# Online Appendix for Campaign Finance Vouchers do not Expand the Diversity of Donors: Evidence from Seattle

For online publication only

## Contents


**1 Data Related** **1**
- 1.1 Elections used to construct Voting Index . . . 1
- 1.2 Data merging . . . 1
- 1.3 L2's demographic predictions against known data . . . 2
  - 1.3.1 Income against administrative data . . . 2
  - 1.3.2 Imputed race against hand-coded race . . . 2
  - 1.3.3 L2 Partisan estimates . . . 3
- 1.4 Candidate-Level Data (in-text Figure 3) . . . 3
  - 1.4.1 Data sources . . . 3
  - 1.4.2 Merging procedure . . . 3
  - 1.4.3 Candidate diversity over time . . . 4
- 1.5 Match Rates . . . 4
  - 1.5.1 Getting a true match rate . . . 5
- 1.6 Does matching bias results? Cash vs. Vouchers users . . . 5
- 1.7 Does matching bias results? Ethnorace . . . 5
- 1.8 Is census data accurate under the selection (panel) process? . . . 6

**2 Methods and Models** **7**
- 2.1 Block bootstrapped standard errors for Table 4 . . . 7
- 2.2 Parallel trends for block group quartiles . . . 7
- 2.3 Relaxing the linearity assumption . . . 8

**3 Dollar-weighted participation** **9**

**4 Mechanisms for persisting inequality** **10**
- 4.1 Income inequality in participation . . . 10
  - 4.1.1 Substitution . . . 10
  - 4.1.2 Time trends . . . 11
- 4.2 Are program rules inhibiting more diverse participation? . . . 12
  - 4.2.1 Excluding cash donors after candidates hit their voucher limits . . . 12
  - 4.2.2 What if late voucher contributions were counted? . . . 13
- 4.3 Evidence for learning? . . . 13
- 4.4 Are nonwhite potential donors more participatory in the presence of nonwhite candidates? 14

**5 Alternate comparisons and identification issues** **16**
- 5.1 Voucher-to-cash comparisons . . . 16
- 5.2 Two-year comparisons . . . 16
  - 5.2.1 Effects for people outside the panel using a 2-year comparison . . . 17

**6 Does panel start-year matter?** **19**

**7 Advances in the DiD method and placebo tests** **21**
- 7.1 Difference-in-Differences with a Continuous Treatment: Callaway et al., 2021 . . . 21
- 7.2 De Chaisemartin and d'Haultfoeuille papers . . . 21

7.3 Doubly robust difference-in-differences estimators Sant'Anna and Zhao, 2020 . . . . . 22
7.4 Main tables with placebo treatment time . . . . . 22
7.5 Two year comparisons with placebo treatment times . . . . . 22

# 1 Data Related

## 1.1 Elections used to construct Voting Index

Congressional primaries, 2008-2020 as well as general elections in even-years are used to construct the voting index. These elections are among those with the highest levels of turnout in Washington. Presidential primaries are omitted because presidential primaries often had no meaningful Democratic competition and many presidential primaries were caucuses.

## 1.2 Data merging

For legal and anonymity reasons, original L2 and WA voter files cannot be included in replication files. See the "README" file for further details.

**Voter file to Local campaign contributions (name and ZIP):** The only link is names (single string from SEEC and first/last/middle in separate fields within the voter file) and zip code. Zip codes are exact matched. An iterative procedure is used to match based on names. Those not matched using single-string full names (with or without middle initials) had their names split up according to the number of spaces within the name string and matched alongside other names with the same number of strings. All procedures involve matching both a possible first name and possible last name to the voter file. The following table shows all procedures used. Only one-to-one matches are considered: if a single SEEC line yields multiple records in the voter file, the match is rejected. Table A1 describes the order that I use to make matches. Unlisted types of procedures, such as [ab,c] are omitted as they yielded no matches.

| Spaces | [a,b] | [a,b,c] | [a,b,c,d] |
|---|---|---|---|
| **Proc 1** | [a,b] | [a,c] | [a,d] |
| **Proc 2** | | [a,b] | [ab,d] |
| **Proc 3** | | [ab,c] | [ab,cd] |
| **Proc 4** | | [a,bc] | [a, bcd] |

**Table A1** – An example of a match procedure for the hypothetical MARY KATE OLSON, who donated cash as MARY KATE OLSON but is registered to vote as [MARY, K, OLSON]. First, "MARY KATE OLSON" == "MARY K OLSON" and "MARY OLSON" are tried and both fail. Next, the string "MARY KATE OLSON" from the SEEC data is deconcatenated by white space into [MARY, KATE, OLSON], and I try to match the first name-last name pair [MARY, OLSON] from the voter file. The first procedure then tries to match [a,c] = [MARY, OLSON] to [MARY, OLSON], which is indeed a match. Assuming that this is the only MARY OLSON within the zip code, this match succeeds and MARY KATE OLSON is removed from the pool of names to continue onto subsequent matching procedures. Procedures such as [a,bcd] indicate concatenating the non-comma-separated letters, for example, for procedure [a,bcd], "SERENA VAN DER WOODSEN" in the SEEC data (three spaces) gets searched as [SERENA, VANDERWOODSEN] and procedure [ab,d] is searched as [SERENAVAN, WOODSEN].

**FEC Data to Voter File:** FEC data is used to make Figure 2 in the main text. Nicknames, misspellings, and inconsistent addresses are common. FEC data was downloaded from here, consisting of donations only from Seattle, to authorized committees of House, Senate, and presidential candidates from 2005-2020. These were matched to the voter file using full name and zip code.

**Census Data to Voter File:** Based on an individual's December 2021 address, I use ArcGIS's geocoding service. I use coordinates and census shapefiles (2010 and 2020 census linse) to get block groups. Block group data from the ACS/Census is then merged on by block group.

**L2 to Voter File:** L2 and voter files are merged using unique WA Voter ID numbers.

## 1.3 L2's demographic predictions against known data

Estimates of individual data from vendors is often described as a "black box", where it's unclear how values are calculated and whether these values are accurate. I validate these measures by comparing data from L2 to their true values.

### 1.3.1 Income against administrative data

To validate L2 income, I obtained F-1 Personal Financial Affairs Statements via a public records request for all candidates who filed them 2017-2019. I find candidates within the L2 donor file, and compare true reported incomes to estimated incomes. Binned L2 income estimates are very accurate.

**Table A2** – Binned income esimates are generally close to true binned income estimates. On the right, SEEC dollar codes from an F-1 Personal Fiancial Affairs Statement.

| | *Dependent variable:* |
|---|---|
| | L2 est income (binned) |
| Self reported income bin | 1.029*** |
| | (0.026) |
| Observations | 51 |
| $R^2$ | 0.969 |
| *Note:* | *p<0.1; **p<0.05; ***p<0.01 |

| SEEC DOLLAR CODE | AMOUNT | | |
|---|---|---|---|
| (1) | $0 | -- | $999 |
| (2) | $1,000 | -- | $4,999 |
| (3) | $5,000 | -- | $9,999 |
| (4) | $10,000 | -- | $24,999 |
| (5) | $25,000 | -- | $99,999 |
| (6) | $100,000 | -- | $199,999 |
| (7) | $200,000 | -- | $999,999 |
| (8) | $1,000,000 | -- | $4,999,999 |
| (9) | $5,000,000 or more | | |

In this regression, an intercept of 0 is forced. This model shows that L2 estimated incomes are indeed accurate representations of an individual's true income, with an $R^2$ value of 0.97. Error is largely driven by four candidates.[1] A limitation in this regression is that incomes are binned, leaving less room to confidently run analysis with the raw values of L2 estimated incomes. For this reason, in my analysis using incomes, I bin L2-estimated incomes into deciles. Block group demographic data from the census, however, confirms similar results to those found using modelled L2 data.

### 1.3.2 Imputed race against hand-coded race

**Table A3** – Confusion matrix of candidate racial classification

| | *Asian* | *Black* | *Hispanic* | *White* |
|---|---|---|---|---|
| **Asian** | 4 | 0 | 0 | 1 |
| **Black** | 0 | 2 | 0 | 4 |
| **Hispanic** | 0 | 0 | 11 | 1 |
| **White** | 3 | 0 | 0 | 59 |

This section presents a confusion matrix, comparing L2-estimated race against hand-coded race for city council and city attorney candidates 2005-2019. Hand coding involved searching the internet for websites and photos of candidates. True values are on the rows (bold) and modeled values are in the columns (italics). Not all individuals are categorized by L2. Only 57% of imputed Asians are Asian (however this error is largely due to two candidates in interracial marriages). The L2 algorithm

[1] Among these are an individual who clearly hastily filled out the F-1 financial affairs disclosure, and another who reports a household income of less than $5000 but a net worth of $900,000.

also struggles with correctly classifying Blacks. While everyone it classifies as Black is indeed Black, it misses 67% of individuals who are Black. Significant error comes from misclassifying Blacks as white. L2 fares better with Hispanics, correctly categorizing 92% as Hispanic.

### 1.3.3 L2 Partisan estimates

Washington does not record party information. This subsection describes how party information is extracted by L2. I also provide a check that L2 party data does capture partisanship. L2 states: *"For the most part, the Washington State parties have run their presidential nominating processes by caucus rather than by presidential primary. However, during the presidential cycles of 1992, 1996, 2000, 2008 and 2016, the state also ran presidential primaries. In each of these presidential primaries, the state temporarily recorded the party ballot that was picked up by each voter...L2 acquired all of it after each primary when it was still available and I believe we are the only source for this information."*

To verify that L2 party is calculated as anticipated, I merge L2 partisan data with individual CF Scores from Bonica's DIME (2008-2018). Positive CF scores suggest that an individual has primarily donated to Republicans (likely themselves a Republican), while negative CF scores usually belong to Democrats. As shown to the right, individual CF scores and party match on quite well. Individuals represented in this chart are from the 2011-2021 panel, and L2-identified non-partisans are excluded. In the main analysis I binarize *Republican*, as non-partisans skew towards negative CF scores and are more similar to Democrats.

