina and La Ferrara 2005; Alesina et al. 2003; Easterly and Levine 1997).

References

[1] Alesina, A., A. Devleeschauwer, W. Easterly, S. Kurlat, and R. Wacziarg (2003): "Fractionalization", Journal of Economic Growth, 8, 155-194.

[2] Alesina, A., and E. La Ferrara (2005): "Ethnic diversity and economic performance", Journal of Economic Literature, 43(3), 762-800.

[3] Ashraf, Q., O. Galor, M. Klemp (2021): "The ancient origins of the wealth of nations", The Handbook of Historical Economics, 675–717.

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

[5] Bates, R. (1983): "Essays on the Political Economy of Rural Africa", University of California Press.

[6] Berezkin, Y. (2015): "Folklore and Mythology Catalogue: Its Lay-out and Potential for Research", The Retrospective Methods Network Newsletter, 10, 58–70.

[7] Betti, L., F. Balloux, W. Amos, T. Hanihara, and A. 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.

[8] Betti, L., N. von Cramon-Taubadel, A. Manica, and S. 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.

[9] Betti, L., and A. 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), 20181807.

[10] Easterly, W., and R. Levine (1997): "Africa's growth tragedy: policies and ethnic divisions", Quarterly Journal of Economics, 112(4), 1203-1250.

[11] Galor, O., and O. Ozak (2016): "The Agricultural Origins of Time Preference", American Economic Review, 2016, 106(10): 3064–3103.

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

[13] Manica, A., W. Amos, F. Balloux, and T. Hanihara (2007): "The Effect of Ancient Population Bottlenecks on Human Phenotypic Variation", Nature, 448(7151), 346–348.

[14] Michalopoulos, S. (2012): "The Origins of Ethnolinguistic Diversity", American Economic Review, 102(4), 1508–1539.

[15] Michalopoulos, S., and M. M. Xue (2021): "Folklore", Quarterly Journal of Economics, 136(4), 1993–2046.

[16] Purzycki, B. et al. (2016): "Moralistic gods, supernatural punishment and the expansion of human sociality", Nature, 530, 327–330.

[17] Ramachandran, S., et al. (2005): "Support from the Relationship of Genetic and Geographic Distance in Human Populations for a Serial Founder Effect Originating in Africa." Proceedings of the National Academy of Sciences, 102(44), 15942–15947.

[18] von Cramon-Taubadel, N., and S. J. Lycett (2008): "Brief Communication: Human Cranial Variation Fits Iterative Founder Effect Model with African Origin," American Journal of Physical Anthropology, 136(1), 108–113.