Participatory budgeting and voter turnout

Carolina Johnson H. Jacob Carlson Sonya Reynolds

Working Paper Draft: September 28, 2018

Do not cite or circulate without permission

1 Introduction

The "participation hypothesis" suggests that the act of participating in civic activities encourages further engagement in other spheres of civic life (Finkel 1985, Gastil et al. 2008, Mansbridge 1995). Yet Mansbridge (1995) advances this claim while observing "[p]articipation does make better citizens. I believe it, but I can't prove it. And neither can anyone else" (p. 1). Testing the hypothesis that participation begets more participation is made difficult by the problem of selection bias in favor of those otherwise predisposed to participate. Some have attempted to isolate causal estimates of the effect of prior participation on subsequent participation (Gastil et al. 2008, Gerber et al. 2003, Plutzer 2002). While the specific form of the participation outcome varies, a common measure is later voter turnout. In this paper, we examine the effect of individual involvement in participatory budgeting in New York City on subsequent voter turnout.

Participatory budgeting (PB) is a process where everyday people come up with ideas for and vote on how to spend public money. By getting people directly involved in issues that affect their community, many hope that this form of participatory democracy will have broader impacts on the civic and political life of communities, including on electoral participation. Indeed, some youth in Boston report that as a result of participating in PB, they are more likely to vote in regular elections (Grillos 2014: 25). Many PB participants from other localities say that they are more likely get involved in their communities as a result of PB (Crum, Salinas, Weber 2013; Jovanovich and Russell 2016). Yet existing data on this question is relatively sparse, and ultimately relies on self-reported data about future behaviors.

In this project, we use New York State voter file to measure the effect of PB participation on voter turnout in regular elections. We link individual participants in PB in New York City to their state voter records in order to assess, via logistic difference-in-difference regression models, whether people who participate in PB are subsequently more likely to vote in regular elections. To help isolate the effect of PB, we use coarsened exact matching (CEM) to match PB voters to otherwise similar voters in districts where people did not have the opportunity to participate in PB. Comparing PB voters to similar individuals that we would expect to have participated in PB if they had the chance, we find that engaging with participatory budgeting increased individuals predicted probability of voting by over 7%, on average. This is a substantial effect, which is amplified for minority voters and those in less well educated and wealthy neighborhoods.

The paper proceeds as follows: First we offer a brief introduction to participatory budgeting, in particular as implemented in New York City. We provide an empirical and theoretical justification for our hypothesis that participatory budgeting participating may be associated with an increase in regular voter turnout. Along the way, we also review other factors known to influence individual voter turnout, as key elements to control when identifying the non-PB comparison group and in our final analysis. Next, we describe our data and the procedures for matching PB voters to similar non-PB voters. Finally, we present the results of our analysis, including detailed breakdowns of the expected effect from PB for specific subgroups in our population, including breakdowns by individual and neighborhood characteristics.

2 Background

2.1 Participatory Budgeting

Participatory budgeting has grown in the US since it first came to Chicago in 2009. Since then, over 433,000 people have voted across over 30 cities in the US and Canada, directly deciding how to spend almost $300 million in public money. In PB, potential budget items are directly collected from residents, rather than generated by government officials. Through a process of deliberation and analysis, the collected budget ideas are then developed and refined by fellow residents, in partnership with government staff. A narrowed list of proposals are then placed on a ballot for other community members to directly vote for which items should be funded. Winning items are then implemented. While the PB process can be applied to almost any type of budget, most cases of PB in North America have been applied to portions of the capital budgets, often distributed by city council districts or neighborhoods. In such cases, rather than the city council members deciding on how to spend discretionary money in their district, they use a PB process to give decision-making power over to their city residents.

Each local PB process has its own rules for who is eligible, but in North America, people without traditional voting rights - non-citizens, people under the voting age, or those disenfranchised due to felony conviction - are usually eligible to vote in PB. Authorization to vote or participate in PB typically has minimal requirements, making it easy for people to partake in the process. Since the process is unfamiliar for many residents, those implementing the PB process go to lengths to publicize it, including street outreach in public places, text message and phone recruitment, advertisements, and other methods.

Many have examined PB as a form of participatory democracy with the ability to reshape governance institutions in more democratic ways (Fung and Wright 2003, Cabannes 2004). It is a way to create avenues for civic participation and political voice for underrepresented or marginalized people (Baiocchi 2003), though the depth of these effects are not guaranteed (Baiocchi and Ganuza 2014). While the participants of PB in the US tend to be more highly educated, there also often is an over-representation of black voters and low-income voters, relative to the neighborhood in which it takes place (Hagelskamp et al. 2016: 6).

In general, research on participatory budgeting has been mostly descriptive, providing accounts of the process as well as documenting the composition of who participates. There has been less systematic investigation on the measurable impacts of PB on individuals and their communities (Wampler et al. 2018). The few projects that have examined PB's measurable impacts include studies exploring PB's impacts on the structure of civic communication and collaboration networks (Johnson 2017), numbers of civil society organizations (Wampler 2012), and health care spending and infant mortality rates (Gonçalves 2014, Touchton and Wampler 2014). Lerner and Schugurensky (2007) document the ways that PB educates individuals with new knowledge and skills and shifts attitudes and practice, serving as schools of citizenship that may support future civic or political engagement.

