INSTITUTE for REPLICATION

No. 63
DISCUSSION PAPER SERIES

Response to Gonzalez and Özak's (2023) Replication Report

Andrew Dickens

This paper responds to:

Gonzalez, Javier, and Ömer Özak. 2023. Replication of Dickens (2022) "Understanding Ethnolinguistic Differences: The Roles of Geography and Trade". I4R Discussion Paper Series No. 62. Institute for Replication.

September 2023

INSTITUTE for REPLICATION

I4R DISCUSSION PAPER SERIES

I4R DP No. 63

Response to Gonzalez and Özak's (2023) Replication Report

Andrew Dickens[1]

[1]Brock University, St. Catharines/Canada

SEPTEMBER 2023

E-Mail: joerg.peters@rwi-essen.de
RWI – Leibniz Institute for Economic Research
Hohenzollernstraße 1-3
45128 Essen/Germany
www.i4replication.org

ISSN: 2752-1931

Response to Gonzalez and Özak's (2023) Replication Report[†]

Andrew Dickens[‡]

27 July 2023

Abstract

Gonzalez and Özak (2023) provide a direct and successful replication of Dickens (2022). Using a reconstructed version of the main independent variables from the same original sources, in addition to an updated version of the source data, the replicators confirm the main finding of the original study. In addition to the replication, Gonzalez and Özak (2023) develop an alternative measure of potential gains from inter-ethnic trade. They use this new measure in an interesting extension that delves deeper into the the specifics of the inter-ethnic trade mechanism proposed and tested by Dickens (2022). In this response, I clarify two minor points about how the original data set was constructed, and contrast the potential shortcomings of the original and alternative measures of inter-ethnic gains from trade.

[†]The report documents a replication of my paper "Understanding Ethnolinguistic Differences: The Roles of Geography and Trade," published in *The Economic Journal*, Volume 132, Issue 643, April 2022, Pages 953–980.

[‡]Brock University, Department of Economics, St. Catharines, ON. E-mail: adickens@brocku.ca

I would like to thank Gonzalez and Özak (2023) for the thoughtful and detailed replication report on my paper "Understanding Ethnolinguistic Differences: The Roles of Geography and Trade," published in *The Economic Journal* (henceforth Dickens (2022)).

In the aforementioned paper, I study the impact of inter-ethnic trade on linguistic differences. To do this, I use Lewis's (2009) mapping of ethnolinguistic groups from across the world to extract the border segment connecting each neighboring group pair, and construct a buffer zone around each border segment as my unit of observation. As a proxy for the the potential gains from trade, I measure the standard deviation in agricultural productivity within each buffer zone. I find that ethnic groups separated across geographic regions (i.e., buffer zones) with high variation in land productivity are more similar in language than groups separated across more homogeneous regions. As evidence of the proposed mechanism—inter-ethnic trade—I show that ethnolinguistic homelands characterized by high variation in land productivity relied more on inter-ethnic trade for food and subsistence than low variation homelands in pre-modern times.

Replication Clarifications

The main takeaway from the report is that Dickens (2022) is replicable. Following the same methodology, the authors conclude: "we have successfully replicated his central result establishing the negative effect of agricultural variation on linguistic distances. We have shown the results are robust to changes in the construction method and the underlying linguistic maps finding that the empirical analyses remained qualitatively unchanged, although the estimated effects tended to be larger." (Gonzalez and Özak, 2023, p. 13) In other words, the main result of Dickens (2022) is robust to updated data, and in most instances the estimated effects are more precisely estimated and larger in magnitude. Despite the fact that the authors were able to successfully replicate the study, I would like to address two lingering issues noted in the replication report.

First, the authors point out that, in the original paper, I do not clarify the geographical projection used in the analysis. Throughout the entire analysis, I use the Goode homolosine projected coordinate system—an equal-area pseudocylindrical projection commonly used for world maps. As the report authors rightfully point out, using the WGS84 geographic coordinate system for area calculations is not the correct approach and would bias the results. However, based on the findings of the replication report, it seems that the main result of Dickens (2022) is robust to this alternative coordinate system.

Second, the authors point out that I do not specify which Caloric Suitability Index (CSI) is used for the pre-1500 period. The CSI data come from Galor and Özak (2015, 2016), and for the pre-1500 period, these raster data are available for Africa with and without Asian crop varieties included in the CSI calculations. In the original paper, I use the pre-1500 measure that *includes*

Asian crops in Africa. However, once again, the replication report documents that the main result of Dickens (2022) is robust to using any version of the CSI data.

