

# Crossing the Line Replication

Angela Fu

5/8/2020

## 0.1 Abstract

In their paper “Crossing the Line: Local Ethnic Geography and Voting in Ghana,” Nahomi Ichino and Noah Nathan found that the local ethnic demographics of the area in which Ghanaian voters live affected who they chose to support in the 2008 Ghana presidential election. I successfully replicated Ichino and Nathan (2013)’s results. I have also reexamined their models and added additional variables to explore the effect of a voter’s trust in members of their own ethnicity and members of other ethnicities. I found that the inclusion of this variable does not significantly alter the model. As a result, trust in one’s co-ethnics or in others of a different ethnicity do not impact one’s voting decisions as much as the demographics of one’s location do.

## 0.2 Introduction

Researchers have long sought to explore how a voter’s ethnicity affects their voting preferences. This issue becomes more salient in countries where patronage politics is an issue. It is commonly thought that voters will support candidates who share their same ethnicity because they expect that candidate to then favor other co-ethnics. Ichino and Nathan (2013), however, shows that the demographics of the area in which a voter lives also matter. Using voting data from the 2008 Ghana presidential election and survey data from the Afrobarometer, they found that a voter who lives in an area dominated by one ethnicity is more likely to support a candidate whose ethnicity matches that of the area in which they live, not their own ethnicity. In replicating Ichino and Nathan (2013), I used R. The data and code come from Harvard Dataverse. The code in this paper can be found on GitHub<sup>1</sup>.

I have chosen to reexamine the regressions Ichino and Nathan (2013) built to predict the likelihood that an individual support the New Patriotic Party or the National Democratic Congress. I have added variables

---

<sup>1</sup>The repo for this project is here

that measure whether an individual trusts co-ethnics and whether they trust non-co-ethnics, as those two factors will likely influence their voting decisions. After running the new regressions, however, I found that the primary results from Ichino and Nathan (2013) did not change significantly.

### 0.3 Literature Review

The effect ethnicity has on elections in sub-Saharan Africa has long been studied by researchers. This is because many politicians engage in patronage, and voters have come to expect that the provision of public goods will be determined in part by ethnicity. Essentially, voters expect that politicians will provide more goods to people who share their same ethnicity.

The view that ethnicity and voting are linked has come to be widely accepted. Research has shown, for example, that countries that are ethnically heterogeneous have resulted in the under-provision of non-excludable public goods but a high prevalence of patronage goods (Kimenyi 2006).

Though this view has been widely adopted, increasing research has shown that the link may not be so simple. For example, research using results from the Afrobarometer, a series of surveys distributed across the continent, shows that voters take other factors into consideration. For example, one study found that voters in addition to considering ethnic ties, will also base their votes on economic outcomes (Michael Bratton and Chen 2011).

The case of Ghana in particular is interesting because the results of its 2008 presidential election do not conform with the theory that ethnicity and voter preferences are so simply linked. The two parties, the National Democratic Congress and the New Patriotic Party, saw support from voters across ethnic lines. One theory for this phenomenon is that the perceptions of the individual parties and candidates had more of an effect on how voters cast their ballot (Barak Hoffman and Long 2009).

However, Ichino and Nathan have found instead that while ethnicity does play a role, it does not do so in the way most researchers would expect. Though some voters certainly do vote along ethnic lines, others vote across those lines when they live in an area that is heavily populated by members of another ethnic group. The reason behind this is that they expect the candidate they choose to better provide public goods to areas in which they saw the most support. In Ghana, the president has a high amount of influence over how public goods are distributed. These goods are often nonexcludable, so everyone living in the area benefits, regardless of their ethnicity. Thus, it is in voters' best interest to vote across ethnic lines if they are surrounded by members of a different ethnicity.

## 0.4 Paper Review

Ichino and Nathan (2013) first examines data at the polling station level. Using 2008 presidential election results from the Brong Ahafo Region, a rural, ethnically diverse region of Ghana, they created models predicting support for the NPP based on characteristics of the area surrounding a polling station. They found that the proportion of Akans in the area and at the polling station were a key factor in predicting NPP vote share. Areas with a high percentage of Akans had polling stations report significantly greater support for the NPP, even after controlling for the ethnic demographics at the individual polling stations. This makes sense because in Ghana, the NPP is closely associated with the Akan ethnic group. Similarly, the NDC is closely associated with the Ewe ethnic group. Other ethnic groups are split fairly evenly in their support for either party.

Though the first part of Ichino and Nathan (2013) confirms the trend one expects to see if local demographics influence voter choices, it does not confirm individual voter behavior. As a result, the second part of Ichino and Nathan (2013) examines survey data from the 2005 and 2008 rounds of the Afrobarometer to create models at the individual level. The Afrobarometer includes information about enumeration areas, which allowed Ichino and Nathan (2013) to determine the local demographics of the respondent's area.

