

Statement on removal of certain data for “The Causal Effect of Intergroup Contact on Exclusionary Attitudes”

Ryan D. Enos, March 10, 2022

In the replication data¹ for my article “The Causal Effect of Intergroup Contact on Exclusionary Attitudes” (PNAS, 2014)², I inadvertently included the IP addresses of survey respondents. These were not used in any analysis. I have removed the data with those IP addresses because they contain potentially personally identifying information about my research subjects and posting them was potentially a violation of human subjects protections. Nothing else about the replication data has changed. I have contacted my university Institutional Review Board and PNAS about the matter. If anybody would like to request the original files, including IP addresses, they can contact me to arrange viewing privileges.

This came to my attention because a journalist at a website contacted my university public relations office about receiving information from an anonymous source about alleged fraud in an article I had published in PNAS. An internet search revealed that somebody had posted these IP addresses on an online rumors site and invited people to geocode them. When geocoding the addresses, not all of the addresses are geocoded as being in Massachusetts. Because the research reported in the article was conducted in Massachusetts, this was interpreted as fraud.

The explanation for why the IP addresses are outside of Massachusetts is simple: IP addresses are linked to Internet Service Providers (ISPs) that are often located somewhere other than where the person accessing the internet is physically located. In fact, when I Google “are ip geocodes accurate”, the first reported result [“says IP-based geolocation services provide 55 percent to 80 percent accuracy for a user’s region or state.”](#)

I took the addresses that were collected in my survey and geocoded them using this website: <https://www.showmyip.com/>. Of the 271 addresses, 90 (33.2%) are geocoded as outside of Massachusetts. Looking at the ISPs for these out of state IP addresses, many are for large corporations, such as major consulting firms, banks, and insurance companies that have offices in Boston but are headquartered in other states. Some are also for large departments of the federal government that also have offices in Boston. In the experiment reported in the article, conducted in 2012, subjects were recruited on train platforms in the morning when most people were on their way to work and provided with a URL at which they could complete a survey on the platform Qualtrics, which also recorded their IP address. Many of these subjects likely completed the survey in the offices of one of these corporations. These large entities have their own ISPs, which is likely causing the geolocator to locate them at the ISP out of state. To verify that this phenomenon of mis-geocoded IP addresses is general, using the same website I looked at the IP addresses of respondents from another sample who I had surveyed using email addresses linked to a commercial voterfile. All of these respondents are registered to vote in Massachusetts. Out of 200, 41 (20.5%) are geocoded as outside of Massachusetts.

All of the non-personally identifying data for this project and my other research projects are publicly available. I am always willing to engage with other scholars about any comments, questions, or critiques about my research and welcome the opportunity to do so.

¹ <https://dataverse.harvard.edu/dataset.xhtml?persistentId=doi:10.7910/DVN/DOP4UB>

² <https://www.pnas.org/doi/10.1073/pnas.1317670111>