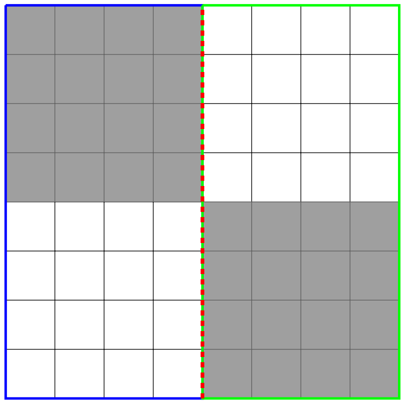
Response to the Mechanism Robustness Analysis

In Section 4 of the replication report, Gonzalez and Özak (2023) provide an interesting and thoughtful extension to the analysis in Dickens (2022). As an alternative measure of potential gains from trade, Gonzalez and Özak (2023) calculate the absolute difference in average productivity on each side of the border, within each buffer zone. This is in contrast to the approach that I take, where I measure the standard deviation of productivity within the entire buffer zone. The authors argue that my measure of inter-ethnic trade potential "fails to capture cross-ethnic economic specialization," and thus is argued to be a measure of *generalized* trade because it cannot "directly provide support for [the] *inter-ethnic* aspect of trade." (Gonzalez and Özak, 2023, p. 10)

To make their point, the authors show that their alternative measure still significantly predicts linguistic distance on its own, but when both measures are entered jointly in a horse race regression, the alternative measure loses significance while the original measure is largely unchanged and remains statistically significant. If the alternative measure is indeed a better measure of inter-ethnic trade, then the evidence suggests *generalized* trade is the underlying mechanism connecting land productivity variations to linguistic differences, not *inter-ethnic* trade.

Finding a perfect proxy for cross-ethnic economic specialization is difficult given available data sources, and although I acknowledge the potential shortcomings of my own measure, I believe the alternative measure similarly suffers from its own shortcomings that make drawing any clear conclusions from the report about generalized vs. inter-ethnic trade difficult. For example, both approaches rely on the accuracy of contemporary linguistic border locations. As the report authors rightfully point out, the movement of a border's location between the historical and contemporary period would introduce measurement error into the analysis (Depetris-Chauvin and Özak, 2020). This potential measurement error resulting from "fuzzy" borders is a major reason why I rely on buffer-level variation throughout my analysis—knowing the exact location of a historical border is not essential for my measurement strategy, assuming that contemporary border movements are not so large that the historical border falls outside of the contemporary buffer zone (my unit of observation). In other words, my approach should arguably allow for minor changes in border location, while still capturing the potential gains from trade in the broader geographic region where two ethnic groups would presumably meet and trade. To the contrary, the alternative measure takes the placement of contemporary borders as given, relying of differences across the exact (and potentially incorrect) border location. To be clear, the authors are upfront about this potential shortcoming in their report. However,

Figure 1: Productivity and Inter-Ethnic Trade Potential in a Hypothetical Buffer Zone



Similar cross-border productivity yet high potential for cross-ethnic economic specialization.

it is noteworthy that the measurement error introduced by fuzzy borders will differ across the two measures, and the measurement error is potentially greater when comparing average productivity differences across the exact location of a border. This different form of measurement error is one potential explanation for why the alternative measure loses significance in a horse race against the original measure.

It is similarly difficult to separate these two mechanisms because both the original and alternative measures share a correlation coefficient of 0.65 or higher. This high level of correlation is not entirely surprising, since a history of generalized trade could build the foundation of an inter-ethnic trading network, and vice versa. As the report authors suggest, "it may very well be the case that linguistic distances decrease because people from *different* ethnic groups interact when trading *similar* goods, i.e., these interactions are not driven by ethnic-level specialization," but a history of such interactions would also incentivize and facilitate further interactions driven by ethnic-level specialization (Gonzalez and Özak, 2023, p. 10). In this context, the inability to observe this dynamic process of social interactions and linguistic change from available data sources introduces a challenge for interpretation—it is likely that both measures, to some extent, capture aspects of generalized and inter-ethnic trade interactions, as the large correlation coefficients suggest. The difficult task of separating these two mechanisms is an interesting avenue for future research.

How do the Original and Alternative Measure Differ?

Of course a correlation coefficient of 0.65 suggests there are some buffer zones where the two measures of potential gains from trade are uncorrelated, or possibly even negatively correlated, which is an interesting opportunity to explore when and where these two related measures differ. The aforementioned horse race regression estimates are identified from buffer zones where

Table 1: Correlation Coefficients of the Original and Alternative Measure at Various Cutoffs

	Ethnologue v16		Ethnologue v17	
Abs Diff in Decile Rank	Correlation Coefficient	% of Sample	Correlation Coefficient	% of Sample
0	0.85	24.28	0.85	24.56
1	0.73	31.50	0.72	31.02
2	0.42	21.45	0.42	20.99
3	0.12	11.45	0.11	11.74
4	-0.12	4.92	-0.12	4.87
5	-0.07	2.48	-0.15	2.54
6	-0.34	1.79	-0.34	2.00
7	-0.65	1.07	-0.58	1.14
8	-0.80	0.79	-0.78	0.87
9	-0.81	0.29	-0.78	0.28

This table documents various subsamples according to the absolute difference in decile rank between pre-1500 land productivity variation (i.e., the original measure) and pre-1500 cross border change in land productivity (i.e., the alternative measure). Correlation coefficients represent the correlation between the original and alternative measure for each subsample.