Ichino and Nathan created four main models. The logistic models predict the probability that a respondent supports either the NPP or the NDC. Their models include variables representing the proportions of various ethnic groups in the respondent's area, as well as the respondent's gender, their opinion of the current economic condition compared to a year ago, their poverty level, and the survey round in which they participated. The models showed statistically significant relationships between the population shares of the Akan and Ewe and support for the NPP and NDC.

## 0.5 Replication

I was mostly able to replicate the major results of Ichino and Nathan (2013). There are some discrepancies between the clustered standard errors in my models, compared to theirs, which likely stem from their decision to use Stata to create the models. I used `multiwaycov` to create clusters instead and received very similar standard errors. Ichino and Nathan (2013) also included some calculations for the clusters done in R, but I had to modify the code as theirs used R packages that are no longer supported.

## 0.6 Extension

Though Ichino and Nathan (2013)'s primary models do not include the variables measuring trust, the paper does test several models with the addition of `trustother`. However, these models are limited in that they are only examined in the context of non-Akans. Furthermore, only the variable `trustother`, which measures the degree to which respondents reported trusting members of other ethnic groups, is added.

I added both `trustown`, which examines the degree to which respondents reported trusting members of their own ethnic group, and `trustother` to the primary models. The additions of these variables allows for exploration as to whether voters are primarily driven by their desire for club goods or their personal perceptions of the candidate. Furthermore, a voter who does not trust a candidate may decide they they cannot count on the candidate to reward their supporters. Because I have added `trustown` and `trustother`, I eliminated `r4`, which indicates whether the respondent came from round 3 or round 4 of the Afrobarometer. This is because only round 3 included questions about trust.

I also split up the data so that each model is applied to a different ethnic group, and I removed ethnic group variables accordingly. The first table shows the models for members of the Akan ethnic group.

|   | NPP     | NDC       | NPP      | NDC       |
|---|---------|-----------|----------|-----------|
|   | (1)     | (2)       | (3)      | (4)       |
| % Akan in 30 km (spatially weighted)                  | 1.212*  | −2.580*** |          |           |
|   | (0.714) | (0.836)   |          |           |
| % Akan in 30 km (spatially weighted) * Central Region | −0.388  | 2.447*    |          |           |
|   | (1.334) | (1.294)   |          |           |
| % Ewe in 30 km (spatially weighted)                   |         |           | −2.316** | 4.362***  |
|   |         |           | (1.154)  | (1.429)   |
| Male  | 0.001   | 0.023     | 0.031    | −0.030    |
|   | (0.205) | (0.258)   | (0.207)  | (0.260)   |
| Economic approval                                     | 0.219** | −0.506*** | 0.226**  | −0.524*** |
|   | (0.103) | (0.157)   | (0.103)  | (0.154)   |
| Poverty (individual)                                  | −0.098  | 0.149     | −0.091   | 0.124     |
|   | (0.106) | (0.137)   | (0.105)  | (0.137)   |
| Development index (EA)                                | −0.319* | −0.038    | −0.311*  | −0.049    |
|   | (0.174) | (0.214)   | (0.175)  | (0.218)   |
| Central Region  | 0.009   | −1.706*   | −0.252   | 0.112     |
|   | (1.133) | (0.955)   | (0.277)  | (0.320)   |
| Trust Own   | 0.146   | 0.199     | 0.140    | 0.242     |
|   | (0.259) | (0.362)   | (0.257)  | (0.354)   |
| Trust Other   | −0.007  | −0.718*   | 0.003    | −0.724*   |
|   | (0.290) | (0.381)   | (0.287)  | (0.380)   |
| Intercept   | −0.966  | 1.229*    | 0.024    | −0.833*   |
|   | (0.653) | (0.661)   | (0.355)  | (0.486)   |
| N   | 470     | 470       | 470      | 470       |

Notes: \* $p < 0.05$ ; \*\* $p < 0.01$ ; \*\*\* $p < 0.001$ . Logistic regression coefficients with standard errors clustered at enumeration area level in parentheses. Outcome in columns 1 and 3 is support for NPP; outcome in columns 2 and 4 is support for NDC. Data is from Rounds 4 of the Ghana Afrobarometer, excluding respondents from local areas estimated to have more than 1000 people per sq. km

The second table contains models applied to respondents who are not Akan.