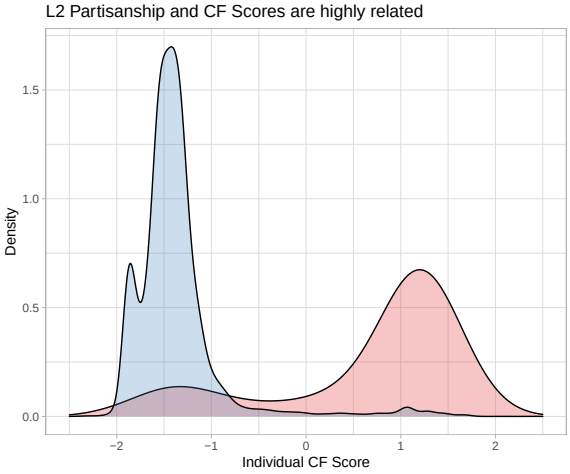



**Figure A1** – CF scores match onto partisanship quite well.

## 1.4 Candidate-Level Data (in-text Figure 3)

Contingent upon qualifying for the Democracy Voucher program, candidate participation in the Democracy Voucher is high. Those choosing to not participate are often well-known or well-off.

### 1.4.1 Data sources

The voter file was merged to a list of candidates from 2005-2021. Candidates include only those who stayed within races until at least the primary election, and names and election-level information are obtained from Seattle's campaigns archive. I match and obtain voter file (2005-2021) data for over 95% of candidates (nearly 200). Candidates who resided in Seattle in years that I have an L2 file (2017-present) also have estimates of income and race. For candidates who left Seattle prior to 2017, I used online information to gather race data. As a final check, I hand-corrected race data for candidates, as L2 (and other name-based race prediction algorithms) often mischaracterizes white and black individuals and women in interracial marriages.

### 1.4.2 Merging procedure

Because there are few candidates and thereby little room for error in calculating candidate demographics, candidates are hand-matched to records from the voter file. In ambiguous cases, candidate information from news articles, websites, or other publicly accessible data are used to match candidates to the most likely individual in the voter file.

### 1.4.3 Candidate diversity over time

There are two common, related questions on candidate diversity. First, did vouchers make candidates more diverse? Second, does the presence of diverse candidates improve inequality within the donor pool? I investigate the former here, and the latter in Section 4.4.

While the causal mechanics of the first question are beyond the scope of this paper, I make two points here with descriptive data. First, candidate diversity has been somewhat increasing over time. Second, by the mid-2010s, there isn't a substantial difference in between those who ran prior to voucher rollout and those who collected vouchers as part of their candidacy.

| | Primary candidates | | | General candidates | | |
|---|---|---|---|---|---|---|
| | non-voucher | voucher cand | se | non-voucher | voucher cand | se |
| Income | 169408.2 | 143785.98 | 11677.07 | 187637.9 | 148967.39 | 16397.13 |
| Prop female | 0.32 | 0.46 | 0.08 | 0.39 | 0.52 | 0.13 |
| Prop nonwhite | 0.23 | 0.31 | 0.07 | 0.25 | 0.38 | 0.12 |
| Age | 49.64 | 46.58 | 2.01 | 52.82 | 47.05 | 2.63 |

**Table A4** – **Candidate differences, starting 2005** – Base n=179. All candidates have race/gender data. Prim cands: Party n=149, Age n=172, Inc n=131. Gen elec base n=87, age n=83, party=70, inc=67.

| | Primary candidates | | | General candidates | | |
|---|---|---|---|---|---|---|
| | non-voucher | voucher cand | se | non-voucher | voucher cand | se |
| Income | 164347.77 | 143785.98 | 13168.99 | 194877.05 | 148967.39 | 18988.97 |
| Prop female | 0.4 | 0.46 | 0.09 | 0.57 | 0.52 | 0.15 |
| Prop nonwhite | 0.26 | 0.31 | 0.08 | 0.3 | 0.38 | 0.15 |
| Age | 47.02 | 46.58 | 2.34 | 48.67 | 47.05 | 3.03 |

**Table A5** – **Candidate differences, starting 2015** – Base n=116. All candidates have race/gender data. Primary candidates: Age n=115, Income n=93. General elec base n=43, age n=42, income=37.

In the 2005-2021 table, there are mild (but generally not statistically significant) differences in between voucher and non-voucher candidates. These differences lessen with the 2015-2021 data. As shown through the columns on female and nonwhite candidate candidates (especially in general elections), varying the start date actually closes a lot of the difference in between non-voucher and voucher accepting candidates. The demographics of Seattle are dynamic, and it appears that through time, viable candidates have already been becoming more diverse along race and gender dimensions. The differences in individual candidate wealth is however noteable, though the mean income of a voucher-accepting candidate is still well-above the mean income of a Seattle household.

## 1.5 Match Rates

In Table 1 of the main text, I describe my data sources and merges. Here I describe loss from merges, and whether this loss can significantly alter results from the main text.

**Voter file onto Census/ACS data:** To merge voter file and census data, I first geocoded each individual's address using the ArcGIS Geocoder, and obtained a latitude and longitude for each address. For about 200 addresses, the ArcGIS Geocoder was unable to find a location, and through additional geocoding means, whittled the un-geocoded addresses to fewer than 10. Using King County census tract shapefiles, I extract census tract from an address' latitude and longitude. Beyond the fewer than 10 addresses unable to be geocoded, this procedure should encounter no data loss.

**Voter file to L2:** Secretary of State voter files and L2 are perfectly linked using state voter ID. There should be no data loss here.

**Voter file to campaign finance data:** In contrast to the two prior merges, data loss is inevitable in this merge because of a lack of a perfect key in between the two datasets. For the rest of this section, I detail attempts to recover a match rate, and assess consequences and possible bias in the main result (I focus on results from in-text Table 4) as a result of SEEC-Voter file matches.

**Voter file alone:** I construct the Seattle panel by extracting the city component of an individual's residential postal address, and requiring the individual to be assigned (by L2's measure) to a Seattle City Council district in 2021. This can cause two types of issues. First, individuals may register to vote with an incorrect address. Second, individuals might live in a residence in unincorporated King County with "Seattle" postal addresses prior to 2021 (thereby previously being ineligible to participate in earlier years), and residing in a true Seattle City Council district by 2021. Neither should pose significant issues: incorrect registration is likely random, and I believe that only White Center population 16.6k), Boulevard Park (5.3k), and Bryn-Mawr Skyway (17.4k) uses Seattle as part of addresses without being able to participate in Seattle city elections, and residing in one of these CDPs and then moving into Seattle proper should also be a comparatively rare event. Additionally, both White Center and Bryn-Mawr have lower average incomes and more racial minorities relative to Seattle, so if their previous residents moved into Seattle and used Democracy Vouchers, differences would be biased towards 0.

### 1.5.1 Getting a true match rate

Getting a match rate in between the donor file and the Seattle panel is not necessarily possible because we cannot be sure that a lack of match is because of a) the matching algorithm fails to capture true matches or b) the individual is not within the Seattle panel. Because the Seattle panel is a subset of the population, most matches will fail because individuals fail to join the panel. A workaround is to extract a single-year match rate. Assume that all 2021 donors appear in the December 2021 voter file: this will slightly deflate the match rate, as some donors may move away, die, or become incarcerated, resulting in removal from the 2021 December file. Additionally, non-citizen residents may request and use Democracy Vouchers, but again, would not show up in the voter file. Using the same procedure used to merge data in the main analysis, we can see what percentage of 2021 donors who state that they live in Seattle can be identified within the December 2021 voter file. Using this procedure, I find that around 83% of donations from 2021 are matched with an individual in the voter file. 73% of cash donations are matched, and 87% of voucher donations are matched. This discrepancy is because cash donations report name and location via user-filled forms, while Democracy Voucher reporting uses voter-file supplied name and zip codes.

## 1.6 Does matching bias results? Cash vs. Vouchers users

Would results change if we randomly deleted voucher donors matches to meet the lower match rate of cash donors? In Table A6 I replicate the results of cols 1-4 of Table 4 in the main text, while randomly simulating fewer and fewer voucher matches (changing in between 0-18% of voucher matches to non-participation). The reported numbers are coefficients and their t-values.

## 1.7 Does matching bias results? Ethnorace

Are members of different ethnoracial groups getting matched at the same rate? Using wru (Imai & Khanna, 2016), I use last names from the donor file to impute race (the geography used here is King County, for which Seattle is the central city). Match rates are indeed higher for whites relative to non-whites.

As in the section on the cash-voucher match discrepancy, I simulate results by varying the level of matches for whites, shown in the right side. There is a 6% difference in between white and Asian

**Table A6** – Changing the balance of cash-voucher matches doesn't change the direction nor significance of main results."Best guess" donation-type merging-bias corrected estimates are in bold. Prop kept = proportion of voucher donations not deleted within the simulation.

| Prop | **White** | | **Income** | | **Voting freq.** | | **Republican** | |
|---|---|---|---|---|---|---|---|---|
| Kept | coef. | t-value | coef. | t-value | coef. | t-value | coef. | t-value |
| 1 | 2.525 | 25.951 | 0.207 | 14.647 | 0.948 | 111.617 | -3.253 | -28.717 |
| 0.98 | 2.502 | 25.865 | 0.206 | 14.604 | 0.929 | 109.939 | -3.187 | -28.254 |
| 0.96 | 2.434 | 25.293 | 0.204 | 14.579 | 0.914 | 108.696 | -3.106 | -27.65 |
| 0.94 | 2.401 | 25.088 | 0.204 | 14.596 | 0.898 | 107.248 | -3.092 | -27.76 |
| 0.92 | 2.388 | 25.074 | 0.201 | 14.513 | 0.879 | 105.59 | -3.016 | -27.182 |
| 0.9 | 2.362 | 24.986 | 0.198 | 14.346 | 0.867 | 104.512 | -2.939 | -26.555 |
| 0.88 | 2.339 | 24.871 | 0.2 | 14.556 | 0.84 | 102.062 | -2.893 | -26.314 |
| 0.86 | 2.296 | 24.563 | 0.194 | 14.195 | 0.825 | 100.806 | -2.818 | -25.799 |
| **0.84** | **2.19** | **23.481** | **0.189** | **13.919** | **0.811** | **99.553** | **-2.8** | **-25.825** |
| 0.82 | 2.176 | 23.577 | 0.19 | 14.027 | 0.793 | 97.792 | -2.679 | -24.768 |

**Table A7** – The left half compares match rates by wru-predicted ethnorace. On the right, "best guess" race-based merging-bias corrected estimates is in bold.

| Wru-predicted race | Proportion matched |
|---|---|
| Asian | 0.804 |
| Black | 0.829 |
| Hispanic | 0.811 |
| white | 0.853 |

| Prop white retained | t-val | Coef |
|---|---|---|
| 1 | 25.951 | 2.525 |
| 0.980 | 24.482 | 2.378 |
| 0.960 | 22.866 | 2.217 |
| **0.940** | **21.326** | **2.064** |
| 0.920 | 20.076 | 1.939 |
| 0.900 | 18.729 | 1.806 |
| 0.880 | 16.891 | 1.625 |

match rates (85.3/80.4), the largest white-nonwhite match rate difference. Reducing the amount of white matches by 6% gives a "race-matching bias corrected estimate," (in bold) and this estimate does not substantively change results.

