Supplementary Materials for:

Plutopopulism: Wealth and Trump's Financial Base

Sean Kates[1] Eric Manning[2] Tali Mendelberg[3] Omar Wasow[4]

November 27, 2024

A Matching

A.1 Record Linkage Procedure

In this subsection, we cover the main aspects of our record linkage procedure we use to match individuals within and across our datasets. The procedure is similar to those described in recent work by Yoder (2020) and Hall and Yoder (2022), though with several notable differences. The full technical specifications of each linkage model are available to interested readers in a separate online appendix.

Unlike existing work (Bonica 2023) that identifies unique donors in FEC data, our procedure targets the identification of unique donor–addresses. We take this same approach for property owners in the CoreLogic data. This allows subsequent linkage across each of our datasets at the person–address level. We adopt an approach for identifying person–addresses that allows for two records to be partial matches across more than one field (e.g., partial matches on name *and* address), a more flexible approach than that described by Bouton et al. (2022). By constructing our own person–address identifiers in the FEC and CoreLogic data, we also establish consistency in our approach across datasets.

To create these identifiers, we enforce fairly strict criteria for partial matches to avoid incorrectly pooling together the assets of distinct individuals, while also allowing for cross-state matches (given that some mailing addresses in both the CoreLogic and L2 data are in states different than their corresponding parcel or residential addresses). The latter allows for the aggregation of interstate land portfolios for donors and non-donors alike, as well as the identification of donors who report contributions from mailing addresses *and/or* P.O. boxes. Matches *across* years are determined solely by the stable donor (FEC), owner (CoreLogic), or voter (L2) identifiers in each of the constituent datasets. We treat the matches generated by this procedure as edges in a graph of person–property nodes. We further populate this network with matches drawn from the co-occurrences of 1) parcel addresses and their corresponding mailing addresses for individuals listed as owners in the property records, and 2) residential addresses and their corresponding mailing addresses for registered voters.

We then identify and label each of the graph's connected components (Linacre et al. 2022), which we treat as identifiers for unique *individuals*—the level of observation of our final dataset. Where multiple person–addresses are connected in this graph, we aggregate that individual's contributions and property portfolios across each associated address. When a voter registration record is included in a cluster, we consider that individual a registered voter and include the available demographic covariates from the voter file.[5]

For example, suppose there is a parcel owner with name X at address A who reports mailing address B. We therefore know with certainty that person–addresses X-A and X-B are the same underlying person. Suppose we then observe a registered voter with name X at address B with mailing address C. We again know with certainty that X-B and X-C are the same person. Suppose we also observe a donor X-C. Our procedure enables us to resolve all person–addresses in this example (X-A, X-B, and X-C) to the same underlying individual—a registered voter, donor, and property owner.

There are two aspects of this procedure worth discussing further. First, we only assign a property's assessed value to an individual's property portfolio if they are listed on that property's tax record as an owner. For example, if a married couple resides at a property, but only one person is listed on the record (as is

[1]University of Pennsylvania, skates@sas.upenn.edu, ORCID ID: 0000-0003-4315-685X
[2]Princeton University, ericmm@princeton.edu, ORCID ID: 0009-0008-0234-4284
[3]Princeton University, talim@princeton.edu, ORCID ID: 0000-0002-4494-7541
[4]UC Berkeley, owasow@berkeley.edu, ORCID ID: 0000-0002-1104-4610

[5]When an individual is identified via this procedure as matching to more than one registered voter record, we retain the most recent registrant from this set.

sometimes the case), we assign the total assessed value to that one person. (In the case where both are listed, we assign the total assessed value to each.) This fails to ascribe property wealth to any additional family or non-family members who may reside at the address. While conceptually it may be correct to assign the total value to each family member, that raises a variety of technical considerations that have no principled answer. To name a few: Would this value be assigned only to those who share the same last name? If so, that may include family members who may not inherit the property. If not, we may incorrectly assign value to a renter. Vice versa, what about family members who do *not* share the same last name? (There is no way to identify familial relations in our datasets beyond the use of names.) In general, by conservatively assigning property wealth only to those listed on the property record itself, we potentially *underestimate* the wealth of some individuals in our dataset, and exclude others entirely (by failing to identify them as owning *any* property). The former would bias observed relationships between wealth and contribution behavior toward zero, and we have no reason to believe the magnitude of this bias in our dataset would vary over time. Likewise, the latter would reduce absolute, but not relative estimates of donor participation—though we rely entirely on relative estimates in our findings.