PB first came to New York City in 2011 through four city council members, and has since been implemented in 31 of the 51 districts.[1] There is a central city-wide steering committee that oversees and supports PB across the city, though a district's implementation ultimately depends on the local city council member's staff. In 2018, the typical district PB uses $1 million, and has around 7,000 voters in an average district size of about 164,000. The average project costs almost $263,000, and range from new school facilities, to new park renovations, to security cameras, to street repairs, and more.[2] PB voters in New York City are diverse. Surveys conducted throughout the first 4 years of PB in New York City have described PB voters as broadly representative of their districts in terms of racial composition and household income. One exception was the common pattern that PB voters often reported higher levels of education than their districts in general (Community Development Project 2015).

2.2 Participatory democracy and voter turnout

One of the core aspects of participatory democratic theory is that people learn how to participate by participating (Pateman 1970). While there are many different types of participation (see Rowe and

[1] For more information on Participatory Budgeting in New York City, particularly in the earlier years of its implementation, see the accounts in ??Harvard grad; Johnson 2017; Evaluation reports; and Celina/Josh accounts in JPD and elsewhere

[2] With one exception, only capital projects are eligible for PB funding, not operations funding.

Frewer 2005), which could have differential effects, the idea is that participation creates a virtuous cycle that can lead to more participation.

One area of interest for the secondary effects of participation is how it leads to participation in more traditional democratic practices, such as voting in elections. Dvořák et al. (2017) find that direct democracy in the form of citizen initiatives and referenda in the Czech Republic led to increased subsequent turnout in the short term. Gastil et al. (2008) evaluate a specific case of civic participation, that of jury participation, to estimate causal effects of jury deliberation on reported attitudes as well as actual election turnout. They find that the effect on turnout increases with more intense forms of jury deliberation. While the specific psycho-social mechanism of participation require more research, activating one's sense of civic duty has been shown to have positive effects on voter turnout (Gerber et al. 2008, Blais and Achen 2018).

2.3 Factors that affect turnout

The confounding factor in determining whether participation causes people to vote more is that there is likely selection bias in determining who participates and who votes - people who are more likely to participate in other avenues of civil society are also more likely to vote in regular elections. Given that random assignment of treatment for a full experimental design is not feasible, we use matching methods to construct a comparison group from the full population of voters in New York City. In order to adequately specify matching conditions and fit appropriate models to identify any effect from PB, we need to account for other factors that are known to affect turnout rates.

One of the most consistent predictors of voter turnout is one's prior voting history (Cancela and Geys 2016). Denny and Doyle (2009) find that voting in a previous election increases the probability of voting in a subsequent election by 13%. People are often habitual voters or habitual non-voters (Plutzer 2002). Fowler (2002) estimates that half of voters either always vote or always abstain. A meta-analysis by Cancela and Geys (2016) shows that 86% of studies measuring the effect of past turnout on subsequent turnout found a positive and statistically significant effect (p. 267). Gerber et al. (2003) attempt to estimate a causal effect of voting habit on subsequent elections. Their study finds a 45%

greater chance of voting among those registered voters who had voted the year prior compared to those who did not.

Education level is a consistent predictor of voter turnout. More education correlates with increased voter turnout with "law-like regularity" (Sondheimer and Green 2010: 174). Some have argued that education is not the mechanism that affects voter turnout, but rather is a proxy for other factors that cause people to be more likely to vote (Kam and Palmer 2008, Tenn 2007). Sondheimer and Green (2010) draw on experimental and quasi-experimental data to attribute causal effects of education on voter turnout.

Race of the voter as well as the context in which they live have been shown to be related to voter turnout. Many find that being a member of the majority racial group in a district has a positive effect on turnout (Fraga 2016a, Fraga 2016b, Barreto et al. 2004), while Geys' (2006) meta-analysis finds that higher shares of ethnic and racial minorities are negatively correlated with turnout.

Turnout is often higher in competitive races. Indeed it is one of the strongest predictors in aggregate level turnout (Franklin 2004). Meta-analysis suggests that closeness of an election is positively associated with turnout (Geys 2006). There are competing explanations of why this is the case. One common explanation is the Downsian argument that people are more likely to vote when they believe that they might cast a decisive ballot (see Matsusaka and Palda 1993). Others argue that close elections induce more elite mobilization and campaigning (Cox and Munge 1989). Regardless of the underlying causal mechanism, competitiveness is largely a consistent predictor of voter turnout.

3 Data and Method

We use individual level Participatory Budgeting voter record data, acquired from the central Participatory Budgeting-New York City body, as well as directly from districts. Through our partnership with NY Civic Engagement Table, we access Catalist to get voter histories as well as other individual level data that we use to match the PB voters with their voter histories. We then use coarsened exact matching to identify comparison group of voters that is approximately identical in districts where PB is not available. This comparison enables us to estimate the effect of participating in PB by seeing how PB voters' turnout for elections changes after having participated in a PB election, compared to the voting behavior of similar people with similar past voting behavior who did not have the opportunity to participate in PB. To make these comparisons, we then use logistic difference-in-difference models to estimate predicted probabilities of voting for PB voters. We additionally estimate predicted probabilities for various subgroupings of voters.