## 1.8 Is census data accurate under the selection (panel) process?

| | *Dependent variable:* |
|---|---|
| | Likelihood of donating |
| VPP*In-sample prop. white | 0.049 |
| | (0.008) |
| Observations | 5,104 |

| | *Dependent variable:* |
|---|---|
| | Likelihood of donating |
| VPP*In-sample mean income | 0.026 |
| | (0.004) |
| Observations | 5,101 |

**Table A8** – The use of in-sample block group ethnorace and income (L2) instead of census-based values does not substantially change results from the main text. Cluster robust standard errors. Income in thousands.

Because of the selection process to enter the panels entails being continuously registered to vote for the past 10+ years, individuals within the panels are likely to be wealthier and whiter than the average individual within their neighborhoods. This might lead to incorrect results, especially if there is strong heterogeneity in who is dropping out of the sample within block groups. While the use of census data indeed overestimates average income and underestimates the proportion of nonwhite residents, results are substantively and numerically similar to results from Table 3 in the main text when swapping census block group income with within-panel mean income.

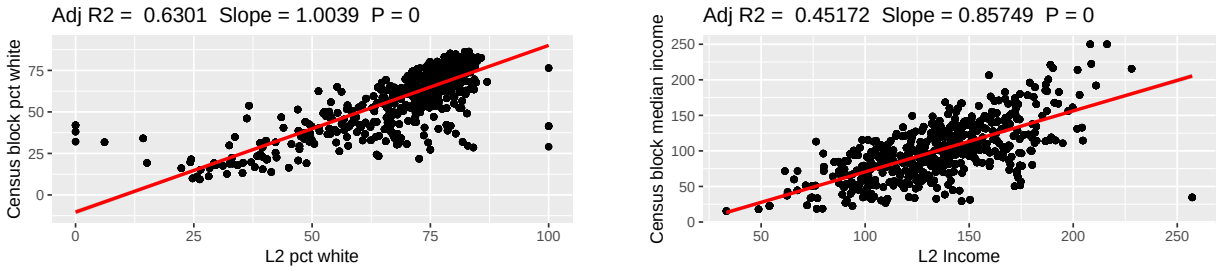



**Figure A2** – In-sample L2 income (right) and percent white (left) are somewhat larger than census-provided demographics.

# 2 Methods and Models

## 2.1 Block bootstrapped standard errors for Table 4

It's difficult to cluster standard errors in Table 4 because of its size. Using the block bootstrap (sampling based on the individual, instead of on the observation), I run 100 iterations of the four main individual-level regression and obtain $\beta$ means and their standard errors.

**Table A9** – Block bootstrapping does not have a large impact on substantive results or standard errors.

| | *Dependent variable:* | | | | |
|---|---|---|---|---|---|
| | Pr. donating | | | | |
| | (1) | (2) | (3) | (4) | (5) |
| White*VPP | 2.54 | | | | 0.44 |
| | (0.12) | | | | (0.13) |
| Income decile*VPP | | 0.21 | | | 0.18 |
| | | (0.02) | | | (0.02) |
| Voting index*VPP | | | 0.95 | | 0.94 |
| | | | (0.01) | | (0.01) |
| Republican*VPP | | | | −3.27 | −3.31 |
| | | | | (0.14) | (0.14) |

## 2.2 Parallel trends for block group quartiles

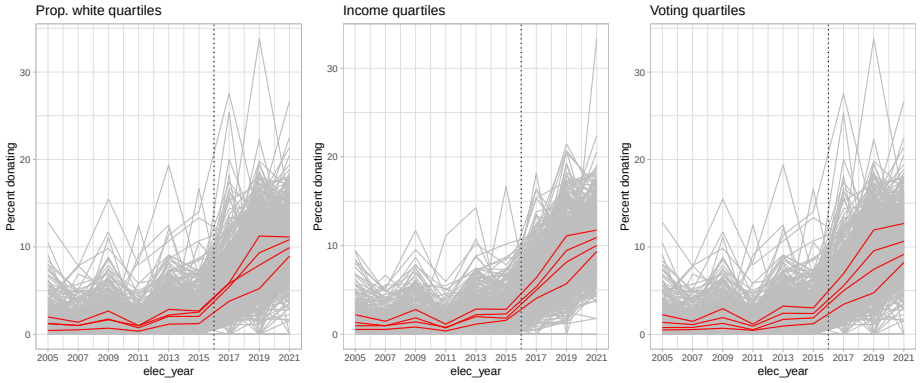



**Figure A3** – Light gray lines are individual census blocks. Red lines are quartile averages. While not perfectly parallel, the quartiles all move in tandem. Block group's donation behavior is different for income quartiles as they're constructed with 2019 ACS block groups (2010 lines), while others are constructed with 2020 census block groups.

## 2.3 Relaxing the linearity assumption

**Table A10** – In-text Table 3 using quartiles. All models include block group and cycle fixed effects. Cluster Robust standard errors.

| | % Donating | | | |
|---|---|---|---|---|
| | (1) | (2) | (3) | (4) |
| VPP*votes (2nd quartile) | 1.34 | | | |
| | (0.33) | | | |
| VPP*votes (3rd quartile) | 2.25 | | | |
| | (0.32) | | | |
| VPP*votes (4th quartile) | 3.50 | | | |
| | (0.31) | | | |
| VPP*% white (2nd quartile) | | 1.09 | | |
| | | (0.36) | | |
| VPP*% white (3rd quartile) | | 1.86 | | |
| | | (0.34) | | |
| VPP*% white (4th quartile) | | 2.19 | | |
| | | (0.31) | | |
| VPP*med income (2nd quartile) | | | 0.95 | |
| | | | (0.34) | |
| VPP*med income (3rd quartile) | | | 1.53 | |
| | | | (0.35) | |
| VPP*med income (4th quartile) | | | 2.12 | |
| | | | (0.32) | |
| VPP*Republican (2nd quartile) | | | | −0.17 |
| | | | | (0.36) |
| VPP*Republican (3rd quartile) | | | | −1.30 |
| | | | | (0.35) |
| VPP*Republican (4th quartile) | | | | −1.85 |
| | | | | (0.35) |
| Observations | 5,104 | 5,104 | 4,464 | 5,104 |

**Table A11** – Treating income decile and voting history as a factor instead of assuming linearity confirms results from the main text. Cycle and individual fixed effects and robust standard errors.

| | *Dependent variable:* |
|---|---|
| | Likelihood of donating (%) |
| Vtg idx =1*VPP | 0.42 (0.14) |
| Vtg idx =2*VPP | 0.75 (0.15) |
| Vtg idx =3*VPP | 1.20 (0.15) |
| Vtg idx =4*VPP | 1.23 (0.14) |
| Vtg idx =5*VPP | 1.65 (0.14) |
| Vtg idx =6*VPP | 1.98 (0.13) |
| Vtg idx =7*VPP | 2.24 (0.13) |
| Vtg idx =8*VPP | 2.82 (0.13) |
| Vtg idx =9*VPP | 3.51 (0.14) |
| Vtg idx =10*VPP | 4.40 (0.14) |
| Vtg idx =11*VPP | 5.31 (0.15) |
| Vtg idx =12*VPP | 6.40 (0.15) |
| Vtg idx =13*VPP | 7.30 (0.16) |
| Vtg idx =14*VPP | 8.75 (0.16) |
| Vtg idx =15*VPP | 10.86 (0.16) |
| Vtg idx =16*VPP | 13.72 (0.15) |
| Observations | 1,297,683 |

| | *Dependent variable:* |
|---|---|
| | Pr. donating (%) |
| Inc dec 2*VPP | 0.82 (0.17) |
| Inc dec 3*VPP | 0.85 (0.17) |
| Inc dec 4*VPP | 0.94 (0.17) |
| Inc dec 5*VPP | 1.50 (0.17) |
| Inc dec 6*VPP | 1.43 (0.17) |
| Inc dec 7*VPP | 1.68 (0.17) |
| Inc dec 8*VPP | 1.69 (0.17) |
| Inc dec 9*VPP | 2.33 (0.19) |
| Inc dec 10*VPP | 1.96 (0.18) |
| Observations | 1,297,683 |

# 3 Dollar-weighted participation

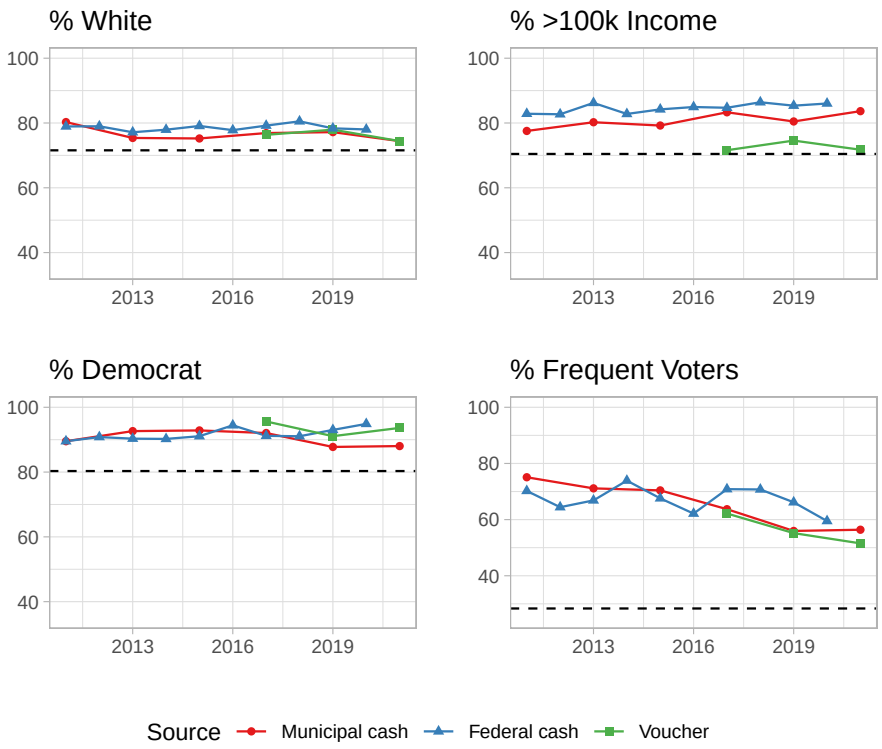



**Figure A4** – Using dollar-weighted contributions does not drastically change results. Except in the case of incomes, dollar-weighted contributions also reveal that already-overrepresented groups utilize Democracy Vouchers similarly or even more than they utilized cash. The dashed line represents the mean among those within the panel.

