Documenting Adherence to/Departure from Pre-Registration

"The Effects of Moving to Opportunity on Youth Political Behavior"

Below is a list of all analysis choices described in the pre-analysis plan, followed by a note describing whether the current analysis follows this choice and, if not, why not.

Key Dependent Variables

1. Registration is measured as the proportion of MTO subjects in each random assignment group who correctly matched to the voter file.
    a. **Followed**.
2. To measure turnout, we examine general and primary elections from 2000 through 2018. In examining turnout, rather than looking at a single election, we average turnout over a number of elections in order to minimize the influence of idiosyncratic variation.
    a. **Followed**.
3. Voter turnout can be examined as proportion voting…among those registered [to vote].
    a. Followed.
4. **Added:** whether a subject ever voted in an election post-treatment.

Primary Analysis Methods

1. comparing the control to the treatment groups using randomization inference that clusters the subjects by MTO site and family
    a. Using OLS rather than randomization inference. The probability of assignment to treatment changed over time, and there is no way to back out all the information we would need to know each subject's probability of assignment to treatment. This makes randomization inference impossible, as we don't know the treatment assignment probabilities for each subject.
    b. Treat MTO sites as blocks, not clusters. This was a typo; analyses were clustered by family, as this was the level of random assignment; randomization sites are treated as blocks and addressed by including site fixed effects in analyses.
2. We will also include pre-treatment covariates available from L2 and the MTO baseline survey, including pre-treatment turnout (if applicable and available), race, and gender.
    a. **We present models with only MTO site controls and with the full suite of controls recommended in the documentation.** We did not control for pretreatment turnout since only 100 subjects ever voted before treatment. The full list of control variables is listed in the appendix.
3. We will also conduct our analysis with and without the weights to adjust for uneven probabilities of randomization
    a. **All analyses in the paper draft are weighted.** Because the weights reflect differences in the probability of assignment to treatment, estimates from models without the weights will be biased.
4. If compliance variables are available, we will measure a complier average causal effect (CACE); if not, we will measure an intent to treat effect (ITT).
    a. Both are reported.

5. [we will be] conducting subgroup analysis, with the sample divided between people over 18, 13 – 18, and under 13 at random assignment…
    a. **Followed.**
6. … and [as an alternative] interacting these age subgroups by treatment.
    a. **Not implemented.** We felt that presenting the models separately by age group was clearer than presenting single models with an interaction effect.

Secondary Analysis Methods

1. We will test the moderating effects of gender, race/ethnicity (for racial subgroups with sufficient power), duration of exposure/moving frequency (if available), and city at random assignment (if sufficiently powered).
    a. **Complied on gender, race/ethnicity, and city of random assignment.** There is no indicator for duration of exposure.
2. Additionally, we will conduct mediation analysis, using nonparametric procedures (Imai et al. 2010) to estimate the average causal mediation effect and average direct effect of each respective mediator: education, income, AFDC use, marital status, parental status, time in jail/prison, index of social connection, index of crime exposure, index of perceived racial discrimination, dissatisfaction with neighborhood, under-served by police, moving frequency and housing stability; also, a set of variables measuring economic distress (percent unemployed; percent TANF-receiving; percent female-headed household; aggregate income per working-age adult; percent under 18); level of racial segregation; and tract-level crime rates. These will be measured for the focal neighborhood, which is the first neighborhood after treatment, and if available, subsequent neighborhoods.
    a. **There is no direct effect of treatment** for adults and young children, nor is there a "first stage" effect of the treatment on education and most of these other variables, so we did not proceed with mediation analysis.
    b. We conducted mediation analyses for the teen sample on variables for which the treatment had a significant effect. However, post-treatment data on the teens is only available through the final roster dataset, so most of these variables are not available. Only the neighborhood-level variables (except crime rates, which we omit, and segregation, which we substitute with share nonwhite), education, parental status, marital status, moving frequency, and work status were available. We also tested whether the treatment affected these variables for children, though since there was no direct effect, we did not continue to mediation analyses.
3. We will apply the FDR correction to the set of mediation tests to adjust for multiple comparisons, as needed.
    a. **Not implemented;** no mediation tests were substantively significant.
4. We can also check our results for sensitivity to false positive and negative matches by restricting our data by various levels of the match probability….we will restrict our dataset to more or less restrictive match thresholds (0.85, 0.9, 0.95, 0.99) and test whether our primary results substantively differ with match criteria.
    a. **Replaced with models weighting by match probability.** Rather than restricting the data based on match quality, we present models in the appendix that include all possible matches and weight each match based on its posterior probability.

