Racial admixture and socioeconomic outcomes: a meta-analysis

John Fuerst¹
Emil O. W. Kirkegaard²

Abstract
We meta-analyzed studies reporting associations between socioeconomic outcomes (S outcomes) and racial ancestries (European, African and Amerindian) in the Americas. 33 studies yielded a total of 120 datapoints and 41 non-overlapping effect sizes.

We found that European ancestry is positively related to S outcomes $r = .17$ [95% CI: .12 to .21], Amerindian negatively $-.11$ [-.16 to -.06] and African negatively $-.09$ [-.17 to -.01].

Key words: ancestry, biogeographic ancestry, race, ethnicity, SIRE, income, education, admixture mapping

1. Introduction

The mean level of important socioeconomic outcomes (S, e.g. income, education, health³) vary by self-identified race/ethnicity (SIRE) in almost all countries examined. Furthermore, it varies in systematic ways, such that SIREs that are more European are better off than SIREs that are less European. Many causes have been proposed for this pattern, especially in terms of discrimination and remnants of slavery (e.g.). Others have proposed that the differences seen in S is an effect of differences in

¹ Independent researcher, USA. Email: j122177@hotmail.com.
² University of Aarhus, Denmark. Email: emil@emilkirkegaard.com.
³ If a collection of socioeconomic measures are analyzed, a common factor usually emerges such that positive outcomes have positive loadings and negative outcomes negative loadings. This factor has been named the S factor (see e.g. Kirkegaard, 2014; Fuerst & Kirkegaard, 2016). We strain language a little by using $S$ both to refer to the general factor of socioeconomic outcomes and socioeconomic outcomes in general.

The link between SIREs and certain health outcomes has been noted in the medical literature. Due to advances in technology, it is now fairly cheap to genotype individuals with thousands of markers (usually SNPs). To determine the causes of differences in disease rates between SIREs, medical researchers have often employed estimates of racial ancestry. Often, ancestry is just included as a covariate and is not the primary variable of interest. Still, this means that there are now many studies that report on relationships between S and ancestry. If the hereditarian model with regards to CA is true and CA causes higher S, one would expect to find that the racial ancestries statistically associated with higher CA are also statistically associated with higher S. Figure 1 shows a path diagram.

An alternative path from ancestry to S (R~~other traits→S; ~~ means correlates with) is included but has not been found or ruled out yet. However, SIRE groups do not seem to differ much in other traits known to be important for S (Dalliard, 2014). For instance, the personality trait Conscientiousness (C) is probably the trait that has the most consistently positive associations with S (Bogg & Roberts, 2004, 2013; Dudley, Orvis, Lebiecki, & Cortina, 2006; Ng & Feldman, 2010; Shaffer & Postlethwaite, 2013). However, a very large meta-analysis of the Black White differences in C found no difference (d=.02, N=193k, K=81) (Tate & McDaniel, 2008). The other differences in the Big Five traits were similarly tiny. Thus, the Black-White difference in S cannot be explained by C or other Big Five traits.

Many recent studies show that CA is itself related to health outcomes at the individual level (Calvin et al., 2011; Deary, 2009; Der, Batty, & Deary, 2009; Gottfredson, 2004; Wraw, Deary, Gale, & Der, 2009).

\(^4\) It is also called general cognitive ability, general mental ability, complex cognitive ability, intelligence, general intelligence, g and so on (Jensen, 1998).
2015). However, medical researchers still do not generally include measures of CA as covariates in their studies (see also Lubinski & Humphreys, 1997; Lubinski, 2009). For this reason, there are few studies that report genomic associations between CA and racial ancestry. Thus, we could not meta-analyze the studies that report on the ancestry−−CA link, but we could meta-analyze studies that report on the ancestry−−S link.

2. **Methods**

2.1. **Collecting studies**

We relied on Google Scholar as our academic search engine and used search phrases such as “admixture African/Amerindian socioeconomic/education”. In total, these searches turned up approximately 20,000 hits in descending order of relevance to the search terms. The first 1,500 abstracts were skimmed. Approximately 250 papers were identified as potential sources and read. Over the next year we expanded this search using multiple search engines, such as Medline.\(^5\)

Many studies did not report effect sizes, or some other statistic that we were able to convert to an effect size. In fact, many studies did not even report directions for relationships. In every case of where we could not find or calculate an effect size, we wrote to the authors to request that they provide us with it, or provide the case-level data so that we may calculate it ourselves. In general, few authors replied to our emails with the missing effect sizes and in no case did they share case-level or other raw data.