**Table A12 – Individual participation** – Using amount contributed as a dependent variable, individuals in overrepresented groups are still benefitting the most from the Democracy Voucher program. Cycle and Individual fixed effects. Robust standard errors. Many coefficients are close to their Table 4 (main text) analogue. This likely has to do with the fact that the median donation is $100.

| | Mean Amount Donated ($) | | | | | | |
|---|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) |
| VPP*White | 2.53 | | | | 0.25 | | |
| | (0.28) | | | | (0.29) | | |
| VPP*Inc Decile | | 0.54 | | | 0.51 | | |
| | | (0.04) | | | (0.05) | | |
| VPP*Vot Idx | | | 0.93 | | 0.90 | | |
| | | | (0.03) | | (0.03) | | |
| VPP*Repub | | | | −2.35 | −2.34 | | |
| | | | | (0.43) | (0.47) | | |
| VPP*Age (<70) | | | | | | 0.02 | |
| | | | | | | (0.01) | |
| VPP*Age (≥70) | | | | | | | -0.26 |
| | | | | | | | (0.07) |
| Observations | 1181601 | 1297683 | 1319634 | 1319634 | 1161927 | 1052541 | 168975 |
| Starting Year | 2005 | 2005 | 2005 | 2005 | 2005 | 2011 | 2011 |

**Table A13 – Census block groups** – Results if Table 3 used dollar values as the dependent variable. Block and cycle fixed effects. Cluster robust standard errors. Income in thousands.

| Mean Donation ($) | (1) | (2) | (3) | (4) |
|---|---|---|---|---|
| VPP*vote index | 1.11 | | | |
| | (0.15) | | | |
| VPP*pct white | | 0.06 | | |
| | | (0.012) | | |
| VPP*med income | | | 0.04 | |
| | | | (0.005) | |
| VPP*Republican pct | | | | −0.19 |
| | | | | (.06) |
| Observations | 5,104 | 5,104 | 5,104 | 4,464 |

# 4 Mechanisms for persisting inequality

## 4.1 Income inequality in participation

Main paper results indicate that when the voucher program is in place, the biggest increases in participation come from wealthier individuals. At the same time, Figure 2 (main text) shows that voucher users are more likely than cash donors to have lower income levels. With respect to income, voucher users are more representative than cash donors. Why?

### 4.1.1 Substitution

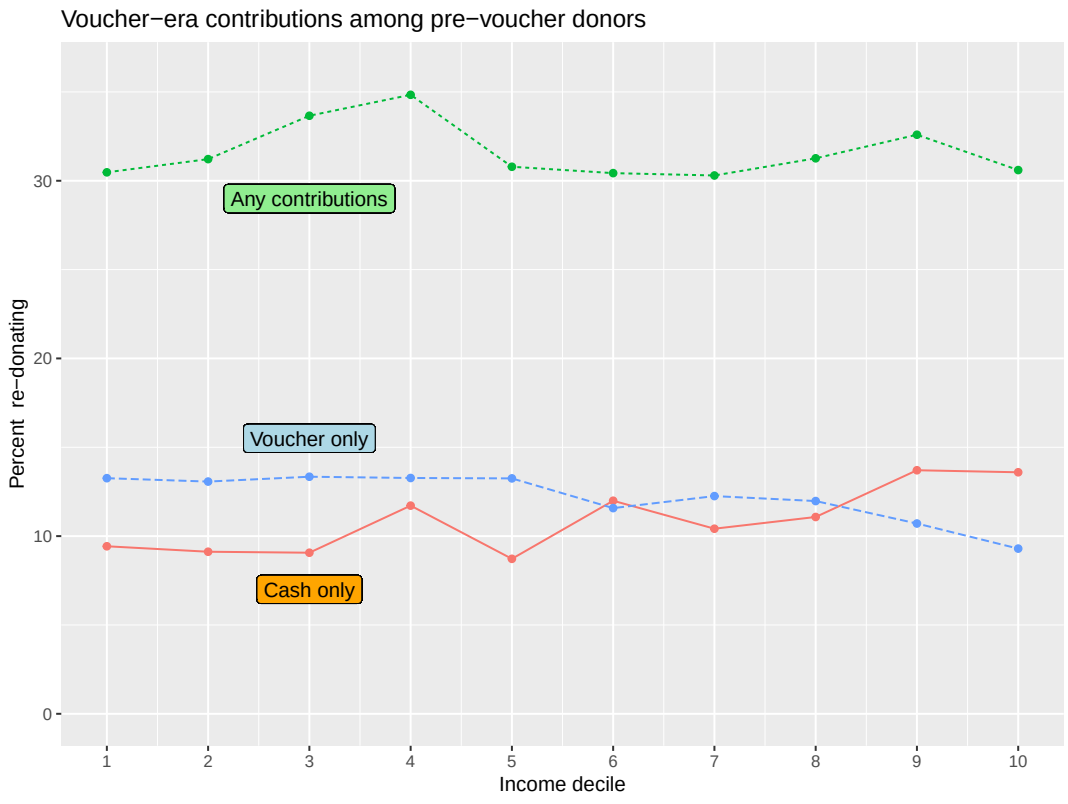



**Figure A5** – Wealthier donors are more likely to re-donate using only cash, and lower-income donors are more likely to re-donate using only vouchers.

Substitution may explain why voucher users have lower mean incomes. Consider the population of 2015 donors from the 2005-2021 panel, who are among the most likely to donate in 2017. For both the highest and lowest income deciles, 2017 non-participation rates are relatively identical at 61.7% and 58.7%, respectively. However, among these 2015-2017 donors, voucher usage is higher at the lowest decile, at 58.44%, relative to 36.8% for the highest. That is, high-propensity donors similarly participate in voucher years, but high-propensity high-income donors contribute cash while high-propensity low-income donors are more likely to use only vouchers. This general pattern manifests when looking at pre-voucher contributors (all years) and their voucher-era donation behavior in Figure A5.

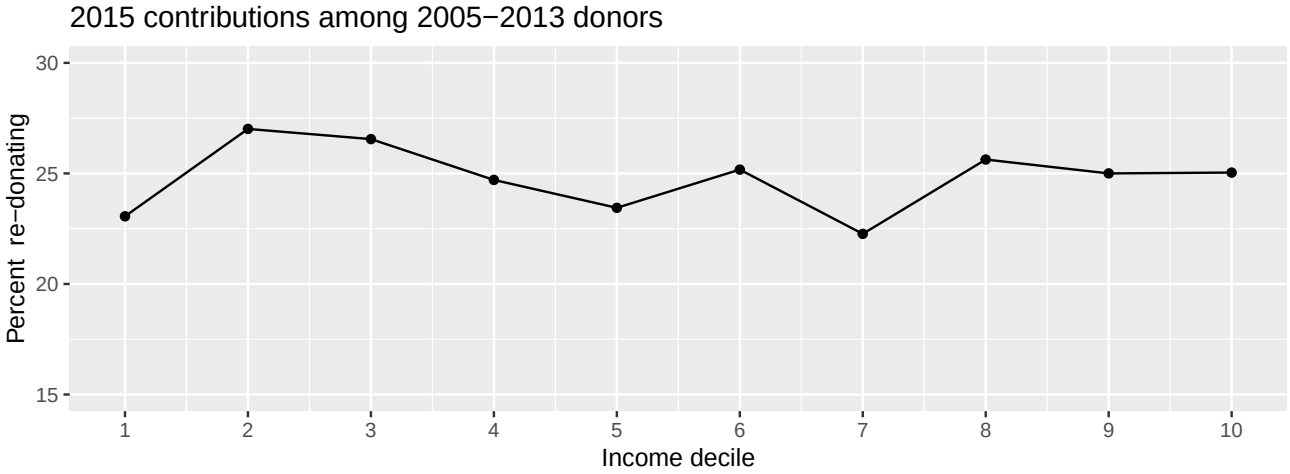



**Figure A6** – Wealthier donors are more likely to re-donate using only cash, and lower-income donors are more likely to re-donate using only vouchers.

One might be concerned that vouchers are the reason that lower-income donors are re-donating, however, this does not necessarily seem to be the case. Among pre-voucher donors from the 2005-2021 panel, somewhere around 30-35% make a voucher-era contribution. This number is relatively consistent across income deciles, and most notably, is virtually similar among the richest and poorest deciles. Examining the 2015 behavior of 2005-2013 donors in Figure A6, while there is some variation by decile, there is no strong monotonic trend in between income decile and propensity to re-donate.

Substitution is further suggested by the top right chart in Figure 2 in the main text. That is, while Tables 3 and 4 suggest that when the voucher program is in place, new donors are wealthier, Figure 2 shows that voucher users are much closer to the median income than are municipal cash donors. This is consistent with the idea that less-wealthy pre-voucher cash donors may be switching to exclusively using vouchers when they become available, substituting what would have been a costly cash donation with a free voucher donation.

### 4.1.2 Time trends

There remains the possibility that in voucher years, the donor pool becomes more diverse despite higher campaign finance participation by overrepresented groups when the voucher program is in place. That is, the donor pool can still trend in more equal directions due to less-disproportionate levels of voucher-year participation. Especially in the case of income, Figure 2 (main text) suggests that this could be the case. Table A14 examines mean donor incomes in block groups in pre- and post-voucher years. In all four regressions, the presence of Democracy Vouchers, either alone or interacted with time, are statistically insignificant. The passing of time, not voucher availability, seems to be most associated with an increasingly economically diverse donor pool.

**Table A14** – Controlling for pre-voucher time trends, block group participatory income inequality remains steady in when the voucher program is in place

| | *Dependent variable (within block group):* | | | |
|---|---|---|---|---|
| | Income Decile | | Mean Donor Income | |
| | (1) | (2) | (3) | (4) |
| voucher_yr | 0.045 | −0.098 | 1,142.633 | −2,998.269 |
| | (0.086) | (0.152) | (2,017.325) | (3,552.300) |
| year | −0.028 | −0.044 | −722.081 | −1,193.405 |
| | (0.013) | (0.019) | (302.515) | (449.729) |
| voucher_yr:year | | 0.030 | | 860.498 |
| | | (0.026) | | (607.667) |
| Observations | 2,687 | 2,687 | 2,687 | 2,687 |

Each line in the data is a census block group-cycle 2011-2021. All models include block group fixed effects. VPP=Voucher program in place. Robust standard errors.