The second key aspect worth more discussion is the strictness of our matching approach. We require a (fuzzy) match on both name *and* address for us to link any two name–address pairs within or across datasets. One implication of this requirement is that we may fail to match donations from individuals who give solely from the address of a property which they do not own (for example, a donor who contributes from their workplace). This would yield a lower overall donor match rate to our other datasets, though this is only a concern if match rates vary by campaign. A second implication is that we may fail to match properties owned by the same person if there is no information (from mailing address co-occurrences in the L2 or CoreLogic datasets) that allows us to resolve these two properties to the same individual. This would yield total portfolio value underestimates for such (wealthier) individuals, consistently *attenuating* observed relationships between wealth and contribution behavior, therefore not threatening our main findings.

An alternative approach to linkage at the name–address level is given in Giraud-Carrier et al. (2015) and underlies observational data in Magleby, Goodliffe, and Olsen (2018), who identify unique donors in FEC data by blocking on metropolitan statistical areas (MSAs) (with a ZIP-code distance penalty) and matching on name only. We disfavor this approach for our purposes. The connected components algorithm by which we resolve links into individuals is prone to over-clustering. Namely, only one false positive link between two individuals' clusters—no matter how dense—would collapse these clusters, creating one observation in our dataset assigned to both individuals' properties and/or donations. The problem would be especially acute for name-only linkage in the presence of common names and large geographic areas (where the likelihood of more than one individual with any given name is high), generating mechanically (and potentially falsely) an exponential relationship between wealth and contribution outcomes, driven by those in large MSAs.

A.2 Comparison to Existing Dataset Merges

In this section, we compare our matching procedures to the most common approach in the political science literature using administrative datasets (e.g., Yoder 2020; 2021). Following the procedure outlined in Yoder (2020), which uses the **R** package `fastLink` (Enamorado, Fifield, and Imai 2019) to match L2 and CoreLogic records in California circa 2018, we begin by subsetting the California state L2 and CoreLogic parcel files to the nearest vintage (2016) included in the panel dataset. We subset to residential addresses listed in the voter file, and to residential site addresses in CoreLogic.

We then block only on ZIP code, as in Yoder (2020) (though a significant departure from the more flexible blocking rules used to merge the dataset described elsewhere). We randomly select 20 ZIP codes with at least 1,000 registrants in the voter file and implement `fastLink`'s probabilistic linkage model within each ZIP code, following the specifications in Yoder (2020), which includes first name, last name, and street address as linkage fields, using default thresholds for Jaro-Winkler string distance measures and a posterior probability cutoff of 0.85.

Across this stratified random subset, we find a match rate of 47.29% at the posterior probability threshold of 0.85. The estimated false positive rate is low (0.79%), though the false negative rate is estimated as 40.56%. This is similar to Yoder (2020)'s overall estimated match rate in the same state-year (50.44%), though the estimated false negative rate reported here is higher (compared to 19% in Yoder (2020)). Yoder (2020) notes that this match rate is close to the U.S. Census estimate of California's homeownership rate (50.44%). On its face, then, it appears that the matching procedure used to generate the dataset (in which we find in our 2020 vintage that 32.8% of California registrants match to at least one property) yields a lower match rate than the ZIP blocking `fastLink` procedure — evidence in favor of the latter.

However, clerical inspection of the matches generated by this approach suggest a high false positive rate (contrary to the model's summary statistics) that can be lowered with additional processing, and one which (expectedly) is higher among matching pairs with lower posterior match probabilities. For example, a significant proportion (7.5%) of the matches estimated by this approach occur between observations that are estimated as full matches on last name and address, but which fall below the "partial match" threshold for first name similarity (an average posterior probability of 90.3%, close to the 0.85 posterior threshold). Clerical inspection of these observations shows that nearly all are false positives.

Further, consider that with an estimated false negative rate of 19% among the matched units and a ground-truth homeownership rate of approximately 50%, the false positive rate would need to be roughly equal to 19% in order to achieve an overall estimated homeownership rate from the matching procedure equal to the ground truth, assuming that the false positive and false negative rates generated by the Fellegi-Sunter model are accurate (i.e., match rate $-$ FPR $+$ FNR $\approx$ ground truth). However, the fastLink output from California approximates the true homeownership rate as approximately 69% — which is too high. As a result, we conclude that the Fellegi-Sunter model generates either a) overestimates of the false negative rate, and/or b) underestimates of the false positive rate. Given that clerical review suggests the false positive rate is much higher, we identify this as the main source of the discrepancy.