Measurement

1. Education will be coded as a set of indicator variables: high school completion (or GED), college attendance, and if sufficiently powered, college completion. We lack full data on college attendance. We will not have a measure of college completion for the teen or children's cohort. Finally, we will have a pre-treatment measure of whether adult household members were high school graduates or had a GED.
    a. **Pretreatment adults (followed):** indicators for having an HS diploma, GED, or being in school at baseline are available; these are included in the set of controls
    b. **Posttreatment adults (followed)**: for analyses in the appendix looking at participation by education, we use the categories of no HS diploma/GED, HS diploma/GED, AA degree, and BA degree or more
    c. **Posttreatment children (followed) :** in analyses of participation by education, some children can be categorized as no HS/GED, HS/GED, and more than HS/GED, if we limit analyses to those who were at least 18 at the time of the final survey; in mediation analyses, we use indicators for high school completion and college attendance (available for children and teens)
2. AFDC use: Available pre-treatment for all heads of households. It will also be used as a post-treatment mediator, if available, for the heads of households in the final household evaluation and the subset of the children's cohort that was administered the final youth evaluation.
    a. **Pretreatment: followed**
    b. **Posttreatment:** not available
3. Marital status: Available pre-treatment for all adults in the household, available for non-head members of the household from the interim household evaluation, and available for non-head members of the household from the final household evaluation. If available, we will also include a measure of martial status for the head of household from the final household evaluation.
    a. **Pretreatment:** followed
    b. **Interim:** not used, as we did not include non-household-head adults
    c. **Posttreatment:** available and used for mediation analyses with children/teens via a question asked of household heads about them; not used for adults, as we did not conduct mediation analyses
4. Parental status: Only indirectly and imprecisely available, and only for the subset of the children's cohort that was administered the final youth evaluation. It can be derived from a question asking who the respondent lives with, since one of the answer choices is their own children. If available, we will also use an item on the final household evaluation measuring the number of biological children each household member has, as well as a direct measure of parental status from the final youth evaluation.
    a. **Posttreatment:** available and used for mediation analyses with children/teens via a question asked of household heads about them
5. Time in jail or prison: Available for household members *but not head of household* from the final household evaluation in 2008. If available, we will also measure for how long and the last time household members were in jail or prison, both appearing on the final household evaluation.
    a. **Posttreatment:** available and used for mediation analyses with children/teens via a question asked of household heads about them; measures of duration and timing were not available
6. Index of social connection: Only non-neighborhood specific measures are available post-treatment, though neighborhood-specific measures are available pre-treatment. For the heads of

household, we measure how often they had friends over, how often they visited friends, and how often they go to church. For the subset of the children's cohort participating in the final youth evaluation, we measure how they met their friends, how often they visit friends they had at the time of random assignment, and how often they attend youth religious activities. For the heads of household, we also have pre-treatment measures of how often they talk to neighbors, how many friends and family members live in their neighborhood, and their likelihood of telling a neighbor if the neighbor's child had gotten into trouble.

   a. **Adults:** pretreatment indicator variables for some of these were used as baseline controls; see appendix for a full list. They were not combined into an index so as to comply with standard practice for control variables with the MTO data. The postreatment variables were not used because we did not conduct mediation analyses for adults.
   b. **Children:** not used, as we did not conduct mediation analyses for children.

7. Index of crime exposure: For the heads of household, we measure in 2008 whether they have recently seen drug use or sales, the safety of streets, and whether abandoned buildings, graffiti, and trash are a problem. For the subset of the children's cohort participating in the final youth evaluation, we measure whether they have recently seen drug use or sales and the safety of the streets. For the heads of household, we also have pre-treatment measures of street safety and, if available, whether they have been victims of crime in the recent past.
   a. **Adults:** pretreatment indicator variables for some of these were used as baseline controls; see appendix for a full list. They were not combined into an index so as to comply with standard practice for control variables with the MTO data. The postreatment variables were not used because we did not conduct mediation analyses for adults.
   b. **Children:** not used, as we did not conduct mediation analyses for children.
8. Dissatisfaction with neighborhood: For the subset of the children's cohort participating in the final youth evaluation and for the heads of household, we measure in 2008 satisfaction with their neighborhood. For the heads of household, we also have a pre-treatment measure of satisfaction with their neighborhood.
   a. **Adults:** the pretreatment indicator variable was used as a baseline control. The postreatment variables was not used because we did not conduct mediation analyses for adults.
   b. **Children:** not used, as we did not conduct mediation analyses for children.

The remaining variables have not yet been used in any analyses.

9. Income will be coded as raw annual earnings. Income data will come from two possible sources: income derived by L2 from commercial data (c. 2017) and income from the MTO data. The latter includes a measure of individual income from the respondent's main job, recorded by the MTO final evaluation surveys for the household and youth (2008). If available, we will supplement the adult measure from the household survey with an MTO measure of income from *all* of their employers in the previous year, measured in raw earnings or income categories depending on whether the respondent supplied the raw dollar amount. If available, we will also use a measure of total household income from the MTO final evaluation (measured in raw earnings). None of the teen cohort will have the 2008 MTO income measure. If available, we will also calculate a pre-treatment measure of income for the head of household from their earnings per hour, their hours worked per week, and the number of months worked in the preceding year.
   a. **Income data from L2 is not in the dataset, nor is income data from MTO.**

10. Index of perceived racial discrimination: For the heads of household, we measure in 2008 whether they have faced discrimination in the past 6 months in their own neighborhood and with police. For the subset of the children's cohort participating in the final youth evaluation, we measure whether they have faced discrimination from the police in the past 6 months.
    a. Not used, as we did not conduct mediation analyses for adults or children.
11. Under-served by police: Available for the heads of household from the 2008 final household evaluation.
    a. Not used, as we did not conduct mediation analyses for adults.
12. Moving frequency and housing stability: If available, we will use the following measures from the 2002 interim household survey to construct measures of moving frequency and housing stability: length of time the head of household has been in their current housing space; whether the head of household was without housing in the past year; the amount of time it has been since the head of household last rented or owned their own unit; the length of time that the head of household has lived in their current neighborhood; and the number of times the head of household moved since the year of random assignment.
    a. Not used, as we did not conduct mediation analyses for adults.

Other decisions

1. Sample inclusion: The approximate numbers of individual MTO participants are: 4,616 adults, 2,331 teens (13 – 18 at random assignment), and 8,945 children (under age 13 at random assignment).
    a. **Followed.** "Other adults" are not included in analyses.