3.1 Data

This analysis uses individual and census tract level data on PB and non-PB voters. The individual data is provided by Catalist. This includes full voter history, age, and address, and registered party, which come directly from the New York State Voter File. Catalist additionally includes an estimated race variable using a proprietary algorithm, which has been used with high success in previous research (Fraga 2016a, Fraga 2016b, Ansolabehere and Hersh 2012). It uses names, local census information, and commercial data to estimate a voter's race.

Along with the individual level data on voting history, age, race, and address (which is used to identify the election district and census tract), we include census tract-level variables from the American Community Survey (US Census Bureau 2015). When matched with individual level data, localized contextual data has been used as proxies for missing individual data, such as income (Hersh and Nall 2016). We include median household income and percent of the population with different levels of educational attainment as a proxy for important unmeasured individual traits of income and education. We also include the percent of the population that is white, to account for the effects of racial compostion, as well as whether a voter's estimated race matches the majority race of their census tract (Fraga 2016b). Construction of the sample population for both PB and non-PB voters, including the matching process, is described below.

3.1.1 PB voters

Identifying PB voters, and matching them to their voting histories, was the most challenging part of data collection. For the first several years of PB, some districts and central city staff collected basic identifying information about voters, including name and date of birth, usually as part of the process of validating voters' eligibility. A complicating factor is that different districts implement PB at different points in time, and voters may, or may not, vote in multiple years of PB. To create our dataset, staff at the Participatory Budgeting Project obtained the PB voter lists for districts who held a PB process in New York City from 2013-2016. In the first years, voter lists were not held in a central location, but rather by local district offices. PBP staff contacted all 27 city council districts who held a participatory budgeting process in 2015 to request their list of voters in their respective districts. 15 districts replied, but only two districts had multiple years of voter lists, while the remainder only had records for the 2016 voting cycle. Missing vote years mean we may incorrectly identify someone's first year of participation in PB, assigning them incorrectly to a later year in which they voted. Misidentifying the year people first voted in PB introduces possible bias in the analysis. To minimize the possibility of this happening, we do not include voters from districts for which we are missing data from early cycles. Our final set of 7,634 PB voters included voters from early-adopting districts 39 and 23 (which were able to provide voters from all years) and the late adopting districts 30, 35, 36, and 40 (whose first PB vote occurred in 2016, the first year for which we have data from all districts.[3]

This analysis uses individual and census tract level data on PB and non-PB voters. The individual data is provided by Catalist through the NY Civic Engagement Table and their partnership with Participatory Budgeting Project and PBNYC. The voter lists had full name, address, birth date, and sometimes phone numbers to assist with linking PB voters to the state voter file. We were able to match 10,030 with their state voter file ID numbers. It is important to note that PB voting procedures are deliberately more inclusive than traditional electoral processes, and extend voting rights to youth, non-citizens, and felons who may not be eligible to vote in regular elections, and thus will not be found in the voting file. For those we did not match, we do not know for certain whether this is because they were not registered to vote, ineligible to register, or we could not match them for some other reason (for example an error in a PB voter's recorded name or date of birth). Finally, for this analysis, we

[3] In addition, a small subset of voters whose votes were recorded in 2012 but not definitely assigned to a PB district are also included, as the first PB votes occurred in 2012 and thus we are confident of recording these voters initial experience with PB.

limited our sample to individuals with a current voting address within New York City. This left us with a final PB voter sample size of 7,634 people.

3.1.2 Non-PB voters

Non-PB voters were identified from the New York State voterfile via Catalist. Having obtained the full voterfile, we excluded all voters in districts that had implemented PB as of 2017 as well as any voters in those non-PB districts who were listed as a PB voter (this could happen in the case of someone moving from a PB to non-PB district during the timeframe of this study). We excluded all voters from districts that had implemented PB to minimize bias from two possible sources: 1) without full PB voter records in all PB districts, we cannot ensure that a voter who would match a PB voter did not themselves vote in PB (which would bias any estimated effect of PB downward) and 2) a non-PB voter in a PB district has demonstrable differences from a PB voter (by virtue of opting out from the PB process). The important comparison is between PB voters and *voters who would by the type to get involved in PB* if it were an option available to them.

20 of the 51 New York City Council districts have never implemented PB. Voters from these 20 districts make up the population of potential matches with the PB voters, as described in the matching process below.