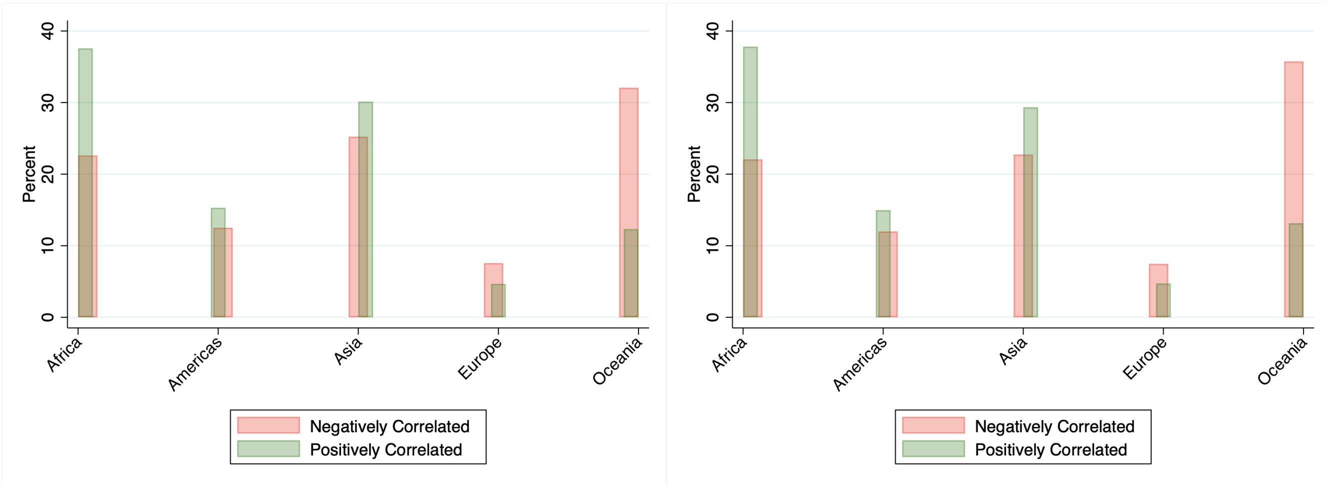
these two related measures differ, so understanding why they differ is essential to learning more about the mechanism. As an example, one could imagine a buffer zone where the original measure captures large variations in productivity, yet the alternative measure captures very little difference in average productivity across the border. Figure 1 depicts this hypothetical buffer zone.[1] In this example, average productivity on each side of the border is identical, yet across the entire buffer zone the productivity level varies considerably. Here, Gonzalez and Özak's (2023) alternative measure suggests an absence of inter-ethnic trade due to the absence of average productivity differences across the border. Yet the distribution of productive land is such that cross-ethnic specialization, and thus inter-ethnic trade, may still arise and, in this context at least, is only identifiable using the original measure.[2]

To delve deeper into this example, I use the reproduced data set from Gonzalez and Özak (2023) to identify the subsample of buffer zones where our two measures differ. I start by dividing the original and alternative measures into deciles, and for each buffer zone, and I calculate the absolute difference in decile rank between the two measures. I then divide the full sample

[1]See Figure 3 in Gonzalez and Özak (2023) for a relatable figure and example.

[2]The distribution of peoples within each ethnic group is crucial to understanding whether inter-ethnic trade is more likely here than intra-ethnic trade, which is also possible. For example, if trading partners are a function of distance, then peoples livings near the top or bottom of the hypothetical buffer zone in Figure 1 are more likely to engage in inter-ethnic trade, whereas incentives for intra- vs inter-ethnic trade are less clear for peoples living near the center. Unfortunately, data limitations are such that measuring the distribution of peoples during the historical period at such a fine spatial resolution is difficult.

Figure 2: Distribution of Positively and Negatively Correlated Buffer Zones by Region



(a) Ethnologue v16 (b) Ethnologue v17

into 10 subsamples, according to the absolute difference in decile rank, and calculate the correlation coefficient for pre-1500 land productivity variation (i.e., the original measure) and pre-1500 cross border change in land productivity (i.e., the alternative measure) in each subsample. Table 1 reports these correlation coefficients. The correlation between the two measures becomes negative for differences in decile rank greater than three, which corresponds to 11.3% of the v16 sample and 11.7% of the v17 sample.