2.2. **Coding of studies**

JF did most of the coding. When coding was mostly completed, both authors went over a large fraction of the studies and reached agreement in disputed cases. We recorded the following pieces of information for each datapoint:

- Author-year (APA format).
- Type of sample (medical, control, combined, etc.).
- Country.
- First order administrative division within country (variously called states, departments, regions).
- Specific region such as city.
- The subpopulation examined (African American, Hispanic American, Puerto Rican, etc.).
- The sample name.

\(^5\) We used searches such as: (admixture) AND (socioeconomic or education or income or SES or poverty) AND (African or European or Amerindian) AND (Antilles OR Latin America OR South America OR Central America OR Caribbean OR Anguilla OR Antigua OR Aruba OR Barbuda OR Argentina OR Bahamas OR Barbados OR Belize OR Bolivia OR Brazil OR Chile OR Colombia OR Costa Rica OR Dominica OR Dominican Republic OR Ecuador OR El Salvador OR Grenada OR Grenadines OR Guadeloupe OR Guatemala OR Guyana OR Haiti OR Honduras OR Jamaica OR Martinique OR Mexico OR Montserrat OR Nevis OR Nicaragua OR Panama OR Paraguay OR Peru OR Puerto Rico OR Saint Kitts OR Saint Lucia OR Saint Vincent OR Suriname OR Surinam OR Trinidad OR Tobago OR Uruguay OR Venezuela).
- A sample ID within each study.
- Sample size.
- Ancestry examined (European, Amerindian, African).
- Mean level of admixture (arithmetic mean).
- Standard deviation of admixture.
- Outcome category (e.g. SES, income).
- Outcome literal. The specific type of outcome mentioned in the source.
- The association direction (positive, negative, null, not stated).
- The p value.
- Which test was used to derive the p value.
- The correlation.
- The conversion method. If the correlation was derived from other statistics, note the method.
- Details for the conversion.
- Details for author contact attempts (multiple columns).

2.3. **Descriptive statistics of studies**

A total of 33 studies were coded for the meta-analysis yielding a total of 120 datapoints. Not all of these actually reported meta-analyzable information such as direction of effect or effect size.

2.3.1. **Country**

Of the datapoints, about 2/3 came from the US. Table 5 shows a breakdown by country.
<table>
<thead>
<tr>
<th>Country</th>
<th>Frequency</th>
</tr>
</thead>
<tbody>
<tr>
<td>U.S.</td>
<td>78</td>
</tr>
<tr>
<td>Brazil</td>
<td>11</td>
</tr>
<tr>
<td>Mexico</td>
<td>7</td>
</tr>
<tr>
<td>Uruguay</td>
<td>6</td>
</tr>
<tr>
<td>Chile</td>
<td>3</td>
</tr>
<tr>
<td>Colombia</td>
<td>3</td>
</tr>
<tr>
<td>Costa Rica</td>
<td>3</td>
</tr>
<tr>
<td>Peru</td>
<td>3</td>
</tr>
<tr>
<td>Trinidad and Tobago</td>
<td>3</td>
</tr>
<tr>
<td>Peru and Chile</td>
<td>2</td>
</tr>
<tr>
<td>Argentina</td>
<td>1</td>
</tr>
</tbody>
</table>

*Table 1: Datapoints by country.*

2.3.2. Year of publication

While we only searched for papers published 2003-2015, we also followed references found in the identified papers. For this reason, not all papers were published in the designated window. Figure 2 shows a histogram of the publication year.

![Histogram of datapoints by publication year.](image)

*Figure 2: Histogram of datapoints by publication year.*

Most datapoints were from papers published in the last few years with the median being 2013. We can expect a lot more relevant papers being published in the next few years.

2.3.3. Sample sizes

Studies varied widely in the size of their samples. Figure shows a combined histogram-density plot of the sample sizes.
2.3.4. Sample type

Many studies had a medical theme and often included both case and control samples. Table 2 shows the breakdown for each sample type.

<table>
<thead>
<tr>
<th>Sample type</th>
<th>Frequency</th>
</tr>
</thead>
<tbody>
<tr>
<td>case</td>
<td>14</td>
</tr>
<tr>
<td>case and control</td>
<td>34</td>
</tr>
<tr>
<td>control</td>
<td>19</td>
</tr>
<tr>
<td>non-medical</td>
<td>53</td>
</tr>
</tbody>
</table>

Table 2: Datapoints by sample type.

