Pre-Registration for "The Causal Effects of Moving to Opportunity on Voter Registration and Turnout"
Using the "AsPredicted" Template on the OSF

1. **Have any data been collected for this study already?**
   - **Yes, at least some data have been collected for this study already**
   - **No, no data have been collected for this study yet**

   Yes. All data for this study have already been collected, but the research team does not have access to the data at the time of pre-registration. The research team will be licensing data about the Department of Housing and Urban Development's Moving to Opportunity (MTO) program from the National Bureau for Economics Research. The MTO data will be matched at the individual level with voter registration records. The matching process has not yet begun at the time of pre-registration.

2. **What's the main question being asked or hypothesis being tested in this study?**

   This paper uses the Moving to Opportunity (MTO) housing experiment to evaluate the causal impact on political participation of being assigned to — and for some, complying and living in — less-distressed neighborhoods. Our primary outcome of interest is registration and voter turnout. We examine outcomes for participants who were adults, teens, and children at the time of random assignment.

   We hypothesize that when children are assigned to a low-poverty voucher condition, their political participation will be higher later in life than those who are assigned to a no-voucher condition. Those assigned to a standard voucher condition (unrestricted) will exhibit weaker effects. We expect those assigned as adults or teens to be much less positively affected, likely exhibiting null or negative effects. This is consistent with studies of childhood anti-poverty interventions on adult voting (Holbein 2017), and with prior studies exploring the effect of MTO on other outcomes, such as income, that are immediate predictors of voting (Chetty, Hendren, and Katz 2016).

   We will also test the effect of treatment on party registration (if available). We hypothesize that children and adults who are assigned to the low-poverty voucher to be more likely to be registered with a political party. We will also examine the difference between treatment and control groups in proportion registered as Republican. We do not have a strong directional hypothesis with respect to which party individuals are more likely to join, but it may be that moving to a low-poverty neighborhood will crystalize political interest and engagement.

3. **Describe the key dependent variable(s) specifying how they will be measured.**

   Our primary outcomes of interest are registration and voter turnout. We will also test the effect of treatment on party registration. We examine outcomes for adults, teens, and children at the time of random assignment.

Registration is measured as the proportion of MTO subjects in each random assignment group who correctly matched to the voter file. This estimand is unbiased if we assume that our ability to match is equal across treatment groups and attrition via incapacitation is equal across groups. However, it is possible that a subject who was registered to vote at some point after random assignment was subsequently removed from the voter file (death, name change, imprisonment, subsequent moves, etc.), but this cannot be observed. To that end, we define the primary estimand as the percent registered among all possible voters in each group and thus incorporate attrition into this estimand. However, with this approach, bias can still result from differences in ability to match the voter file with MTO data, but this cannot be observed.

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. Because we know the date when the subject registered to vote, subject age, and the date of assignment in the MTO experiment, we can measure participation before 2000 as a pre-treatment covariate (if applicable and available), post-MTO participation, and post-registration date participation.

Voter turnout can be examined as proportion voting among all subjects or only among those registered. Both approaches are subject to potential bias from differential non-matching across treatment and control. We will evaluate turnout both ways, for robustness. We plan to test for balance in pre-treatment covariates among missing and non-missing subjects, and if missingness is balanced between treatment and control, this will provide some assurance that missingness will not cause bias. However, such balance does not account for unobserved covariates that may be related to both missingness and potential outcomes; thus false negative and false positive are both possible even after accounting for covariate balance.

We will measure party registration as the proportion of MTO subjects who are (1) registered with a political party (vs. nonpartisan) and (2) who register as Republican in each treatment group. If available, in states with non-partisan registration, we will examine partisanship imputed by L2 or we will examine which party primary in which individual voted, for the more recent primary on record (if any), as a proxy for partisanship.

4. **How many and which conditions will participants be assigned to? (optional)**

Households were assigned to one of three groups: 1) receive a traditional Section 8 housing voucher (traditional voucher condition); 2) receive a housing voucher that could only be used in a Census tract with less than 10% poverty (low-poverty voucher condition); or 3) receive no voucher (with the ability to remain in public housing) (control condition).

5. **Specify exactly which analyses you will conduct to examine the main question/hypothesis.**

All treatment effects can be measured by comparing the control to the treatment groups using randomization inference that clusters the subjects by MTO site and family. 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. We will also conduct our analysis with and without the weights to adjust for uneven probabilities of randomization (Orr et al. 2003). If compliance variables indicating whether the housing voucher was used are available, we will measure a complier average causal effect (CACE); if not, we will measure an intent to treat effect (ITT).

Each analysis will be conducted separately by age at the time of random assignment. As noted above, a person who was younger at the time of random assignment, especially in years prior to adulthood, is likely to be more affected by the treatment. As such, we will test for different effects by age in three ways: 1) conducting subgroup analysis, with the sample divided between people over 18, 13 – 18, and under 13 at random assignment; and 2) interacting these age subgroups by treatment.

6. **Any secondary analyses?**

We will test the moderating effects of race/ethnicity (for racial subgroups with sufficient power) and city at random assignment (if sufficiently powered).

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. The matching procedure (see below) will return a posterior probability that two entries are a match, as well as a false discovery rate and false negative rate at different probability cutoffs for the fuzzy matching algorithm. These cutoffs correspond to lower bounds for the posterior probability at which a match will be accepted as true. Thus, 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.

7. **How many observations will be collected or what will determine the sample size? No need to justify decision, but be precise about exactly how the number will be determined.**

The MTO initiative randomized approximately 4,600 poor families. 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).

8. **Anything else you would like to pre-register? (e.g., data exclusions, variables collected for exploratory purposes, unusual analyses planned?)**
The matching procedure is as follows:

To perform the match, we use the fuzzy-matching algorithm (see the *fastLink* package in R; Enamorado, Fifield, and Imai 2018), which returns multiple matches with a probabilistic estimate for each pairwise match of being correct. We will exact match on gender, such that all female MTO participants will be matched to individuals in the voter file who are listed as "female" or "unknown," and all male MTO participants will be matched to individuals in the voter file who are listed as "male" or "unknown". To reduce the number of pairwise comparisons, we will then derive clusters of individuals in the voter file and the MTO dataset with maximally similar first names (see Enamorado, Fifield, and Imai 2018). We then derive partial matches based on first name, last name, middle name, suffix, and birthdate. The probabilistic estimate returned by the algorithm will allows us to restrict our final dataset to different thresholds to check the sensitivity of our results to potential false positives.