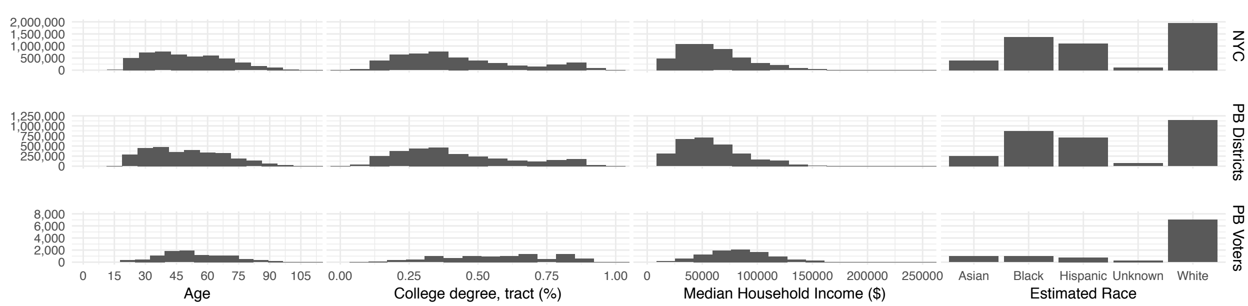

Figure 1: Characteristics of voters in New York City as a whole, PB districts in NYC, and PB voters whose data was collected and included in our sample.

Figures 1 and 2 illustrate the general characteristics of NYC voters as a whole, voters in PB districts (regardless of their participation in PB), and the PB voters that made it into our final sample. While PB districts in general resemble the city as a whole, PB voters in our sample tend to come from more

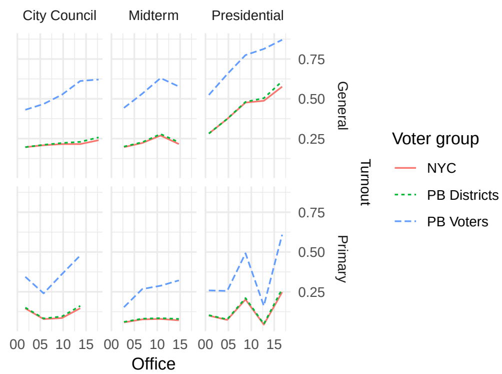

Figure 2: Voter turnout of voters in New York City as a whole, PB districts in NYC, and PB voters whose data was collected and included in our sample.

well educated neighborhoods and to be much more often white than the population. The matching method described below ensures that our analysis compares our PB voters to similar non-PB voters elsewhere in the city.

3.2 Matching

The objective of this analysis is to understand the effect of participation in PB on future electoral participation. Since we cannot directly compare PB voters behavior to the counterfactual scenario in which they did not experience PB, we use matching methods to construct a control group of registered voters in districts where PB is not available that is balanced with our 'treatment group' of PB voters on all covariates that are expected to predict voting, while restricting our analysis only to those individuals who would be likely to to participate in PB if given the chance. While matching methods do not replicate experimental control, such an approach gives us the best opportunity to estimate the observable impact of participatory budgeting with purely observational data.[4]

Given our large set of potential matches and high level of overlap in population characteristics between our PB voters and the whole city voting population, we use one-to-one coarsened exact matching (Iacus et al. 2012), a technique that directly matches two subjects on a set of relevant covariates, reducing model dependence and ensuring balance on even a large number of covariates (such as the

[4] An extensive literature on statistical matching methods exists. For a good review of different methods, see King et al. 2011; Iacus et al. 2012; or Stuart 2010.

Table 1: Match Criteria

	Individual-level	Tract/Election District-level
Age	Age[i]	
Gender	Male/Female[ii]	-
Race	White, Black, Asian, Latinx, Unknown[ii]	
	Membership in majority racial group in census tract [ii]	
		Tract: % of population in identified as white[iii]
Voting History	Indicator for each 2008-2012 election[ii]	
	Count of total votes cast 2000-2007 [v]	
Education		% population with college degree [iii]
Income		Tract: Median household income[iii]
Electoral competitiveness		Election District: % Margin of victory, top vote getting race in each ED 2014 & 2016[iv]

[i]Matching ages within 5 year intervals (15-19, 20-24, 25-29 etc.) [ii]Exact match. [iii]Match within quintiles (Bottom 20% of tracts, next 20-40% percent of tracts, etc). [iv]Match within low, medium, high competition districts in 2014 primary and general and 2016 presidential primary and general. Note that earlier election district-level vote results are not available from the NYC Board of elections. ED-level data are only publicly available for 2014 onward. [v] Voters' total ballots cast in primary and general elections from 2000 - 2007 are separately summed (maximum possible 8 votes for each general and primary counts. Match is by votes in group of 0, (1, 2), (3, 4), (5,6) or (7,8) total votes.

many predictors of individual turnout).[5] A match is only achieved if they are effectively identical on the specified covariates. The match is "coarse," because while some variables like gender are easy to match exactly, continuous variables like age and income are more difficult to match exactly, and thus acceptable ranges for matching are determined.

For this analysis, we use a mix of individual and census-tract level variables, defined in Table 1. Using these strict criteria, we are able to successfully match 3425 , or 45??% of our 7,634 PB voters to a non-PB voter. We can increase our match rate substantially by dropping the census tract level variables and relaxing the coarseness of the match, but were not willing to accept the level of imbalance in contextual neighborhood factors that resulted. After matching, we have a dataset including 151,242 [confirm at end??] voter-election observations. The goal here is to find an adequate sample of comparable voters, pruning our data to be balanced between treatment and control groups, not simply to find a nearest match for every PB voter.

[5]With many more controls than PB voters, we randomly sample one matching control individual for every PB voter in each strata.