## 4.2 Are program rules inhibiting more diverse participation?

There are bounds on how individuals can utilize Democracy Vouchers. Before an individual can allocate a Democracy Voucher to a candidate, the candidate must elect to participate in the Democracy Voucher program, and then must reach a threshold of small cash donations and signatures. This itself is likely not a big barrier to usage: these thresholds are relatively low for viable candidates. The larger threat is on the other end, that candidates are subject to fundraising caps and can't actually take in an unlimited amount of vouchers. The Democracy Voucher program operates on a first-come-first-serve basis, and in the last two elections (2019 and 2021), all but one general election voucher-accepting candidate reached the point of becoming unable to accept more Democracy Vouchers.

### 4.2.1 Excluding cash donors after candidates hit their voucher limits

In order to evaluate whether program idiosyncrasies are to blame for the non-diversification of the donor pool, I ask whether effects disappear (or substantively change) without the post-threshold period, where candidates can no longer accept vouchers. In Table A15 I replicate the first five columns of in-text Table 4, dropping post-threshold donations.

**Table A15** – After-voucher-threshold donors are not driving main results. Individual and cycle fixed effects. Robust standard errors

| | *Dependent variable:* | | | | |
|---|---|---|---|---|---|
| | Likelihood of donating | | | | |
| | (1) | (2) | (3) | (4) | (5) |
| White*VPP | 2.49 | | | | 0.45 |
| | (0.10) | | | | (0.10) |
| Income decile*VPP | | 0.19 | | | 0.16 |
| | | (0.01) | | | (0.01) |
| Voting index*VPP | | | 0.93 | | 0.92 |
| | | | (0.01) | | (0.009) |
| Republican*VPP | | | | −3.23 | −3.31 |
| | | | | (0.11) | (0.12) |
| Observations | 1,181,601 | 1,297,683 | 1,319,634 | 1,319,634 | 1,161,927 |

### 4.2.2 What if late voucher contributions were counted?

I consider what would've happened if all vouchers could be redeemed. Democracy Voucher program data includes records of vouchers which were attempted to be donated, but where a donation did not materialize (officially coded as "accepted" or "received" instead of "redeemed"). An obvious issue here is that individuals might opt to not even attempt to donate their vouchers if they knew that their preferred candidate had hit the spending limit (thereby being unable to accept more vouchers), but I this is unlikely. The dates in which candidates hit their spending limits were not well-advertised by the city nor by media, and the proportion of vouchers received after these dates is relatively in line with the proportion of cash received after these dates.

By comparing successful versus unsuccessful attempts at democracy voucher redemption in 2021, I can gather a sense of whether a fully-funded voucher program would yield more diverse donors. In Table A16, people whose vouchers are non-redeemed and permanently stuck at "accepted" or "received" are more white, richer, and older, though less Democratic than those who successfully assign their vouchers. In Table A17, the inclusion of these received-but-not-redeemed vouchers only very slightly changes coefficients from the main text.

**Table A16** – Demographic Balance of successful and unsuccessful voucher users

| | Successful | Unsuccessful | CI low | CI high |
|---|---|---|---|---|
| Prop. white | 0.78 | 0.81 | -0.04 | -0.03 |
| Income decile | 5.44 | 5.76 | -0.35 | -0.28 |
| Prop. Democrat | 0.98 | 0.94 | 0.04 | 0.04 |
| Age | 47.25 | 53.78 | -6.74 | -6.32 |

**Table A17** – Main effects if received-but-not-redeemed vouchers counted. Robust SE.

| | *Dependent variable:* | | | | |
|---|---|---|---|---|---|
| | Likelihood of donating | | | | |
| | (1) | (2) | (3) | (4) | (5) |
| iswhiteperson:VPP | 2.54 | | | | 0.45 |
| | (0.10) | | | | (0.10) |
| income_perc:VPP | | 0.21 | | | 0.18 |
| | | (0.01) | | | (0.01) |
| sums:VPP | | | 0.95 | | 0.94 |
| | | | (0.01) | | (0.009) |
| VPP:republican | | | | −3.26 | −3.34 |
| | | | | (0.11) | (0.12) |
| Observations | 1,181,601 | 1,297,683 | 1,319,634 | 1,319,634 | 1,161,927 |

## 4.3 Evidence for learning?

The results from the main text show that overrepresented group participation grew at a higher rate than it did for underrepresented groups in voucher years. Was this concentrated mainly when vouchers were first introduced? Did underrepresented groups "catch up" to overrepresented participation in later years of the Democracy Voucher program? We might expect this because information might be slower to disseminate into underrepresented communities, but also because in 2021, mayoral elections (the highest-salience local office) became voucher-eligible. Using a two-year design (causal issues expounded on in Section 5.2), I descriptively examine whether there appears to be evidence for heterogeneous learning over time.

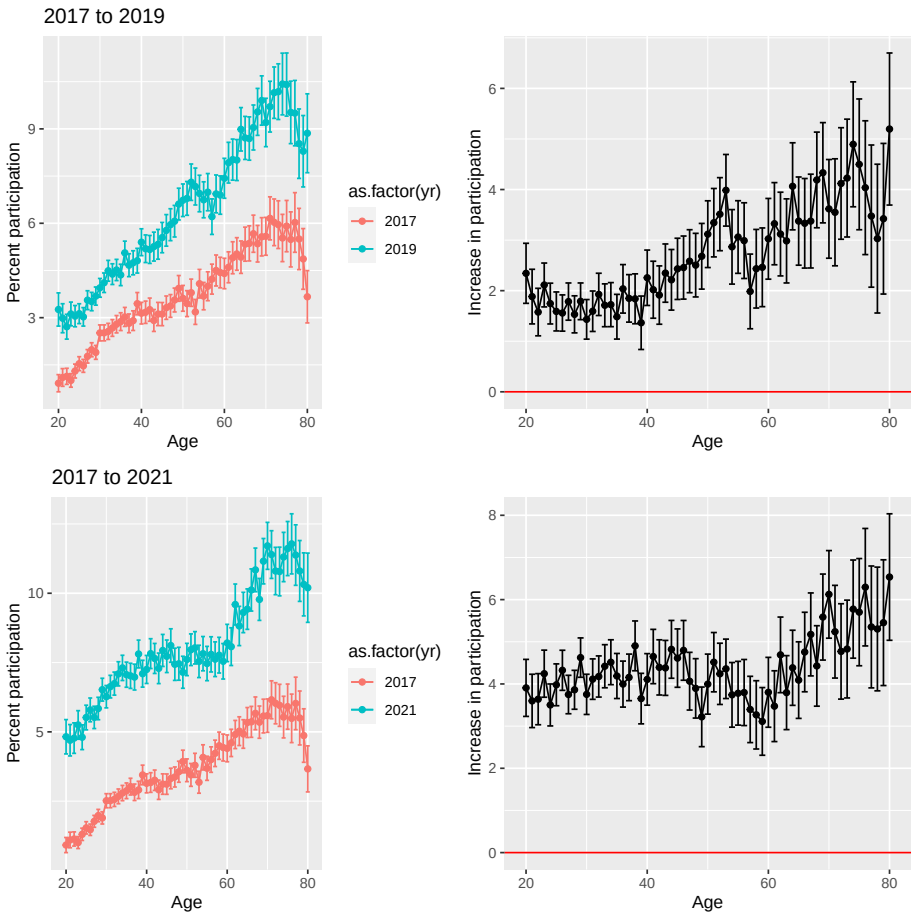



**Figure A7** – Younger individuals do not "catch up" in campaign finance participation in between 2017-2019 and 2017-2021. Plotted values are average rates of participation by age.

In this Figures A7-A8, I examine evidence for learning using the full Seattle population. This limits the independent variables that I can use to age and wru-computed race, as I do not have L2 measures for non-contemporary populations. 2019 and 2021 age gaps are mostly on par with the 2017 gaps, and race gaps widen in 2019, though the black-white gap in 2021 is about equivalent the 2017 gap.

## 4.4 Are nonwhite potential donors more participatory in the presence of nonwhite candidates?

This section evaluates how VPP*White changes when potential donors have the opportunity to contribute to an in-district nonwhite candidate. The key comparison is VPP*white and VPP*nonwhite when there's a nonwhite candidate running in a given district. In column 1 (column 2 shows similar results), I find that the difference in VPP premium when there's a nonwhite candidate present is -.23 - (-1.37) = 1.14 in favor of whites. The VPP premium for whites decreases when there is a nonwhite candidate running in an individual's district (relative to the 1.73 premium when candidates are white), but it is still significantly positive. This evidence suggests that the presence of nonwhite candidates only somewhat reduces, not solves, the VPP premium for whites.