|   | NPP                  | NDC                 | NPP                  | NDC                 |
|---|----------------------|---------------------|----------------------|---------------------|
|   | (1)                  | (2)                 | (3)                  | (4)                 |
| % Akan in 30 km (spatially weighted)                  | 1.427***<br>(0.423)  | -1.160**<br>(0.556) |                      |                     |
| % Akan in 30 km (spatially weighted) * Central Region | -3.235<br>(2.798)    | 3.873<br>(3.323)    |                      |                     |
| % Ewe in 30 km (spatially weighted)                   |                      |                     | -1.976**<br>(0.955)  | 2.454***<br>(0.832) |
| Ewe   | -0.177<br>(0.286)    | 0.850***<br>(0.303) | 0.663<br>(0.623)     | -0.357<br>(0.513)   |
| Dagomba (Mole)  | 0.548<br>(0.365)     | -0.003<br>(0.338)   | 0.213<br>(0.358)     | 0.288<br>(0.328)    |
| Male  | 0.190<br>(0.240)     | -0.298<br>(0.198)   | 0.201<br>(0.233)     | -0.297<br>(0.197)   |
| Economic approval                                     | 0.442***<br>(0.116)  | -0.218*<br>(0.120)  | 0.438***<br>(0.113)  | -0.226*<br>(0.122)  |
| Poverty (individual)                                  | -0.198<br>(0.126)    | 0.125<br>(0.135)    | -0.209*<br>(0.123)   | 0.113<br>(0.142)    |
| Development index (EA)                                | -0.066<br>(0.147)    | 0.0001<br>(0.191)   | 0.049<br>(0.155)     | -0.127<br>(0.200)   |
| Central Region  | 2.045**<br>(0.893)   | -1.551<br>(1.772)   | 0.795*<br>(0.424)    | 0.024<br>(0.709)    |
| Trust Own   | 0.125<br>(0.258)     | 0.133<br>(0.261)    | 0.227<br>(0.258)     | 0.029<br>(0.268)    |
| Trust Other   | 0.382<br>(0.299)     | -0.029<br>(0.285)   | 0.314<br>(0.317)     | 0.057<br>(0.299)    |
| Intercept   | -2.335***<br>(0.428) | -0.0002<br>(0.395)  | -1.796***<br>(0.406) | -0.436<br>(0.389)   |
| N   | 431                  | 431                 | 431                  | 431                 |

Notes: \* $p < 0.05$ ; \*\* $p < 0.01$ ; \*\*\* $p < 0.001$ . Logistic regression coefficients with standard errors clustered at enumeration area level in parentheses. Outcome in columns 1 and 3 is support for NPP; outcome in columns 2 and 4 is support for NDC. Data is from Rounds 4 of the Ghana Afrobarometer, excluding respondents from local areas estimated to have more than 1000 people per sq. km

Though there are some differences between these new models and the old ones, the primary conclusions Ichino and Nathan found do not change. The spatially weighted proportions of Akan and Ewe in the respondent's area still remain significant predictors of support for either party. Similarly, the respondent's view on how the economy has changed is still statistically significant.

Trust in co-ethnics and non-co-ethnics does not appear to be a major predictor for non-Akans. However, for Akans, trust in others does appear to matter. Interestingly, the relationship is negative, meaning Akans who reported trusting non-co-ethnics are less likely to vote for the party associated with a different ethnic group.

Because the addition of `trustown` and `trustother` did not significantly alter the original findings, most of Ichino and Nathan (2013)'s conclusions can be considered more robust than they initially claimed. Their model still produces the same results even after the addition of two other variables, indicating that local demographics play a major role in a voter's decision calculus. This is especially true for non-Akans. However, within the Akan group, trust appears to play some role, but it is still not as significant as local demographics.

## 0.7 Conclusion

Ichino and Nathan (2013) confirm that ethnic voting may not be as straightforward as some believe. Through their analysis of the 2008 presidential election in Ghana, they find that local ethnic geography can affect individuals' voting patterns. For example, an Akan who lives in an area heavily populated by Ewes may decide to vote for the Ewe-associated NDC party. Voters choose to cross ethnic lines in their voting because they expect politicians to reward areas in which they received heavy support, i.e. areas that are heavily populated but co-ethnics.

I replicated Ichino and Nathan's findings using R. All of the code used in this paper can be found on Github, an example of a successful replication of one of their key findings is attached to the Appendix. The original code and data used for Ichino and Nathan's paper can also be found on Github and in the Harvard Dataverse.

After replicating Ichino and Nathan's findings, I added variables measuring trust in co-ethnics and non-co-ethnics to their models. I then ran the new models on two subsets of the data, Akans and non-Akans. In doing so, I found that the primary findings of Ichino and Nathan's original paper—that local geography is closely associated with voter preference—remained largely unchanged. This is a testament to the robustness of their original conclusions.