What regions characterize these positive vs. negative correlation coefficients? Figure 2 plots the percentage of buffer zones by region, separately for buffers where the original and alternative measures are positively vs. negatively correlated. Oceania has the most buffer zones where the two measures are negatively correlated—both in aggregate terms and relative to the number of buffer zones where the two measures are positively correlated. In other words, the original and alternative measure are uniquely different in the Oceania region, suggesting that further exploration of why these measures differ across this region is one avenue towards better understanding the mechanism.

As a first step in this direction, Table 2 makes clear that the homelands associated with Oceania buffer zones are, on average, relatively smaller in population and area than any other region of the world. Both population and area are commonly used proxies for market size in the trade literature, and the implications for trade are interesting in this context. For example, large populations enable producers to offset their fixed costs along a larger consumer base so in equilibrium you get a larger set of unique varieties and specialized products (Krugman, 1980; Melitz, 2003). The small "market size" of Oceania suggests that cross-ethnic specialization may have been historically less common, which might explain why the alternative measure has less predictive power than a generalized measure of trade potential (Gonzalez and Özak, 2023). These cursory findings hint at potential explanations, but clearly more research is needed.

Table 2: Buffer Zone Population and Area by Region

	Ethnologue v16			Ethnologue v17		
Region	Mean ln(population)	Mean ln(area)	Obs.	Mean ln(population)	Mean ln(area)	Obs.
Africa	12.63	9.45	3,015	12.62	9.45	2,721
	(2.01)	(0.44)		(2.00)	(0.44)	
Americas	14.71	9.66	1,259	14.84	9.69	1,106
	(4.23)	(0.60)		(4.20)	(0.62)	
Asia	12.89	9.49	2,485	12.71	9.50	2,161
	(3.40)	(0.52)		(3.44)	(0.53)	
Europe	15.21	9.82	420	15.23	9.84	382
	(2.83)	(0.71)		(2.88)	(0.73)	
Oceania	8.72	9.21	1,223	8.70	9.20	1,194
	(2.56)	(0.31)		(2.51)	(0.29)	

This table documents population and area differences by region. ln(population) is the natural log of aggregate population for both groups associated with a buffer zone, and ln(area) is the natural log of aggregate land area of both ethnolinguistic group homelands (km^2). Standard deviations are in parentheses.

Concluding Remarks

Gonzalez and Özak (2023) provide a direct and successful replication of Dickens (2022) using both reconstructed data and an updated version of the original source data. In addition to the replication, Gonzalez and Özak (2023) make progress on interpretation of the mechanism. In their report, the authors argue that my original measure of trade potential is a better proxy for generalized trade, whereas their alternative measure better approximates the potential for cross-ethnic specialization and thus inter-ethnic trade. While this alternative measurement strategy is a fruitful step forward in better understanding the mechanism, it is not without its own shortcomings, as outlined in this response and as the report authors themselves acknowledge. Indeed, fully understanding the mechanism may require addressing border measurement issues and delving deeper into cross-ethnic crop specialization using more detailed data. It is also necessary to better understand when and where these two measures differ, insofar as this source of variation is an essential piece of the puzzle. Future research along these lines holds the answer to this interesting area of inquiry.

References

Depetris-Chauvin, E. and Özak, O. (2020). (De Facto) Historical Ethnic Borders and Contemporary Conflict in Africa. *IZA Discussion Paper Series*.

Dickens, A. (2022). Understanding Ethnolinguistic Differences: The Roles of Geography and Trade. *The Economic Journal*, 132(643):953–980.

Galor, O. and Özak, O. (2015). Land Productivity and Economic Development - Caloric Suitability vs Agricultural Suitability. *Brown University Working Paper*, 12 July 2015.

Galor, O. and Özak, O. (2016). The Agricultural Origins of Time Preference. *American Economic Review*, 106(10):3064–3103.

Gonzalez, J. and Özak, O. (2023). Replication of Dickens (2022) "Understanding Ethnolinguistic Differences: The Roles of Geography and Trade. *Institute for Replication*, 27 June 2023.

Krugman, P. (1980). Scale Economies, Product Differentiation, and the Pattern of Trade. *American Economic Review*, 70(5):950–959.

Lewis, M. P. (2009). *Ethnologue: Languages of the World*, Vol. 16, SIL International, Dallas, TX.

Melitz, M. J. (2003). The Impact of Trade on Intra-Industry Reallocations and Aggregate Industry Productivity. *Econometrica*, 71(6):1695–1725.