Addendum to

Pre-analysis Plan for

The Effect of Citizenship on Financial Outcomes

Vasil Yasenov[a,b], Michael Hotard[a], David D. Laitin[a,c], Duncan Lawrence[a], Jens Hainmueller[a,c,d]

[a]Immigration Policy Lab, Stanford University.
[b]IZA Institute of Labor Economics .
[c]Department of Political Science, Stanford University.
[d]Graduate School of Business, Stanford University.

February 9, 2021

1 Introduction

We previously registered plan AEARCTR-0006790, an analysis for an experiment that examines the impacts of access to citizenship on the financial outcomes of immigrants. This analysis was based on a sample of lottery registrants who were matched to credit records. Here we pre-register an addendum to also examine the impacts of access to citizenship on a broader set of economic and labor market outcomes that were measured using follow-up surveys.

2 Empirical Methodology

The methodology for this addendum will generally remain the same as with the previous plan. The main differences will be as follows:

- Outcomes - we will use a set of new outcomes that we detail below.
- Sample - the estimation samples will differ somewhat because the survey outcomes are only available for those who responded to the surveys. In addition, we can include registrants that responded to the survey but were not matched to the credit data.
- Covariates - we will use a richer set of pre-treatment covariates from the registration data following Hainmueller et al. (2018).
  - Gender, age, education (high school, some college, college dummies), years since green card, country of origin dummies (Dominican Republic, Colombia and Ecuador), language (English, Spanish dummies), marital status (single, married dummies) and (log) household income.

  - Additionally, we will control for the number of days between the date of the registration and the date of the voucher lottery as well as the pre-treatment outcomes measured at baseline (when available).

- Regression models - we are restricted in the regression models that we can run because most outcomes are not measured every year. Hence, we will focus on regressions with 2-year outcomes and use the panel models only for the limited set of outcomes for which we have baseline and follow-up data.

- Subgroups - we will estimate the model on subgroups based on splits of the pre-treatment covariates that we have.

  - Male and female,
  - Spanish and English speakers,
  - Below and above median age,
  - Low (High School only) and high education level,
  - Below and above median household income.

3 Sample and Follow-up Surveys

The sample for this analysis will consist of the lottery registrants from all three registration cohorts (2016, 2017, and 2018). These registrants were randomly assigned (within blocks) to receive a naturalization voucher. The analyses will draw on two data sources. First, the baseline data measures individual characteristics at the time of registration for the lottery. Second, follow-up surveys provide measures of citizenship and outcomes of interest.

Table 1 below provides the timeline of the experiment and follow-up surveys. Because the citizenship process can take up to a year, we focus on outcomes measured in years two and later. The table shows that we have two year outcomes for all three cohorts, which will be the main focus of the analysis.

Table 1: Timeline of the Experiment Registration and Follow-up Surveys

Cohort	Registration	Survey Year 1	Survey Year 2	Survey Year 3	Survey Year 4
2016	Sept 2016	Nov 2017	July 2018	Sept 2019	Nov 2020
2017	Aug 2017	–	Sept 2019	Nov 2020	–
2018	July 2018	Sept 2019	Nov 2020	–	–

4 Outcomes

The outcome variables from the follow-up surveys are divided into four categories as follows:

- Income - we will use log transformations of:

  - Personal income,

  - Household income,
  - Equivalised household income.

- Labor force status - we will analyze indicators for being:
  - Employed,
  - Unemployed,
  - In the labor force,
  - In school.

  The first two variables are defined only for individuals in the labor force.

- Expenses - we will examine indicators that the respondent can afford an unexpected expense of:
  - $500,
  - $1,000,
  - $10,000.

- Occupation status - we will match each respondent's self-declared occupation to the 2018 American Community Survey (Ruggles et al., 2020) using the Standard Occupation Classification (SOC) system (*occsoc* variable) and examine the following outcomes:
  - Average occupation wage,
  - Average occupation education level,
  - Occupation wage rank,
  - Occupation education level rank,
  - Occupational task content from O*NET (DOL/Employment and Administration, 2018):
    * manual/routine/abstract/cognitive content.

References

DOL/Employment and Training Administration, "O*NET Database, Version 23.0," *dataset]. www.onetcenter.org/db$_r$eleases.html.*, 2018.

Hainmueller, Jens, Duncan Lawrence, Justin Gest, Michael Hotard, Rey Koslowski, and David D Laitin, "A randomized controlled design reveals barriers to citizenship for low-income immigrants," *Proceedings of the National Academy of Sciences*, 2018, *115* (5), 939–944.

Ruggles, Steven, Sarah Flood, Ronald Goeken, Josiah Grover, Erin Meyer, Jose Pacas, and Matthew Sobek, "Integrated Public Use Microdata Series: Version 10.0. Minneapolis: University of Minnesota," 2020.