3.3 Effect Estimation Method

We implement a difference-in-differences design for a binary outcome of whether an individual voted or not in a given election. We use logit models that include random effects to account for the hierarchical clustering of observations within individuals and districts. In addition to the indicators of membership in the 'treatment' group (PB voters) and whether an election occurred before or after an individual's first PB vote (the equivalent of the treatment*time interaction in a simple two period difference in difference design), we include dummy variables for type of election and election years, as well as including individual and tract level controls.

The base model thus takes the form:

$$Pr(y_i = 1) = \text{logit}^{-1}(\gamma PB_i + \lambda_y + \delta A_{iy} + \mathrm{X}_{\mathrm{i}}'\beta + \vartheta M_v + \alpha_D + \alpha_i)$$

where

$$\alpha_D \sim N(0, \sigma_D), \text{for} D = 1, ..., 47$$

$$\alpha_i \sim N(0, \sigma_i), \text{for} i = 1, ..., 6734$$

In this notation, PB_i is the indicator of a individual's membership in the group of PB (treatment) or non-PB (control) voters , λ_y are the fixed effect estimates for each year capturing the baseline time trend, A_{iy} is the interaction between treatment group membership and an election occurring after an individual's first PB vote, making δ the difference-in-difference estimate: the expected effect of PB. $\mathbf{X}_{\mathrm{i}}$ is a vector of individual covariates, including age, gender, race, membership in the local majority racial group, neighborhood median household income, percent of the population with college degrees in the neighborhood. M_v are fixed effects for the type of election, general, primary, or presidential primary, for each individual vote observation. α_D and α_i random effects for district D and individual i.

We also do not assume that the effect of PB is constant across all groups within the population. To test the heterogeneity of effects among groups, we also fit a series of models including separate interaction terms between the indicator for an election taking place after first voting in PB and various subgroup classifications, including by race, youth, gender, high/low education, and high/low median household income.

These interaction models take the form:

$$Pr(y_i = 1) = \text{logit}^{-1}(\gamma PB_i + \lambda_y + \delta A_{iy} + \eta A_{iy}S_i + \tau_{Siy} + \mathrm{X}_{\mathrm{i}}'\beta + \vartheta M_v + \alpha_D + \alpha_i)$$

These interaction models include the same general terms as the base model above, with the additions of the necessary interaction terms for the "triple difference" estimates of the conditional effect of PB on different subgroups. $A_{iy}S_i$ is the interaction term of the 'after PB' indicator with the subgroup, with η providing the estimate of the conditioned effect of PB for the covariate of interest. τ_{Siy} is the interaction of the subgroup variable with the year fixed effect dummies, accounting for the subgroup-specific time trends.

Effect estimates from logit models, especially when interactions between variables are included, require additional interpretation to be meaningful in terms of direct quantities of interest. All results from models below are thus presented in the form of changes in the predicted probability of turnout. Tables of regression coefficients and significance tests are available in the appendix. Confidence intervals for predicted probabilities were generated by simulation.

4 Results

This analysis demonstrates a strong and statistically significant effect of voting in participatory budgeting on the probability of voting in subsequent regular elections. On average, participation in PB is associated with an average increase of 7.8% in an individual's predicted probability of voting, all else being equal.[6] A positive significant effect of PB is evident in a minimal model including only the mini-

[6] 7.8% is the marginal effect of the post-treatment variable in the best-fitting diff-in-diff model, without including any interactions of the post-treatment indicator with other covariates.

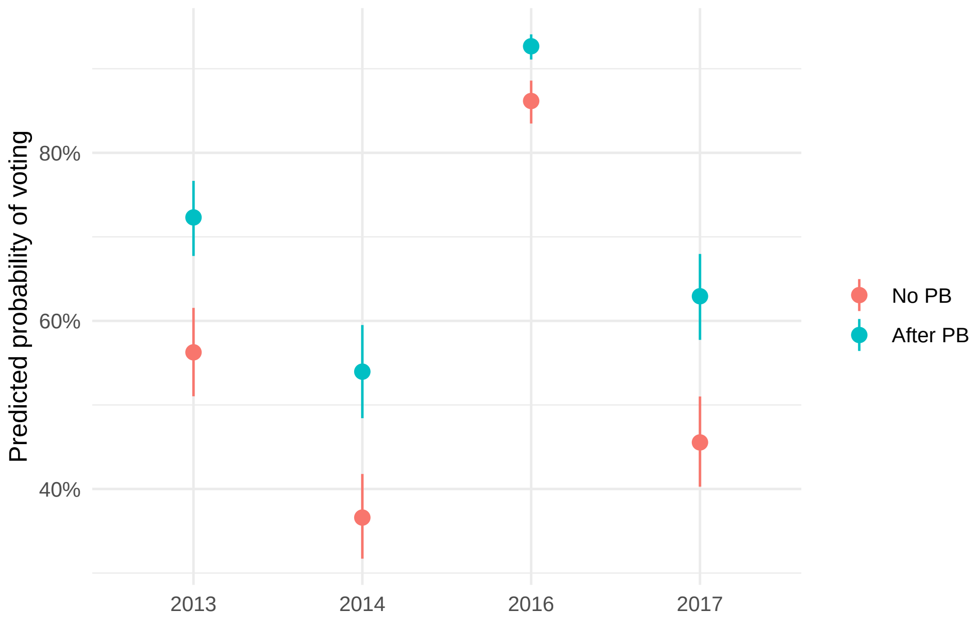

Figure 3: Predicted probability of voting in general elections, with 95% confidencce intervals. Predictions for non-PB voters, showing effect of hypothetical participation in PB vote. 95% confidence intervals generated by simulation.