There are caveats with this modelling strategy and data. These results should not be interpreted as the causal effect of a nonwhite candidate or individual race, but rather as the causal effect of VPP for individuals who are [race] living in districts with [candidate race]. Elections are not that frequent,

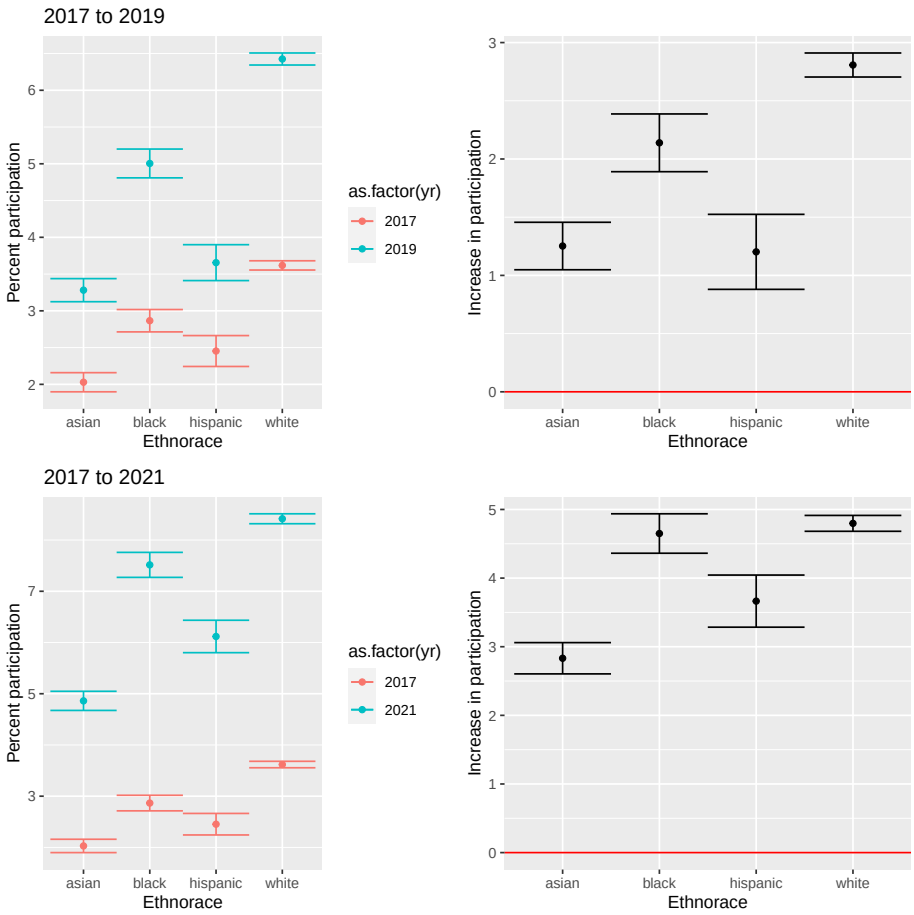



**Figure A8** – Nonwhite individuals do uniformly not "catch up" in campaign finance participation in between 2017-2019 and 2017-2021, though the gap does somewhat decrease. Plotted values are average rates of participation by race.

**Table A18** – White individuals donate more when VPP, even in the presence of nonwhite candidates. Baseline= VPP*White*Nonwhite. Robust standard errors.

| | *Dependent variable:* | |
|---|---|---|
| | Likelihood of donating | |
| | (1) | (2) |
| VPP*White cand*Nonwhite indiv | −1.73 | −1.74 |
| | (0.12) | (0.12) |
| VPP*Nonwhite cand*White indiv | −0.23 | −0.27 |
| | (0.16) | (0.15) |
| VPP*Nonwhite cand *Nonwhite indiv | −1.37 | −1.52 |
| | (0.18) | (0.17) |
| Observations | 1,173,008 | 1,319,634 |
| Years | 2005-2019 | 2005-2021 |

so ultimately this is a low-n analysis. Only 2019 provides any by-district variation of candidate race, and in 2021 and 2017 there were no city council district elections. Finally, individuals are free to donate to candidates outside of their own district.

# 5 Alternate comparisons and identification issues

In previous analyses on Democracy Vouchers, there are two common ways to assess the changing representativeness of donors. I examine the causal assumptions needed for each of these to be valid.

## 5.1 Voucher-to-cash comparisons

Knowing that year-to-year comparisons (next subsection) are problematic, analysts may compare same-year cash and voucher pools. Figure 2 in the main text shows this data. This comparison is causally problematic as it hinges on the assumption that cash donors in voucher years would still donate cash in non-voucher years (which is mildly plausible as the financial outlay is similar, although it is worth noting that the amount of cash donors in voucher years also skyrocketed compared to previous years), and less plausibly, that voucher users would not donate cash if vouchers were not available. As discussed in Section 4.1.1, there is evidence that individuals, especially those with lower economic means, likely substitute vouchers for a cash donation when vouchers are available. In addition to the no-substitution assumption, this strategy further assumes that cash donors are not brought in because of the availability of vouchers. It is possible that a certain type of person gains more interest in local politics because of the vouchers, and not only allocates their vouchers to candidates, but also makes an additional cash contribution. In each the substitution and additional-interest cases, the composition of the cash pool is causally impacted by vouchers. As a result, the cash pool is not an adequate counterfactual for a voucher-free world.

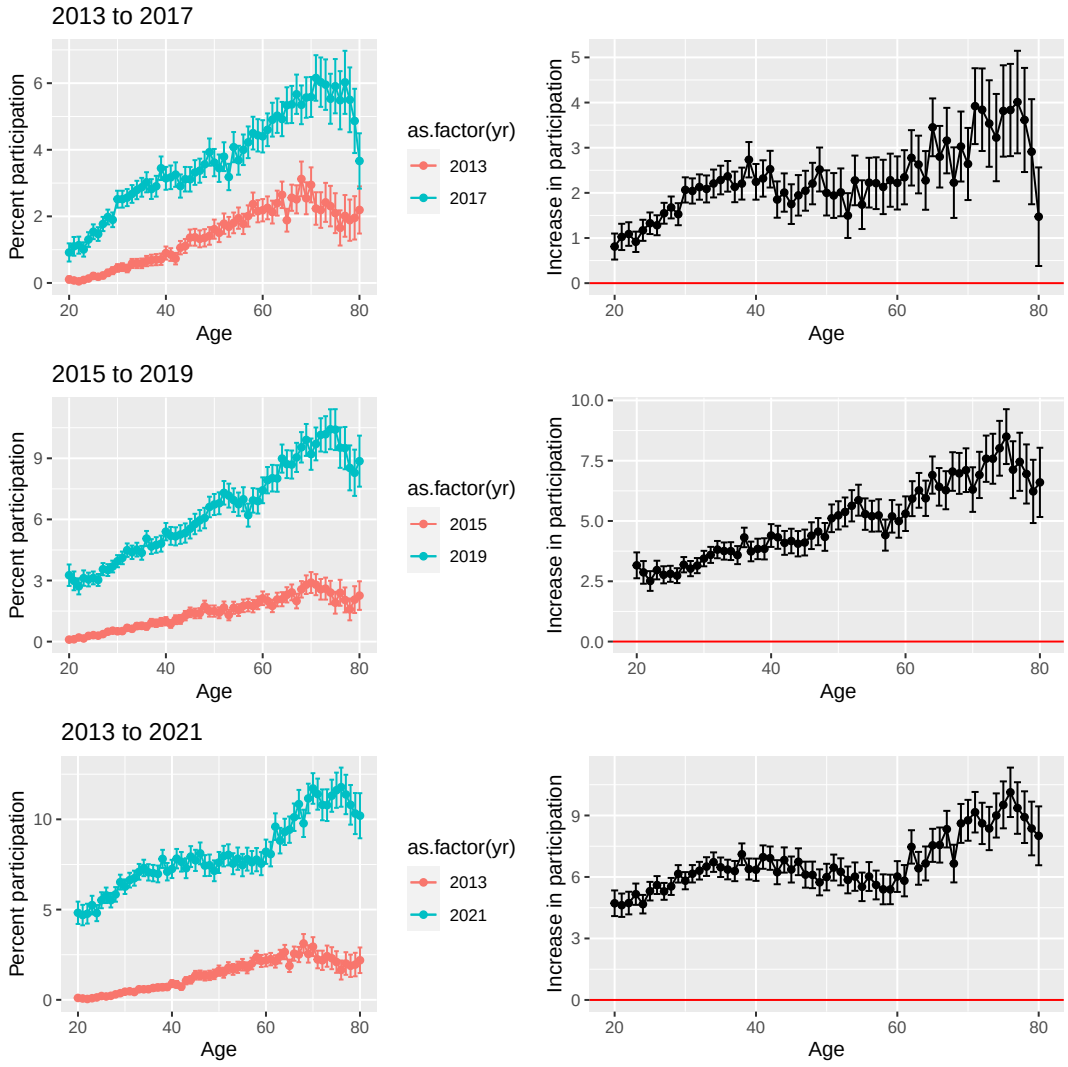
## 5.2 Two-year comparisons

Another way that an analyst could investigate the question is to compare donor pools using two year comparisons of full populations, instead of creating a multi-year panel that ends up excluding movers. I briefly argue why my method leads to more credible causal results, but also do show that the an alternate 2-year differences-in-differences design does not actually substantively change results from the main text.

### Issues with the design

While the design in the main text is restricted in that movers (who are disproportionately young and less white) within the period of study are excluded from the analysis, this is also what enables use of the differences-in-differences model with time and individual fixed effects. Without individual fixed effects, apparent changed donor pool diversity might be as a result of the city's changing demographics over the same time. Seattle is an evolving city: historically highly segregated along racial lines, its nonwhite population is growing. At the same time, rents have risen, and much of this population and economic growth is driven by tech jobs,[2] an industry known to be disproportionately Asian and disproportionately not Hispanic and Black. Thus arises a possible problem: if newcomers into the city are inherently different than those they are replacing in ways that are related to political participation (and it is known that race, age, income, and their intersections are highly related to political participation), and donor pool diversification happens concurrently with the voucher rollout, it becomes impossible to disentangle whether diversification is because of the new population, the vouchers, or a combination of the two. Additionally, this design is highly specific to individual years. In an effort to somewhat control some idiosyncrasy, year-pairings in the following analyses are based on the presence or absence of mayoral elections.

**Figure A9 – Age:** For each pre-post voucher comparison, more donor growth come from near-retirement-age individuals relative to individuals in their 20s and 30s. Plotted values are average rates of participation by age.




### Results using the two-year design

That being said, the results within this section substantively mirror results in the main text. In this section, using age (from voter files) and race (using wru to impute race based on last name), I find that the directionality and significance of results in the main text are preserved even under the 2-year specifications.