2.3.5. Ethnicity

Studies often divided their samples into ethnic groups. Table 3 shows the breakdown.
<table>
<thead>
<tr>
<th>Ethnic group</th>
<th>Frequency</th>
</tr>
</thead>
<tbody>
<tr>
<td>Multi-ethnic</td>
<td>24</td>
</tr>
<tr>
<td>Puerto Rican</td>
<td>19</td>
</tr>
<tr>
<td>African American</td>
<td>16</td>
</tr>
<tr>
<td>Not stated</td>
<td>16</td>
</tr>
<tr>
<td>Mexican</td>
<td>15</td>
</tr>
<tr>
<td>Hispanic-American</td>
<td>14</td>
</tr>
<tr>
<td>Latino</td>
<td>4</td>
</tr>
<tr>
<td>Afro-descent</td>
<td>3</td>
</tr>
<tr>
<td>African American (non-African born)</td>
<td>2</td>
</tr>
<tr>
<td>American-Indian</td>
<td>2</td>
</tr>
<tr>
<td>Native American</td>
<td>2</td>
</tr>
<tr>
<td>Caucasian</td>
<td>1</td>
</tr>
<tr>
<td>Mestizo</td>
<td>1</td>
</tr>
<tr>
<td>Mexican-American</td>
<td>1</td>
</tr>
</tbody>
</table>

Table 3: Datapoints by ethnic group.

2.3.6. Ancestries

When researchers examine the relationship between racial admixture and outcomes, they must necessarily choose one ancestry to focus on. Sometimes, they report correlations with each ancestry type for a sample, but often they only report one, e.g. Amerindian for Mestizos. Table 4 shows the breakdown of each mentioned association.

<table>
<thead>
<tr>
<th>Ancestry</th>
<th>Frequency</th>
</tr>
</thead>
<tbody>
<tr>
<td>African</td>
<td>47</td>
</tr>
<tr>
<td>Amerindian</td>
<td>37</td>
</tr>
<tr>
<td>European</td>
<td>36</td>
</tr>
</tbody>
</table>

Table 4: Breakdown of ancestry associations mentioned.

2.3.7. Outcome categories

Studies reported a large variety of outcome variables. To make things clearer, we coded into them categories so that results can be aggregated within a single category. Table 5 shows the breakdown of outcomes by category.
<table>
<thead>
<tr>
<th>Category</th>
<th>Frequency</th>
</tr>
</thead>
<tbody>
<tr>
<td>SES</td>
<td>63</td>
</tr>
<tr>
<td>Education</td>
<td>44</td>
</tr>
<tr>
<td>Education and SES</td>
<td>5</td>
</tr>
<tr>
<td>Parental SES</td>
<td>5</td>
</tr>
<tr>
<td>Parental education</td>
<td>3</td>
</tr>
</tbody>
</table>

Table 5: Outcomes by category.

Most data concerned some composite measure of SES (socioeconomic status), but a large fraction only had education-related outcomes. Some reported associations with parental outcomes.

2.4. Methods

Meta-analyzing the present dataset presents multiple difficulties.

2.4.1. Meta-analysis method

We used random effects meta-analysis to analyze the effect size results. Standard errors for the correlations were calculated from the sample sizes and the reported correlations. Random effects models are appropriate when the observations cannot be assumed to have been sampled from a single population (J. E. Hunter & Schmidt, 2004, p. 393). Since our datapoints contain information from different sub-populations, different countries and using different outcomes, a random effects model was clearly appropriate.

2.4.2. Multiple outcome measures for a sample

Sometimes associations with multiple outcome measures are reported for a single sample and the same ancestry. Using simple aggregation methods means that samples reporting results for multiple outcome categories count more than those that report fewer. This problem was also encountered in a recent meta-analysis of the association of SES and heritability (Tucker-Drob & Bates, 2016). Their method was to use a complex weighting approach to avoid the double counting. We chose to implement a simpler approach, namely to average (median) values across outcomes within each sample before meta-analyzing them.

For instance, Norden-Krichmar et al. (2014) reported associations between Amerindian ancestry, and SES and income of -.02 and -.11 for one sample. Using a simple analysis, this would get counted as two independent studies. Instead, using the aggregated dataset, it is counted as one study with an association of -.065.