A large number of false positives is particularly concerning given that unique individuals are generated in the dataset by detecting the connected components of the network of person–property nodes (where edges are probabilistically- or deterministically-generated matches). A small number of false positive edges, then, between nodes that in reality represent two distinct individuals can generate a significant amount of overclustering (i.e., labelling these two individuals as the same person). Given the high weight that probabilistic record linkage approaches place on the address field, this is more often than not likely to occur between individuals with similar street addresses, yielding likely overclustering in condominiums, city centers, etc.

Further still, consider evidence suggesting that a significant proportion of individuals who report homeownership in surveys are not actually listed on the parcel record associated with their residence (approximately 1/3, according to the only known study of this phenomenon) (Zhang et al. 2022). As a back-of-the-envelope guide, suppose that the ground-truth match rate between L2 registrants and in-state parcel records in California is truly somewhere closer to $\approx (2/3) \times 50\% = 33\%$ than to 50%. The match rate between the L2 voter file and CoreLogic parcel records in the dataset herein (32.8% in California in 2020) is likely evidence of a superior matching procedure than one which generates a match rate of approximately 50%.

A.3 Match Rates by Year for All Dataset Pairs

Table A1: Number of Observations and % Overlap by Dataset, Year

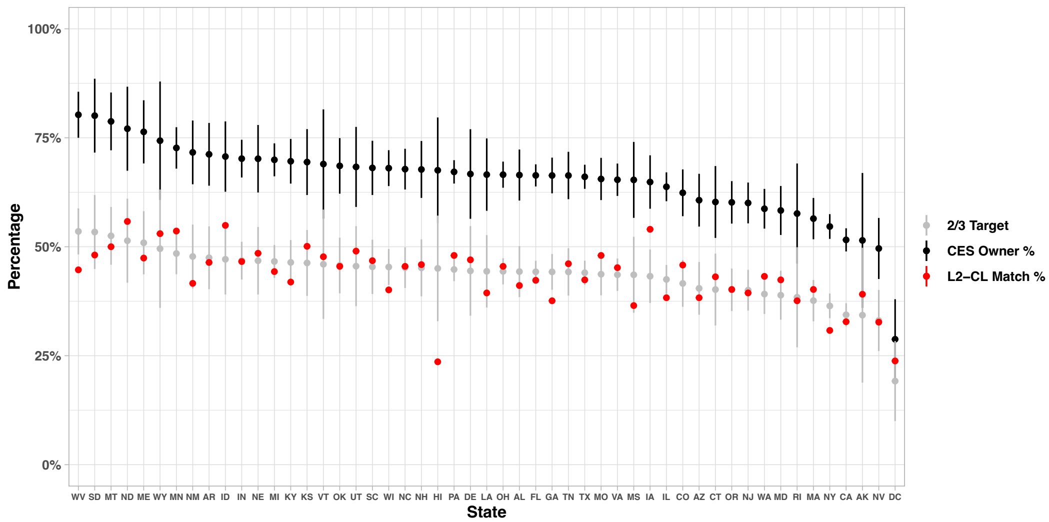
Year	Source	Count	% in L2	% in CoreLogic	% in FEC
2012	CL	129,322,735	52.58		1.24
2012	FEC	3,184,657	53.89	50.42	
2012	L2	150,586,721		45.16	1.14
2016	CL	132,169,078	60.28		1.60
2016	FEC	5,096,556	56.79	41.61	
2016	L2	180,541,518		44.13	1.60
2020	CL	140,094,900	61.17		5.47
2020	FEC	18,896,113	60.07	40.55	
2020	L2	206,859,833		41.42	5.49

Note: 2012 data rely on earliest available L2 snapshots, circa 2013–2014. FEC rows report N identified unique donors.