mal indicators necessary for a difference in difference analysis. It not only persists, but is strengthened by inclusion of individual, election, and tract-level covariates. Table 2 presents the model estimates of both the minimal difference-in-difference model and the models including each level of additional covariates.[7] Note that membership in PB-implementing district has no significant association with voter turnout, validating the effectiveness of the matching method for identifying a comparison group with equal baseline likelihood of voting.

The coefficient estimates demonstrate a robust positive effect of PB. We illustrate the magnitude of this effect with predicted probabilities of voting, since direct substantive interpretation of the logit difference in difference coefficients is not straightforward. We calculate predicted probabilities of voting with and without hypothetical participation in PB to illustrate the magnitude of the effect of PB. Figure 3 illustrates the difference in the predicted probability of voting in general elections since PB was first implemented in New York City. This figure compares the estimated turnout for non-PB

[7]Model selection processes used in-sample percent correctly predicted and AIC/BIC criteria to identify and exclude appropriate transformations and redundant variables, such as a census-tract flag for majority white population (this variable was not informative once race and flag for majority race membership was included).

Table 2

	(1)	(2)	(3)	(4)
pb	−0.350	−0.176	−0.078	−0.083
	(0.246)	(0.182)	(0.155)	(0.153)
after_pb	0.736***	0.709***	0.710***	0.711***
	(0.030)	(0.030)	(0.030)	(0.030)
election_typep	−2.433***	−2.458***	−2.458***	−2.458***
	(0.020)	(0.020)	(0.020)	(0.020)
election_typepp	−2.534***	−2.578***	−2.578***	−2.578***
	(0.033)	(0.033)	(0.033)	(0.033)
RaceA		−1.913***	−1.746***	−1.668***
		(0.090)	(0.092)	(0.094)
RaceB		−1.102***	−0.876***	−0.724***
		(0.130)	(0.129)	(0.132)
RaceH		−1.289***	−1.140***	−0.867***
		(0.125)	(0.127)	(0.138)
RaceU		−1.721***	−1.595***	−1.277***
		(0.222)	(0.222)	(0.231)
Female		0.127**	0.123**	0.125**
		(0.051)	(0.051)	(0.051)
age		0.096***	0.093***	0.090***
		(0.009)	(0.009)	(0.009)
I(age^2)		−0.0004***	−0.0003***	−0.0003***
		(0.0001)	(0.0001)	(0.0001)
I(age_at_vote <18)		−7.663***	−7.665***	−7.671***
		(1.009)	(1.009)	(1.009)
college_pct			2.042***	1.970***
			(0.285)	(0.284)
majmatch				0.374***
				(0.078)
medhhinc_10k			−0.034***	−0.037***
			(0.013)	(0.013)
Constant	1.023***	−2.076***	−2.832***	−3.060***
	(0.208)	(0.286)	(0.290)	(0.292)
Year fixed effects?	Yes	Yes	Yes	Yes
Observations	141414	141414	141414	141414
Akaike Inf. Crit.	109147.900	105197.900	105142.600	105119.500
Bayesian Inf. Crit.	109305.700	105434.600	105398.900	105385.700

Note: *p<0.1; **p<0.05; ***p<0.01

voters in two scenarios - the real historical scenario and the counterfactual alternative where these voters had participated in a PB process in their own district. As Figure 3 shows, depending on the baseline predicted turnout for different election years, the substantive average effect of PB varies from just over 6% increase in the probability of turning out in the 2016 general election to over an increase of over 16% points in the 2014 and 2017 elections.

Logit models appropriately model the outcome of interest as a binary outcome (voted/did not vote). This modelling choice imposes a reasonable non-linearity on any effect of PB: If a person already has a history of voting in every election, PB cannot have a very large impact on their probability of voting, because they no longer have any room to improve. On the other hand, a person who rarely votes has the potential of a much more significant impact on their future behavior. This non-linear effect is why, in Figure 3, the difference between voting with and without PB is smaller in the 2016 presidential election, compared to the other lower turnout years. Presidential election years already see higher turnout in general, so the effect of PB is less evident. This implies a potentially stronger impact from PB during off-cycle election years when turnout is typically lower for most voters.

4.1 Subgroup Breakdowns

The foundational model presented above assume that the effect of PB is the same across all different types of voters. This is an assumption that can be directly assessed by including interactions between the effect of PB and specific groups of interest in the population, effectively making the effect of PB conditional on other variables, for example a voter's race or how well-off their immediate neighborhood is. We explore the conditional effect of PB across several different group breakdowns: by race, gender, youth voters (under 30)[8], and neighborhood characteristics of voters including wealth and education. Figure 4 shows the differential impact of PB within each of these subgroups comparisons. The x axis shows the expected change in the probability probability of voting for typical non-PB voters, were they to have participated in a PB process prior to the general election in 2016. Separate models are fit conditioning the effect of PB on each variable in turn.