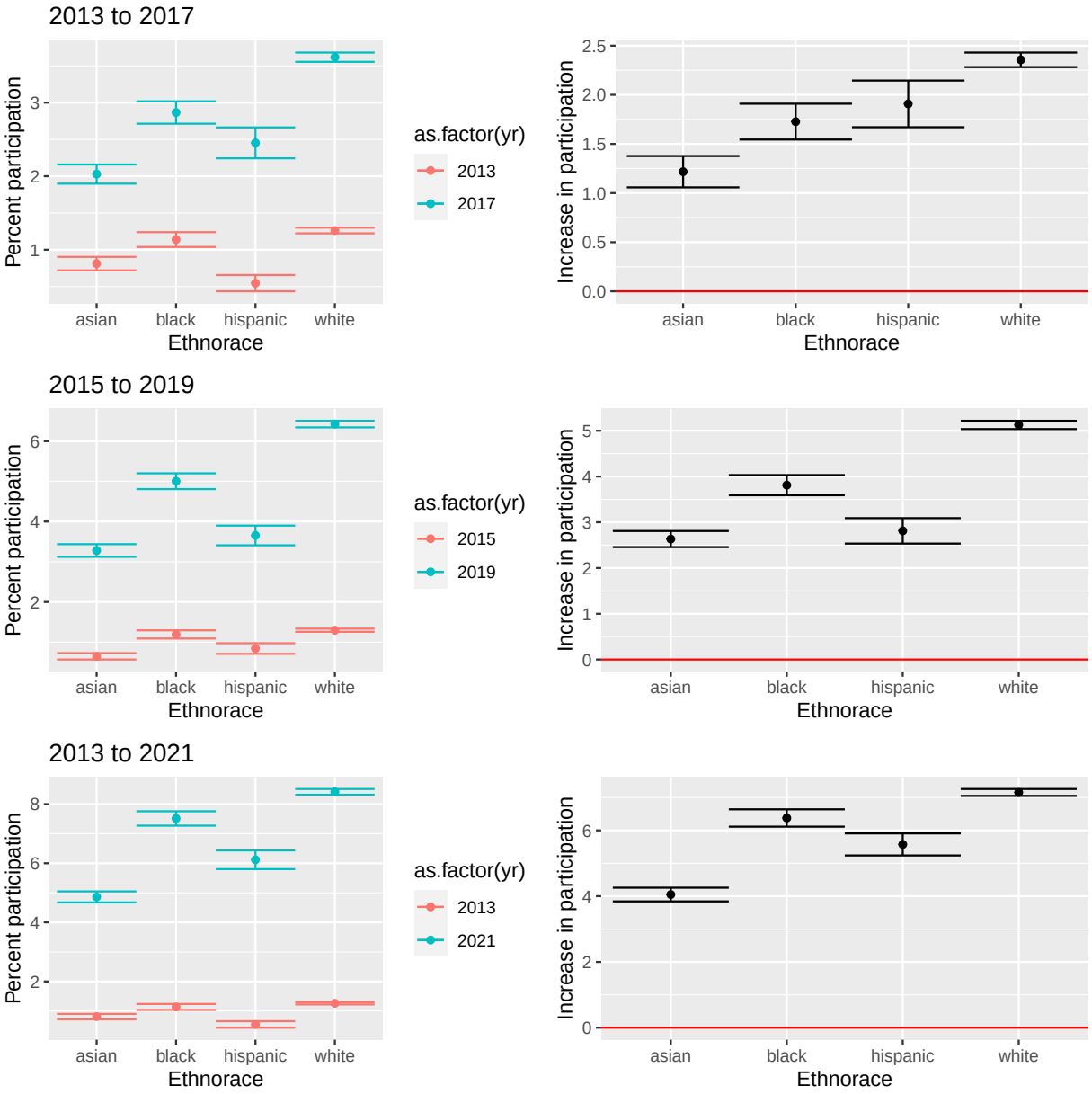
### 5.2.1 Effects for people outside the panel using a 2-year comparison

In (main text) footnote 2, I state that "Those who don't make the panel donate less often, but have similar participation gaps." I provide evidence for this claim here.

The regressions in Table A19 are run using individuals who were registered to vote in Seattle in both 2015 and 2021 (including those who may have not been registered in 2017 or 2019). This is a simple 2-year DiD setup with individual fixed effects. While this is not fully representative of all

[2] https://www.brookings.edu/research/tech-is-still-concentrating/

**Figure A10 – Ethnorace:** For each pre-post voucher comparison, more donor growth come from whites relative other racial groups. Plotted values are average rates of participation by race.




registered voters in Seattle, this subset is more representative than the main panel, who had to be registered to vote continuously 2005-2021. Here, I include an additional interaction effect, "incss2005," which is presence in the 2005-2021 panel.

In all cases, the VPP*incss2005 is positive and significant. The three-way interactions are all insignificant, each well-within a standard deviation of 0. For exposition, I will interpret a comparison using income percentile assuming zero individual fixed effects. Compared to 2015, there are the following amount of new donors per 100 people in 2019:

- For people in the 2005-2021 panel in the 3rd income decile, 6.007 + 1.384 + 0.129*(3) + (3*1*1)*.024 = 7.85
- For people not in the 2005-2021 panel in the 3rd income decile, 6.007 + 0.129*(3) = 6.394

**Table A19** – Voucher-year gains are lower for people not in the 2005-2021 panel. Individual and year fixed effects.

| | *Dependent variable:* | | |
|---|---|---|---|
| | Likelihood of donating | | |
| | (1) | (2) | (3) |
| VPP | 5.936 | 6.880 | 6.007 |
| | (0.112) | (0.074) | (0.151) |
| white × VPP | 1.130 | | |
| | (0.144) | | |
| VPP × incss2005 | 1.436 | 1.667 | 1.384 |
| | (0.189) | (0.116) | (0.230) |
| white × VPP × incss2005 | 0.011 | | |
| | (0.232) | | |
| republican × VPP | | -3.986 | |
| | | (0.221) | |
| republican × VPP × incss2005 | | 0.241 | |
| | | (0.317) | |
| income_perc × VPP | | | 0.129 |
| | | | (0.025) |
| income_perc × VPP × incss2005 | | | 0.024 |
| | | | (0.038) |
| Obs. | 580,726 | 580,726 | 570,214 |

- For people in the 2005-2021 panel in the 8th income decile, 6.007 + 1.384 + 0.129*(8) + (8*1*1)*.024=8.615
- For people not in the 2005-2021 panel in the 8th income decile, 6.007 + 0.129*(8) = 7.039

That is, for every 100 new donors in the 8th income decile in the 2005-2021 panel, there are 91 (=7.85/8.615) donors in the 3rd income decile in the 2005-2021 panel, donors in the 8th income decile but not in the 2005-2021 panel, and 74 donors in the 3rd income decile but not in the 2005-2021 panel.

# 6 Does panel start-year matter?

My design requires that individuals be continuously registered to vote in Seattle in each municipal election. 16 years is a comparatively long time for an individual to reside in a single place. Here, I vary the panel starting date. By moving the starting date forward, individual fixed effects are less plausible. However, we gain additional units, and the sample moves towards representativeness. Standard errors are robust for individual-level regressions and cluster robust for block-level regressions. For conciseness, I only show results for individual-level regressions, but block group-level regressions are similar.

**Table A20** – **Ethnorace** – Start year of the regression does not substantively change results.

| | *Dependent variable:* | | | | |
|---|---|---|---|---|---|
| | Likelihood of donating | | | | |
| | (1) | (2) | (3) | (4) | (5) |
| White*VPP | 2.55 | 2.32 | 2.36 | 2.19 | 2.03 |
| | (0.09) | (0.08) | (0.08) | (0.08) | (0.09) |
| Observations | 1,138,160 | 1,152,060 | 1,091,622 | 1,060,845 | 973,384 |
| Starting Year | 2007 | 2009 | 2011 | 2013 | 2015 |

**Table A21** – **Age, under 70** – Start year of the regression does not substantively change results. Over 70 results are omitted in the interest of space, but replicate similarly.

| | *Dependent variable:* | | | | |
|---|---|---|---|---|---|
| | Likelihood of donating | | | | |
| | (1) | (2) | (3) | (4) | (5) |
| Age*VPP | 0.001 | 0.05 | 0.05 | 0.04 | 0.05 |
| | (0.0000) | (0.004) | (0.003) | (0.003) | (0.003) |
| Observations | 1,145,762 | 1,097,673 | 1,114,006 | 1,025,053 | 942,693 |
| Subset | <70 | <70 | <70 | <70 | <70 |
| Starting Year | 2005 | 2007 | 2009 | 2013 | 2015 |

**Table A22** – **Party** – Start year of the regression does not substantively change results.

| | *Dependent variable:* | | | | |
|---|---|---|---|---|---|
| | Likelihood of donating | | | | |
| | (1) | (2) | (3) | (4) | (5) |
| Republican*VPP | −3.18 | −2.92 | −2.82 | −2.44 | −2.28 |
| | (0.11) | (0.11) | (0.10) | (0.11) | (0.13) |
| Observations | 1,271,776 | 1,288,469 | 1,222,110 | 1,188,950 | 1,093,012 |
| Starting Year | 2007 | 2009 | 2011 | 2013 | 2015 |

**Table A23** – **Income** – Start year of the regression does not substantively change results.

| | *Dependent variable:* | | | | |
|---|---|---|---|---|---|
| | Likelihood of donating | | | | |
| | (1) | (2) | (3) | (4) | (5) |
| Income decile*VPP | 0.20 | 0.18 | 0.18 | 0.14 | 0.14 |
| | (0.01) | (0.01) | (0.01) | (0.01) | (0.01) |
| Observations | 1,250,360 | 1,266,321 | 1,200,840 | 1,167,775 | 1,073,048 |
| Starting Year | 2007 | 2009 | 2011 | 2013 | 2015 |

**Table A24** – **Voting** – Start year of the regression does not substantively change results.

| | *Dependent variable:* | | | | |
|---|---|---|---|---|---|
| | Likelihood of donating | | | | |
| | (1) | (2) | (3) | (4) | (5) |
| Voting idx*VPP | 0.95 | 0.89 | 0.87 | 0.74 | 0.69 |
| | (0.008) | (0.008) | (0.008) | (0.008) | (0.009) |
| Observations | 1,271,776 | 1,288,469 | 1,222,110 | 1,188,950 | 1,093,012 |
| Starting Year | 2007 | 2009 | 2011 | 2013 | 2015 |

**Table A25 – Combined** – Start year of the regression does not substantively change results.

| | *Dependent variable:* | | | | |
|---|---|---|---|---|---|
| | Likelihood of donating | | | | |
| | (1) | (2) | (3) | (4) | (5) |
| White*VPP | 0.43 | 0.26 | 0.36 | 0.42 | 0.38 |
| | (0.10) | (0.09) | (0.09) | (0.08) | (0.10) |
| Income decile*VPP | 0.17 | 0.14 | 0.14 | 0.09 | 0.09 |
| | (0.01) | (0.01) | (0.01) | (0.01) | (0.02) |
| Mean blockgrp voting idx*VPP | 0.93 | 0.89 | 0.86 | 0.74 | 0.69 |
| | (0.009) | (0.009) | (0.008) | (0.008) | (0.01) |
| Republican*VPP | −3.30 | −3.23 | −3.26 | −3.00 | −3.06 |
| | (0.12) | (0.11) | (0.11) | (0.11) | (0.14) |
| Observations | 1,118,984 | 1,132,250 | 1,072,632 | 1,042,065 | 955,676 |
| Starting Year | 2007 | 2009 | 2011 | 2013 | 2015 |

# 7 Advances in the DiD method and placebo tests

The main tables in the text utilize a continuous difference-in-differences specification using panel data. This is certainly not the "standard" 2-groups, 2 (or more)-time period diff-in-diff, cross-sectional setup for which most of the advances in literature have addressed. In this section, I address some of the advances to the diff-in-diff method,[3] and where possible, utilize the method to show the robustness of my findings.