2.4.3. Multiple ancestries x outcomes reported for a sample

Authors may report one or more ancestry x outcome associations for a single sample. Using simple aggregation methods means that some samples will be counted more than once altho not for the same ancestry. We did not take any action against this kind of double counting.
2.4.4. Converting reported statistics to effect sizes

Sometimes we were not sure about our method for converting reported statistics to effect sizes. To guard against conversion errors, we created a subset of data where only studies that reported effect sizes and conversion methods we were sure about were included (e.g. converting from $R^2$ to $r$).

2.4.5. Individual vs. parental outcomes

Some studies reported associations with parental outcomes (cf. Section 2.3.7). We created a subset of the data that only included individual outcomes.

2.4.6. Samples

Two parallel datasets were created. The first (total dataset) includes all the data, including multiple outcomes from a single sample. The second (restricted dataset) includes only effect size conversions we were sure about, only one aggregated outcome for each ancestry per sample, and only individual outcomes (not parental).

3. Results

The present data can be analyzed in many ways. Some of these are presented in the following sections.

3.1. Directions of effects

The mere direction of effect sizes is more frequently reported than the actual effect sizes. Because the effect directions of effects depends on both the true effect size and the distribution of sample sizes, it cannot be used as a measure of effect size. It can, however, give a rough idea of whether the findings are in line with a null hypothesis or not. If the effect directions deviates strongly from that expected by chance, then it is likely that the true effect is in that direction, altho one cannot say how strong it is. If the effect directions do not deviate from chance levels, there may be an effect but sample sizes are too small to reliably see it, or there may be no effect. These conclusions only hold given the assumption of no publication or reporting bias.

We calculated the effect directions in two ways. In the first, we included studies that did not report the direction, and in the second, we excluded them. When effect directions are not reported, it is probably in many cases due to p values being above the cutoff. Tables 6 show the results for all datapoints.
### Table 6: Directions of effects by ancestry.

<table>
<thead>
<tr>
<th>Ancestry</th>
<th>N</th>
<th>positive</th>
<th>negative</th>
<th>null</th>
<th>not stated</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>With unstated directions.</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>African</td>
<td>47</td>
<td>0.06</td>
<td>0.77</td>
<td>0.06</td>
<td>0.11</td>
</tr>
<tr>
<td>Amerindian</td>
<td>37</td>
<td>0.08</td>
<td>0.73</td>
<td>0.08</td>
<td>0.11</td>
</tr>
<tr>
<td>European</td>
<td>36</td>
<td>0.86</td>
<td>0.03</td>
<td>0</td>
<td>0.11</td>
</tr>
<tr>
<td><strong>Without unstated directions.</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>African</td>
<td>47</td>
<td>0.07</td>
<td>0.86</td>
<td>0.07</td>
<td></td>
</tr>
<tr>
<td>Amerindian</td>
<td>37</td>
<td>0.09</td>
<td>0.82</td>
<td>0.09</td>
<td></td>
</tr>
<tr>
<td>European</td>
<td>36</td>
<td>0.97</td>
<td>0.03</td>
<td>0</td>
<td></td>
</tr>
</tbody>
</table>

In general, the directions are as one would expect, namely that African and Amerindian tend to be negatively related to outcomes while European tends to be positively.

Note that because these are based on the total dataset, there is some double counting when a study reports more than one outcome for a single sample x ancestry combination.

### 3.2. Overall distributions of results

To get an overview of the findings, it is useful to plot a density-histogram plot with findings for each ancestry separate. This is shown in Figure 4 for the restricted dataset.

*Figure 4: Density-histogram plot for effect sizes by ancestry.*

Table 7 shows descriptive statistics for the same data.
Table 7: Descriptive statistics for ancestry x outcome correlations.

There is one clear outlier for European ancestry at -.20, which we will return to later. The African effect sizes were more spread out than the others (sd .15 vs. .11), probably due to lack of inter-individual variation in the sample.

### 3.3. Formal meta-analysis

Using the restricted dataset, we analyzed the effect sizes for each ancestry. Figures 5 to 7 show the forest plots.

---

**Figure 5:** Forest plot for European ancestry results. Based on random effects model.

---

6 We used the **metafor** package for R (Viechtbauer, 2015).
Figure 6: Forest plot for Amerindian ancestry results. Based on random effects model.
We note that the far outlier for European ancestry (Zou et al. sample 6) is also an outlier for Amerindian ancestry. If one looks at the study's supplementary material, one sees that R2 values for their sample 6 (Mexicans, Chicago 2) are exactly the reverse of those in their sample 4 (Mexicans, Oakland). Thus, some kind of reverse coding in their analysis or reporting seems likely. We contacted the lead author (James Zou) to ask whether there is a reporting error. He checked the data and reports that this is not so. He urges caution and furthermore explained that there is some degree of self-selection in these samples which might have created the odd result.