A.4 Voter–Homeowner Match Rate by State

The approximate 2/3 target for our matching procedure (in light grey) is motivated by findings in Zhang et al. (2022), who find that approximately 1/3 of survey respondents who report owning their residence are not listed on that property's deed. At the national level, 64% of registered voters reported home ownership, and 73% of homeowners were registered. Our sample-wide match rates of 43.8% of registered voter-years in the L2 dataset to at least one owned property in the CoreLogic dataset and 58.1% of property owners to a voter registration align with these expectations once we again account for that 1/3.

Figure A1: State-level match rates between L2 and CoreLogic datasets correlate strongly with homeownership rates among registered voters (L2–CoreLogic merge).



A.5 Included Committees

Presidential campaigns raise money through other channels besides their principal campaign committees (PCCs). After clinching their party's nomination, candidates typically establish at least one joint fundraising committee (JFC) with their party through which they raise, spend, and distribute their money in coordination with the national and state parties. Beginning in the 2012 election, campaigns are also supported by one or more super PACs that raise and spend money independently and do so entirely (or almost entirely) in support of that campaign (i.e., "single-candidate" super PACs). Capturing individuals' financial support for presidential campaigns requires the inclusion of each of these. Table A2 reports the committees (and their types) we use to construct our contribution outcomes in this paper, along with the related candidates and the contribution date ranges we include for each.

For principal campaign committees and single-candidate super PACs, we include all contributions. (For incumbents, we include contributions made in the full four year period between presidential elections.) For Priorities USA in 2020, a Democrat-supporting super PAC that persists between cycles, we include contributions made only after Biden clinched his party's nomination, as that entity raised and spent money against Trump before Biden's primary victory. We only include super PACs that raised and spent at least $20 million, entirely (or nearly entirely) in support of the presidential nominee alone.

For committee–years starred Table A2, we drop itemized contributions from individuals that have missing transaction codes. We find that nearly all of these transactions, and few others, are recorded as transactions where the donor's cycle-to-date total had yet to exceed $200 (the threshold for mandatory itemization). No other presidential campaign in our dataset reports these contributions. To enable the comparison of campaigns within and across cycles in our dataset, we therefore drop these transactions (and include them in the "unitemized" column in Table A3).

Republican 2012 primary opponents included in Figure 6 are those that raise at least $10 million from individuals: Cain, Gingrich, Ron Paul, Perry, and Santorum. In 2016: Jeb Bush, Carson, Cruz, Fiorina, Kasich, Rand Paul, and Rubio.

Table A2: Included FEC Committees

Cycle	Candidate	Type	Committee ID	Committee Name	Range
2012	Obama	PCC	C00431445	Obama for America	2009-01-01 to 2012-12-31
2012	Obama	JFC	C00494740	Obama Victory Fund 2012	Ending 2012-12-31
2012	Obama	SP	C00495861	Priorities USA Action	2011-01-01 to 2012-12-31
2012	Romney	PCC	C00431171	Romney for President Inc	Ending 2012-12-31
2012	Romney	JFC	C00518282	Romney Victory Inc	Ending 2012-12-31
2012	Romney	SP	C00490045	Restore Our Future Inc	Ending 2012-12-31
2016	Clinton	PCC	C00575795	Hillary for America	2015-01-01 to 2012-12-31
2016	Clinton	JFC	C00586537	Hillary Victory Fund	Ending 2016-12-31
2016	Clinton	SP	C00495861	Priorities USA Action	2015-01-01 to 2012-12-31
2016	Trump	PCC	C00580100*	Donald J Trump for President Inc	Ending 2016-12-31
2016	Trump	JFC	C00618389	Trump Victory	Ending 2016-12-31
2016	Trump	JFC	C00618371*	Trump Make America Great Again Committee	Ending 2016-12-31
2016	Trump	SP	C00637512	America First Action Inc	Ending 2016-12-31
2016	Trump	SP	C00756882	Preserve America PAC	Ending 2016-12-31
2016	Trump	SP	C00574533	Future45	Ending 2016-12-31
2016	Trump	SP	C00618876	Rebuilding America Now	Ending 2016-12-31
2016	Trump	SP	C00608489*	Great America PAC	Ending 2016-12-31
2020	Biden	PCC	C00703975	Biden for President	Ending 2020-12-31
2020	Biden	JFC	C00744946	Biden Victory Fund	Ending 2020-12-31
2020	Biden	SP	C00495861	Priorities USA Action	2020-04-24 to 2020-12-31
2020	Biden	SP	C00701888	Unite the Country	Ending 2020-12-31
2020	Biden	SP	C00669259	FF PAC	Ending 2020-12-31
2020	Biden	SP	C00532705	Independence USA PAC	Ending 2020-12-31
2020	Biden	SP	C00492140	AB PAC	Ending 2020-12-31
2020	Trump		All as above	2017-01-01 to 2020-12-31	

Including FEC lines 11AI for non-PCC, 17A for PCC, all receipt types unless ID marked with * (excludes transactions with null `receipt_tp`).