[8] Voters are flagged as 'youth' voters if they were under 30 at the time of analysis

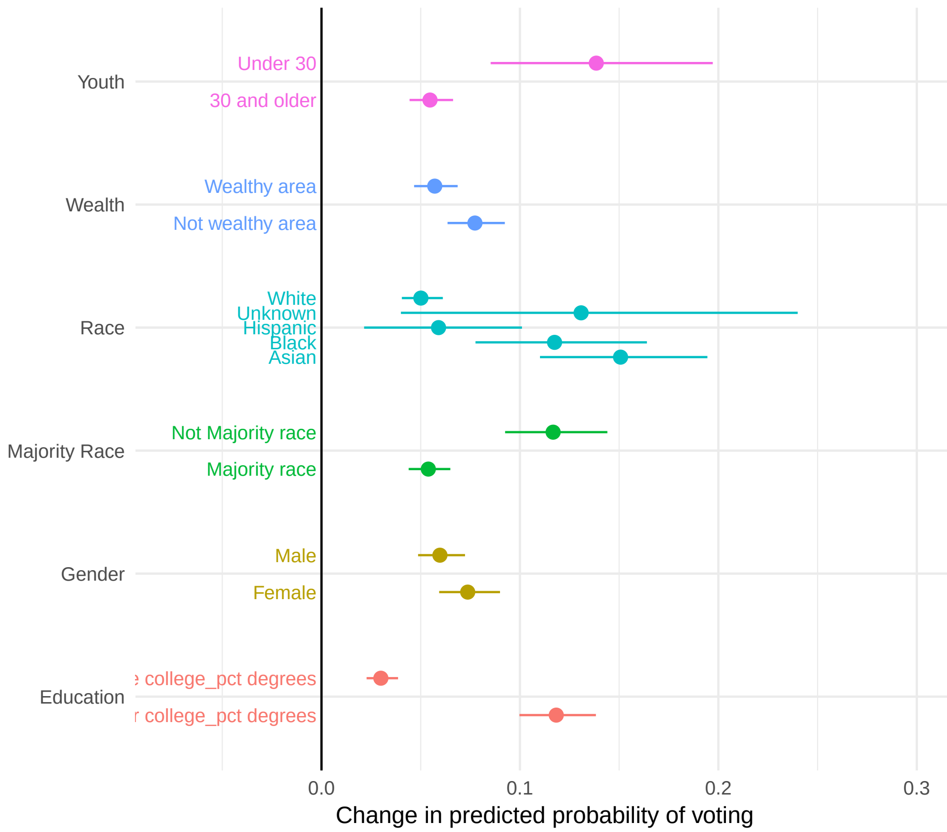

Figure 4: Change in predicted probabilities of voting in the 2016 general election with and without counterfactual participation in PB, with effect conditional on subgroup. Interaction effects for each grouping variable estimated in separate models, including all base terms as covariates. 95% confidence intervals generated by simulation. See Table **??** in the appendix for details of model estimates.

The effect of PB remains consistently significant and positive even when allowed to interact with each key covariate. Within this generally positive effect, however, as Figure 4 shows, PB can indeed have substantially and significantly different effects for different subgroups. The prevalence of post-secondary education in the census tract and membership in the majority race in the census tract are the two variables with the strongest mediating influence on the effect of PB, with individuals in census tracts with fewer college educated people (as a percentage of the population) and who are not members of the majority racial group seeing large increases in the effect of PB - translating to an additional increase in the probability of voting in 2016 of over 10% and 5%, respectively.

Other major dimensions of influence on the effect of PB include the the contrast between older and younger voters and median household income of the voter's census tract. Younger voters, while having a lower predicted probability of voting, demonstrate a significantly and substantially larger effect from PB, while voters in less wealthy areas also experience slightly but significantly larger effects from PB. Figure 4 shows that different racial groups see statistically significantly different effects from PB, with white voters demonstrating the smallest effect and black and Asian voters showing the largest effects. The large expected differences in the probability of voting with and without PB are not directly a result of the interaction of race with PB (the interaction is non-significant in the model), but rather illustrate how the overall effect of PB is amplified more for minority racial groups because of their generally lower probabilities of voting, especially when a member of a minority within their local area. Whites are generally more likely to vote than any other racial group in most elections, while black and Asians are significantly less likely to vote. These difference allows PB to have greater leverage among Asian and black voters than white voters. There is evidence for only very small differential effects of PB on the basis of gender.

