Response paper BRJ:

*Can you move to opportunity? Evidence from the Great Migration – Ellora Derenoncourt (2019)*

* Upward mobility is defined as children’s adult outcomes conditional on parent economic status
* The author focuses on the great migration between 1940 and 1970
  + Point here: are the results robust for these time periods (what if you take 1950-1970 (only post-war, for example?)
* One effect of the migration was white flight and other policies introduced in the north
* **The paper tests whether the Great Migration reduced northern cities’ability to facilitate black intergenerational progress. THIS IS THE MAIN RESEARCH QUESTION!!!**
* The identification strategy is as follows:
  + Comparing cummiting zones in the north that exogenously experienced larger or smaller increase in their urban black population.
  + The impact of these increases on average outcomes are estimated for individuals born in the 1980s
  + Finding: migration lowered upward mobility in the long run (what is long run?)
  + Black men were the most affected subgroup: widening the racial gap in these destinations
  + Using an alternative measure of mobility, it is shown that these effects most likely originate from childhood environment, not unobserved family characteristics (they are based on endogenous policy in the communities.
  + The analysis is based on 130 commuting zones, covering 85% of the non-southern US population in 2000
    - **POINT HERE: is there a reason why the other 15% is not covered! Or is it random. Not stated in the introduction at least. The locations did contain 97% of the non-southern black population (and about 50% of the total black puplation)**
  + There are specific linkages between villages/regions in the south and in north (communities moved together). To address omitted factors, a “shift-share” approach is used. This means that information is combined on the pre 1940 black southern migrants’ location choices with supply-side variation in county outmigration from 1940-1970, predicted from southern economic variables. This set is very large and hence LASSO is used to optimize the set of predictors of net-migration rates from the South.
    - CHECK THIS POINT CAREFULLY. IS THE LASSO BEING OVERUSED?
    - This step is critical in the whole process, since this constructs the shock to a commuting zo: the percentile of predicted black population increase
  + Main result: a 30-percentile greater increase in the black population, (approx. 1 standard deviation of the shock), lowered adult income rank of children from low income families by 3 percentiles (approx. 9% drop in adult income)
  + There is two main mechanisms that could underlie this effect: selection and location.
  + Use Chetty’s database of causal effects on children’ts adult outcomes and examine whether the causal effect of commuting zone varies with exogenous historical increase in black population.
  + No evidence of negative selection of families contributes to the association between historical racial composition shocks and eclines in upward mobility.
  + There are a lot of causal claims in the introduction. How are these backed up by quantitative analysis?
  + Could there still be some sort of selection effect? The black people that were already living in the north, might have had very different social status than the ones that moved there. So by aggregating black people overall, the mobilty might change, but looking at subgroups, it might not?

Regional changes from 1880 and 1940 are relatively constant

Data used to measure upward mobility, construction of the analysis sample of commuting zones and the mearsure of urban black population change.

Upward mobility of a location: average outcomes of children conditional on parent income or educational status. For the 1940 value, the 1940 census is used. Mobility is defined as fraction of 14 to 17 year olds in each commuting zone with 9 or more years of schooling from housholds were the householf had has between 5 and 8 years of schooling. Card et al take the maximum of the parents’ attainment, whereas the author takes the household head.

Income upward mobility for 1980s birth cohorts. Estimated mean individual or household income rank, conditional on parent household income rank. The rank is based on the national income distribution for indicivuals from the same birth cohort, measured at age 26.

* Maybe here can be a point of critique. Should we not compare people in the same stage of life? NO probably not…

How comparable are educational upward mobility in 1940 and income upward mobility in the 2000s? Correlation coefficient is 0.43. Moreover, income upward mobility is strongly correlated with high school graduation rates in low income families today (correlation of 0.53).

* This is definitely something to pay attention to…
* Also it is stated “The two measures are strongly correlated across US CZs where both are available”. What if they are not both available: is the data discarded?

The estimates for childhood exposure effects of commuting zones are deemed to be causal

Race specifi measures of upward mobility come from Chetty et al. They only look at the 25th and 75th percentile of the income distribution though!!!

The analysis is restricte to samples for which cities that are not missing population data in 1940. Moreover, all cities with less thatn 25,000 people are excluded.

* Maybe here something can be said about urbanization rates. Between 1940 and 1980 there is a significant urbanization going on from people moving from rural areas to urban areas. Hence there might be a bigger share of the population missing from the data in 1940 vs to 1980?

The key independent variable is the percentile of the increase. That is to deal with the right skewness of the data. Hence we look at the ranks of the increases, and not the increases themselves.

Six million black migrants who left the south during the south during the great migration, four million migrated between 1940 and 1970. Mid century economic conditions in the northern cities influenced where migrants moved and are thus likely correlated with increase in black population during this period.

Destination factors may influence both black population increases and future levels of upward mobility. Based on this, an instrument is constructed for the population increases that is plausibly exogenous with respect to such pre-1940 characteristics: the shift share instrument.

In the construction of the prediction of outflow, based on the previous characteristics, time series information is not taken into account. This could severely make the predictions worse.

Maybe the pre 1940 migration patterns are also the same as the post 1940 patterns? Hence would that solve the identification issue? Would it be a valid instrument? The overall objective is to find the casusal effect from 1940-1970 increases in migration inflow on the social mobility. For the instrument to be valid, it should affect the outcome variable only through the treatment, right?

The shares of the people going to certain places is kept constant. However, the question is if that is a valid assumption. Maybe this share hovers over time…

SO the way that they come up with these predicted migration flows is done in a very simplistic way. A simple linear prediction in terms of the 10 year lagged economic variables is done. They are simply the fitted values of these regressions. Why the lag of 10 years.

Moreover, the validity of the instrument is highly determined by the

**The main question is if the initial shares are exogenous or not!!!! THAT IS A STRONG ASSUMPTION. INITIAL SHARES CAN ALSO BE CONFOUNDED!!!**

**Would have been cool if the paper also considered to what extend the people moved in patterns. It cannot be that they all moved to the same place. Hence it cannot be entirely orthogonal to pull factors**

**Second order effects: black families might have different reproduction rates. Hence, there will be some endogeneity in the share of black population throughout the migration. A naturally increasing share of black people? Does that still take into account the migration, or is it a second order effect?**

**All variables that have gone up with time are correlated with the rising share of black people. This can be a confounder?**

**Why are only the 25th and 75th percentile looked at. Will results change when we look at all percentiles more elaborately?**

**Family sizes: even after conditioning on parent income,**

**Why is linear regression performed for the prediction of the shares? To increase correlation as instrument with the treatment, you can use different prediction methods (i.e. RandomForest, which also can deal with the high dimensionality of the problem). Lasso assumes that the final solution is sparse, but is that really an assumption that we want to make here?**

**Also make a bit of a summary to introduce the points**

Critical points:

Derenoncourt, E. (2022). Can you move to opportunity? Evidence from the Great Migration. *American Economic Review*, *112*(2), 369-408.

The author aims to investigate the causal link between