A.6 Wealth Distribution Validations

A.6.1 Comparison to L2 Economic Measures

In this subsection, we compare our estimate of individuals' property wealth to modeled estimates of primary residence estimated home value as found in L2 commercial voter data. This serves as a check that our matching procedure is generating links between property owners (CoreLogic) and registrants (L2) that, in general, are accurate, as well as showing that our measure (which aggregates *assessed* values) is positively correlated with L2's estimated market value measure.

We still favor our estimates for several reasons. First, L2 estimates only the home value of a registrant's listed residence, under-estimating wealth for those owning multiple properties. Second, the estimation strategy used for L2's estimated home value variable is proprietary. As a result, we cannot validate this measure absent additional data. Finally, L2 imputes home value to entire households, including those not on the deed.

Figure A2: Comparison of L2 modeled home value estimate for primary residences (x-axis) and our estimate for that individual's (assessment-based) property wealth (y-axis). Blue line is a fitted GAM. Dotted line is the 45 degree line. Based on $N = 500,000$ sample. Marginal distributions shown top and right.

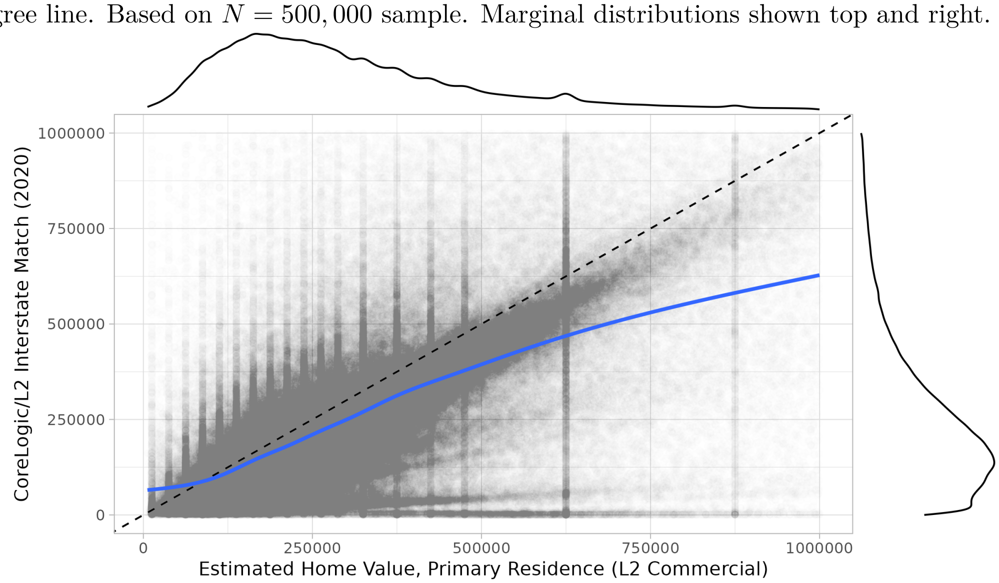
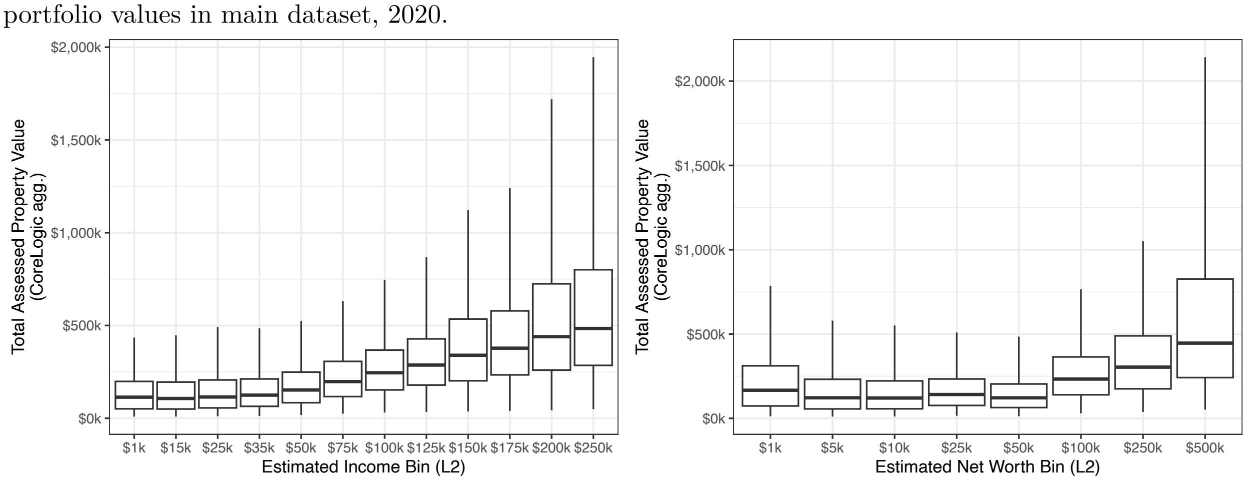