For comparison, Figure 5 shows the estimated effect of PB in the 2017 general election. This was a local election, with overall lower turnout across all groups. This figure highlights how the strongest mediating effects (majority racial group membership, prevalence of post-secondary education, and local median household income) continue to support different total increase in predicted probability of voting as a result of PB. In this election where all groups are less likely to vote, the consistency of

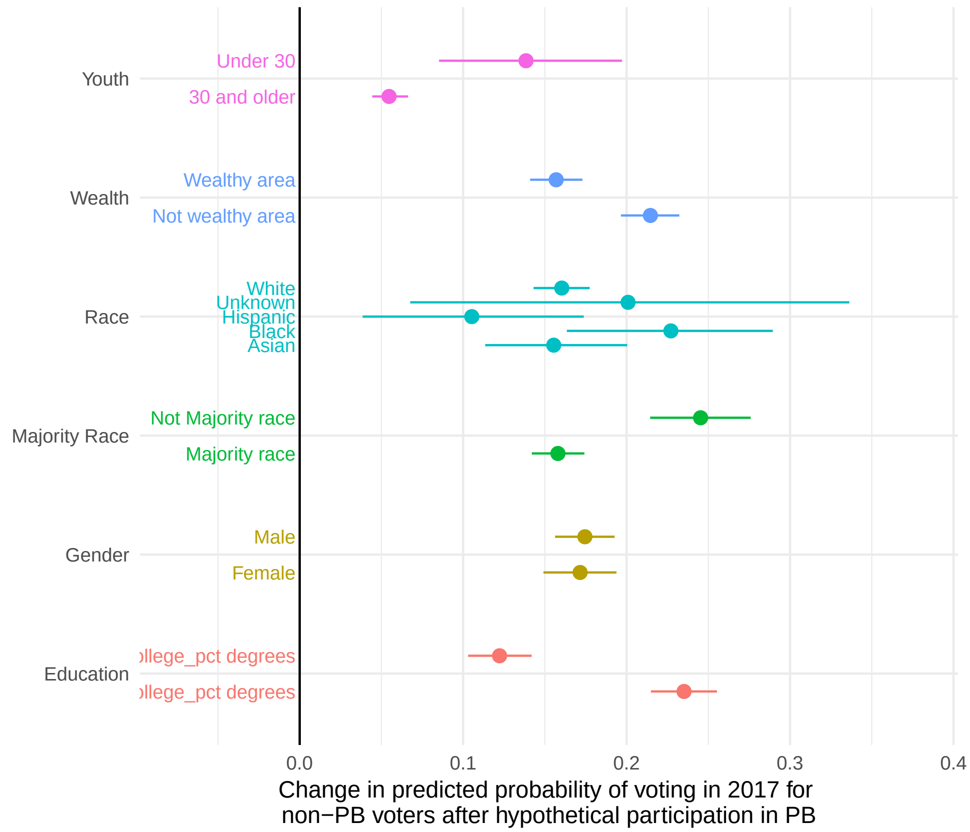

Figure 5: Change in predicted probabilities of voting in the 2017 general election with and without counterfactual participation in PB, with effect conditional on subgroup. Interaction effects for each grouping variable estimated in separate models, including all base terms as covariates. 95% confidence intervals generated by simulation. See Table **??** in the appendix for details of model estimates.

the effect of PB across racial groups is more obvious.

5 Discussion

These results point to a robust positive impact from voting in PB on the likelihood of someone turning out to vote in a future regular election. This effect persists across different types of participants, including across racial groups, genders, age groups, and people from more or less wealthy or well-educated neighborhoods. These results present a confirmation of participant's own direct accounts of increased motivation to get involved in other kinds of citizen activity following their experiences with PB (Lerner and Schugurensky 2006, Gilman 2016) .

Participatory budgeting processes create an intersection between 'less political' volunteering and civic engagement and formal political behaviors. With its community focus and emphasis on public co-design of projects with direct impact on local neighborhoods, PB is uniquely positioned to bring in residents who may not typically vote or engage in formal electoral politics. Intentional outreach efforts combined with low registration barriers help to bring more people easily into the process. This is the first project that has been able to directly investigate whether the PB voting process itself would be able to translate this local civic motivation into expanded political behavior. At least in New York City, it looks like PB has been able to act as an effective bridge between local civic projects and voter activation.

While this analysis has been conducted making the most of available data, these results are of course subject to a range of limitations. One of the major limitations is that we cannot generalize these effects beyond the population of voters likely to vote in PB. Participatory Budgeting has so far been an opt-in process and this analysis should not be extrapolated to claim that requiring PB participation would necessarily result in increased voter turnout. That being said, our sample of PB voters include many individuals without a history of voting who do begin voting following their participation in PB, lending some support to expectations that the civic learning in PB can support an increased propensity to vote even among those who do not have a previous history of voting.

Another caveat on our results is its limitation to PB voters in New York City, and specifically to districts in New York City that had robust data collection practices early in their PB processes. We know that there is diversity in the implementation of PB across New York districts, and across cities that have chosen to implement PB (Johnson 2017). Are the successful data collection practices of these districts indicative of distinct procedures that may shape the voters' experiences and affect our estimates of PB? This is likely. It iss clear from the wide array of case studies and the more limited comparative work (e.g. Wampler 2008) that experiences of PB processes can be highly variable. The results of the analyses presented here must be interpreted in a context that understands that every process that calls itself 'PB' may not be expected to have such strong results. It will be important to replicate and extend the analysis presented here in other settings where individual voter data is available. Even if these results are limited to particular settings or implementations of the PB process, the positive impacts reported here can be informative for helping to identify elements of PB design that are most likely to lead to individual transformations in political behavior.