Though the primary findings did not change much, I did find that trust in non-co-ethnics did play some role in voter choice among Akans, especially in determining their support for the NDC party. Surprisingly, the

relationship was negative and could be the subject of additional research in the future.

## 0.8 Appendix

Results from Ichino and Nathan (2013) were successfully replicated. As an example, here is Table 2 from page 351.

| <b>TABLE 2. Local Ethnic Geography and NPP Vote Share in the 2008 Presidential Election</b> |                               |                     |                     |
|---|-------------------------------|---------------------|---------------------|
|   | (1)                           | (2)                 | (3)                 |
| % Akan in 30 km (spatially weighted)  |                               | 0.316***<br>(0.082) | 0.336***<br>(0.091) |
| % Akan in 30 km (spatially weighted) *  |                               |                     | -0.055<br>(0.109)   |
| % Akan at polling station   | 0.386***<br>(0.055)           | 0.286***<br>(0.061) | 0.325***<br>(0.098) |
| % Mole-Dagbon at polling station  | 0.120 <sup>†</sup><br>(0.063) | 0.075<br>(0.064)    | 0.075<br>(0.064)    |
| % Minor ethnic groups at polling station  | 0.228***<br>(0.054)           | 0.260***<br>(0.054) | 0.265***<br>(0.055) |
| % Public sector employment  | 0.086<br>(0.121)              | 0.083<br>(0.120)    | 0.070<br>(0.123)    |
| Development index (EA)  | 0.025***<br>(0.003)           | 0.024***<br>(0.003) | 0.024***<br>(0.003) |
| $R^2$   | 0.377                         | 0.383               | 0.383               |

*Notes:* <sup>†</sup>Significant at  $p < 0.10$ ; \* $p < 0.05$ ; \*\* $p < 0.01$ ; \*\*\* $p < 0.001$ .  $n = 1590$ . Weighted least squares with weights equal to total votes cast in the first round of the 2008 presidential election by polling station. All models include parliamentary constituency fixed effects (not shown). Standard errors in parentheses.

Figure 1: Table 2



|  | (1)                 | (2)                 | (3)                 |
|--|---------------------|---------------------|---------------------|
| % Akan in 30 km (spatially weighted)                             |                     | 0.316***<br>(0.082) | 0.336***<br>(0.091) |
| % Akan in 30 km (spatially weighted) * % Akan at polling station |                     |                     | -0.055<br>(0.109)   |
| % Akan at polling station  | 0.386***<br>(0.055) | 0.286***<br>(0.061) | 0.325***<br>(0.098) |
| % Mole-Dagbon at polling station                                 | 0.120*<br>(0.063)   | 0.075<br>(0.064)    | 0.075<br>(0.064)    |
| % Minor ethnic groups at polling station                         | 0.228***<br>(0.054) | 0.260***<br>(0.054) | 0.265***<br>(0.055) |
| % Public sector employment                                       | 0.086<br>(0.121)    | 0.083<br>(0.120)    | 0.070<br>(0.123)    |
| Development index (EA)   | 0.025***<br>(0.003) | 0.024***<br>(0.003) | 0.024***<br>(0.003) |
| R <sup>2</sup>   | 0.377               | 0.383               | 0.383               |

Notes: Significant at  $p < 0.10$ ; \* $p < 0.05$ ; \*\* $p < 0.01$ ; \*\*\* $p < 0.001$ .  $n = 1580$ . Weighted least squares with weights equal to total votes cast in the first round of the 2008 presidential election by polling station. All models include parliamentary constituency fixed effects (not shown). Standard errors in parentheses

## Bibliography

- Barak Hoffman, Karen E. Ferree, Clark Gibson, and James D. Long. 2009. *Explaining the African Vote*. American Political Science Association. [https://papers.ssrn.com/sol3/papers.cfm?abstract\\_id=1450123](https://papers.ssrn.com/sol3/papers.cfm?abstract_id=1450123).
- Ichino, Nahomi, and Noah L. Nathan. 2013. *Crossing the Line: Local Ethnic Geography and Voting in Ghana*. American Political Science Association. <https://www.jstor.org/stable/43654018>.
- Kimenyi, Mwangi S. 2006. *Ethnicity, Governance and the Provision of Public Goods*. Journal of African Economies. [https://academic.oup.com/jae/article-abstract/15/suppl\\_1/62/707092?redirectedFrom=fulltext](https://academic.oup.com/jae/article-abstract/15/suppl_1/62/707092?redirectedFrom=fulltext).

Michael Bratton, Ravi Bhavnani, and Tse-Hsin Chen. 2011. *Voting Intentions in Africa: Ethnic, Economic or Partisan?* Afrobarometer. <https://afrobarometer.org/publications/wp127-voting-intentions-africa-ethnic-economic-or-partisan>.