## 7.1 Difference-in-Differences with a Continuous Treatment: Callaway et al., 2021

This paper most closely addresses the research design utilized within the main text. It deals with utilizing TWFE diff-in-diff with treatments that are continuous, like VPP*income. TWFE estimands are problematic because they are weighted in unclear ways (and weights on observations can be negative), are unable to account for heterogeneous causal responses over time (and groups, when treatment is staggered), and TWFE output under standard parallel trends are mired by selection bias. The paper suggests estimands to overcome these first two issues. However, this is still currently a non-peer reviewed working paper. As of writing, no Stata or R package exists to implement these estimands, and the implementation is not straightforward.

## 7.2 De Chaisemartin and d'Haultfoeuille papers

There are multiple papers by these authors (De Chaisemartin & D'Haultfœuille, 2020; De Chaisemartin & D'Haultfoeuille, 2020; De Chaisemartin & d'Haultfoeuille, 2020). Each of these papers suggests a substitution of the TWFE results with results from packages did_multiplegt (Stata and R) or fuzzydid (Stata). While one of the core issues raised in these papers are variable treatment timing (which is not applicable in this case, as all become treated in 2017 with Democracy Vouchers), these estimands created by the authors are desirable for additional reasons as well.

As mentioned above, a TWFE issue is that weights may be negative, leading to the possibility that linear regression coefficients and ATE may be of different signs. A new estimator called the DIDM, immune to this issue, is proposed, focusing on when the switch occurs. They implement a placebo estimator, which evaluates the parallel trends assumption, looking at "switchers' and non-switchers' outcome evolution before switchers' treatment changes." A desirable outcome, indicating that the

[3] Asjad Naqvi's website, https://asjadnaqvi.github.io/DiD/, was incredibly helpful in aggregating papers on developments in the diff-in-diff literature and their corresponding software packages.

TWFE model results stand despite potential undesirable TWFE properties, is to find significance in the estimated effect of the treatment at the time period when switchers switch ("effect" in the table below) and for the placebo effect ("placebo") to not differ from 0. This is implemented through the R package did_multiplegt., I recover the following (specifications are placebo=1 and brep=4):

Here, most coefficients under "Effect," which represents DIDM that is immune to the negative weighting issue, are significant and in the same direction as in the TWFE model. All values under "Placebo" are statistically insignificant or in the opposite direction as the effect.

**Table A26** – Estimators from De Chaisemartin and d'Haultfoeuille's DIDM (left) and Sant'Anna and Zhao's DRDiD (right)

| | **DIDM** | | | | **Doubly Robust DiD** | | | |
|---|---|---|---|---|---|---|---|---|
| | **Effect** | **SE** | **Placebo** | **SE** | **15-19 Eff** | **SE** | **13-17 Eff** | **SE** |
| White | 1.38 | 0.111 | 0.102 | 0.079 | 4.547 | 0.232 | 2.015 | 0.196 |
| Republican | -2.63 | 0.158 | 0.004 | 0.064 | -3.069 | 0.309 | -3.939 | 0.221 |
| Freq voter | 4.24 | 0.132 | -0.065 | 0.115 | 10.369 | 0.204 | 6.97 | 0.169 |
| Lowest inc qrt | -0.510 | 0.079 | 0.096 | 0.097 | -2.044 | 0.223 | -1.314 | 0.183 |

## 7.3 Doubly robust difference-in-differences estimators Sant'Anna and Zhao, 2020

The parallel trends assumption is hard to interrogate and often implausible. This paper introduces new estimators for the ATT in diff-in-diffs where "the parallel trends assumption holds after you condition on a vector of pre-treatment covariates." These estimators can be calculated with the R package DRDID. Using 2015-to-2019 and 2013-to-2017 comparisons with the drdid_imp_panel command (panel data), doubly robust estimators are constructed using propensity scores to calculate the likelihood of being treated. Race (is white binary), Party (is Republican binary), voting frequency, and income are used to create propensity scores, with the outcome of interest being excluded (for example, in the first line, only party, voting frequency, and income are fed as covariates into the propensity score model). In all but one case (lowest income quartile for 2015-2019), effects are statistically significant and in the anticipated direction.

## 7.4 Main tables with placebo treatment time

**Table A27** – On the left, placebo test with 2005-2015 data, assuming 2013-15 are treated. On the right, placebo test with 2005-2015 data, assuming 2011-15 are treated. Compared to in-text Table 3-4, coefficients are much closer to 0. Beyond the placebo design, all specifications are the same as in-text Table 3-4.

| | *Dependent variable:* | | | | *Dependent variable:* | | | | |
|---|---|---|---|---|---|---|---|---|---|
| | % of block group donating | | | | Likelihood of donating | | | | |
| | (1) | (2) | (3) | (4) | (1) | (2) | (3) | (4) | (5) |
| VPP*Vtg idx | 0.18 | | | | 0.04 | | | | 0.04 |
| | (0.04) | | | | (0.005) | | | | (0.006) |
| VPP*White | | 0.005 | | | | 0.03 | | | −0.07 |
| | | (0.003) | | | | (0.06) | | | (0.06) |
| VPP*Med inc | | | 0.002 | | | | 0.004 | | 0.006 |
| | | | (0.0014) | | | | (0.009) | | (0.009) |
| VPP*Rep | | | | −0.043 | | | | −0.57 | −0.55 |
| | | | | (0.015) | | | | (0.06) | (0.07) |
| Observations | 3,429 | 3,429 | 3,003 | 2,857 | 879,756 | 787,734 | 865,122 | 879,756 | 774,618 |

## 7.5 Two year comparisons with placebo treatment times

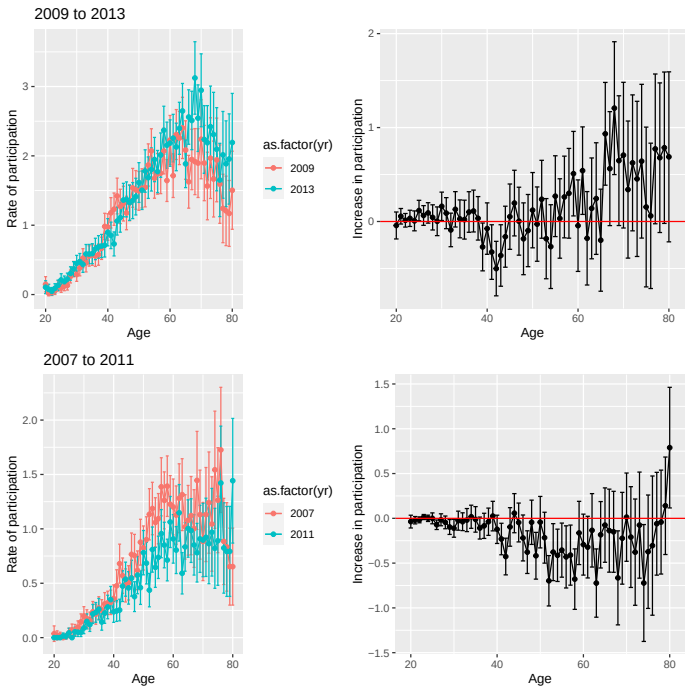



**Figure A11** – Along the dimension of age, there are not really any clear trends in the pre-treatment periods. Plotted values are average rates of participation by age.

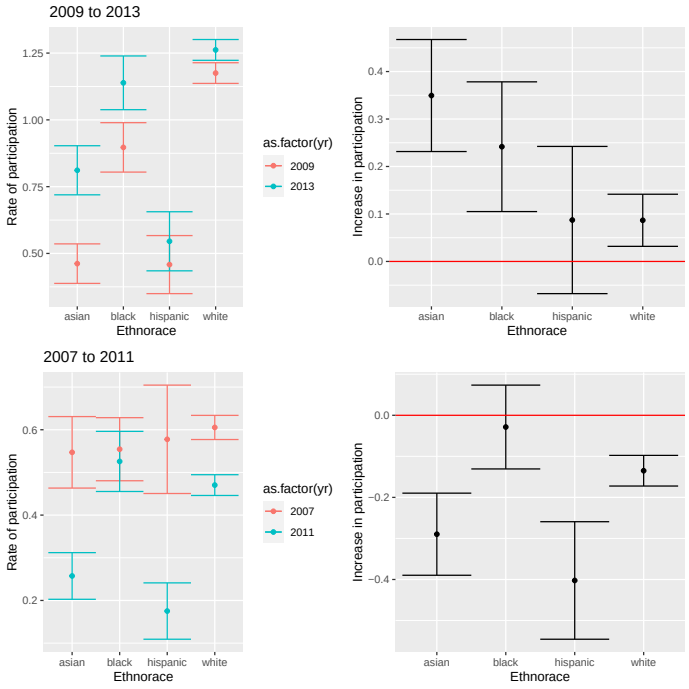



**Figure A12** – By race, there is some statistically significant year-to-year variation in white participation compared to nonwhite participation, but they don't mirror voucher-era trends. Notably, within each interval, the point estimates for increased participation is never highest among whites, whereas in Section A5.2, the point estimates among whites is always the highest within the differences graphs. Magnitudes are also substantively small. Plotted values are average rates of participation by race.

## References


Callaway, B., Goodman-Bacon, A., & Sant'Anna, P. H. (2021). Difference-in-differences with a continuous treatment. *arXiv preprint arXiv:2107.02637*.

De Chaisemartin, C., & D'Haultfœuille, X. (2020). Two-way fixed effects regressions with several treatments. *Available at SSRN 3751060*.

De Chaisemartin, C., & D'Haultfoeuille, X. (2020). Difference-in-differences estimators of intertemporal treatment effects. *arXiv preprint arXiv:2007.04267*.

De Chaisemartin, C., & d'Haultfoeuille, X. (2020). Two-way fixed effects estimators with heterogeneous treatment effects. *American Economic Review*, *110*(9), 2964–96.

Imai, K., & Khanna, K. (2016). Improving ecological inference by predicting individual ethnicity from voter registration records. *Political Analysis*, *24*(2), 263–272.

Sant'Anna, P. H., & Zhao, J. (2020). Doubly robust difference-in-differences estimators. *Journal of Econometrics*, *219*(1), 101–122.