We created a third dataset by excluding the outlier and rerunning the analyses. This gave similar and slightly stronger results and the between study heterogeneity ($I^2$) was somewhat reduced as expected (85%, 95%, 80% to 75%, 95%, 64% for European African and Amerindian ancestry, respectively).

---

7 Specifically, the results were .18 [.14 to .21], -.13 [-.17 to -.09] and -.11 [-.19 to -.02] for European, Amerindian and African ancestry, respectively.
4. Discussion and conclusion

4.1. Implications for epidemiological studies

We found fairly robust effects of European and Amerindian ancestry on S outcomes, and somewhat less robust results for African ancestry. To avoid spurious effects in regressions due to omitted variable bias, it is thus important to include these as co-variates in studies of medical outcomes.

4.2. Untangling effects of racial ancestry and SIRE

In many of the studies included in this meta-analysis, individual outcomes are associated with genomic ancestry within SIRE groups from which we can conclude that SIRE membership is not mediating the relationship perfectly. For example, genomic ancestry correlates with outcomes in US African and Hispanic SIRE populations and within African Trinidadians.

In other cases, particularly in Latin America, the issue is less clear and genomic ancestry may be confounded with SIRE. For example, Leite, et al. (2011) found that African ancestry was negatively correlated with SES in Brasilia. This could be because it was negatively correlated with ancestry net of SIRE or because it was negatively correlated with SIRE but not with ancestry net of SIRE group. The analysis of Ruiz-Linares et al. (2014) has helped clarify the issue. The authors looked at the association between genotype, color and SIRE in a multi-country sample from Brazil, Chile, Colombia, Mexico and Peru (mean age 20 to 25, country depending). The authors found modest correlations between racial ancestry and SIRE (e.g., 0.48 in the case of both European/White and Amerindian). The authors found that wealth and educational attainment robustly correlated with European ancestry (r = .12 for the full sample). However, net of genotype, education was not associated with SIRE. Wealth was only marginally so (B = 0.00291, p = 6.1×10^-4) and only with regards to the European/White group. Despite this, net of genomic ancestry, SIRE was found to be a significant predictor of racially associated phenotypes such as melanin index, hair shape, eye color and eye fold. Figure 8 shows a proposed model consistent with these findings.
In this way, the statistical association between SIRE and S outcomes is thru the common association with racial ancestry, hence the disappearance of the relationship when that link is controlled. The causal path from racial phenotype to SIRE reflects the fact that this association is still found when racial ancestry is controlled.

It may be necessary to modify the above model to incorporate additional nodes, such as discrimination based on racial phenotypes (as argued by e.g. M. Hunter, 2007; Telles, 2014). Quite possibly, racial phenotype based discrimination is present in some countries and absent in others or reversed in direction. In the US, one would expect that groups that benefit from affirmative action policies (African Americans, Hispanics) to do better than expected based on their levels of human capital. In general, this topic is understudied, but several datasets exist that could shed light on it.

4.3. Moderator analyses

We did not conduct analyses for publication bias because the results we aggregated were generally not the focus point of the studies and often reported in supplementary materials. There would thus seem to be little reason to expect publication bias for them. In case, the number of studies was probably too small to detect it if it was present.

The results varied quite a bit across studies (mean effect size heterogeneity = 87%). This means that there are probably moderators that affect the effect size. However, due to the small number of studies in the restricted dataset, we did not conduct moderator analyses. In a few years when more studies have been published, one could update this meta-analysis and conduct moderator analyses.

4.4. Limitations

- Many studies did not report effect sizes and for some of them, the authors did either not reply to our emails or replied with excuses. The present meta-analysis is thus smaller that it could have
been, if scientists had published their results in a manner consistent with scientific principles.

- It is possible that there is publication/reporting bias in the literature. The present dataset was judged to be too small to properly investigate this issue. A future, larger meta-analysis should examine this question.

**Supplementary material and acknowledgments**

The dataset, high quality figures and R analysis code are available in the supplementary materials at the repository at Open Science Framework https://osf.io/ydc3f/files/.

The peer review thread can be found at .

Thanks to ...

**References**