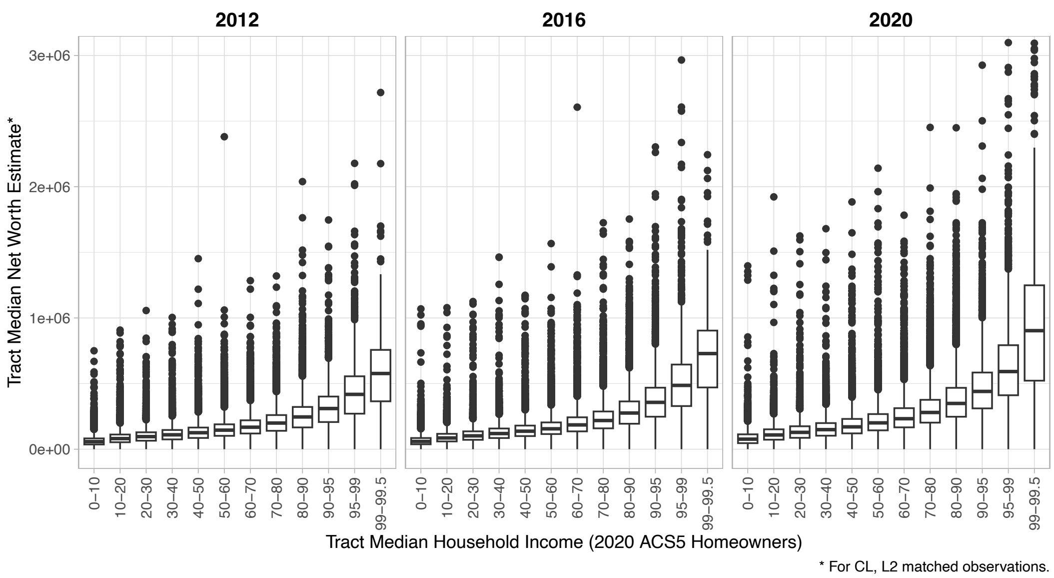
Figure A3: Correlation between individual-level (binned) L2 economic measures and our assessed property portfolio values in main dataset, 2020.



A.6.2 Comparison to Census Measures

In this subsection, we focus on how well our estimates for individuals track known distributions of *income* (net worth is not measured by the Census Bureau) at low levels of geographic aggregation. We do so by aggregating our individual-level wealth measure to the Census tract level and calculating its median. We compare this to estimates of tract-level median income among homeowners from the U.S. Census Bureau's 2020 American Community Survey (ACS). While neither is a perfect proxy for unobservable total net worth, both should be highly correlated. The results are shown in Figure A4.

Figure A4: Relationship between tract-level median household income and tract-level median property portfolio value in L2-CoreLogic intersection (main analysis dataset). Property owners are assigned to the tract where they our registered to vote.



The boxplots here represent the distribution of *tracts* and their median net worth, binned at the same tract's level of median household income according to the ACS. We expect to see tracts in the highest ACS bins also rank higher in median net worth by our estimates and that is in fact what we observe, across all three cycles. This presents strong evidence that, again, our measure is capturing what it is meant to: namely, relative wealth.

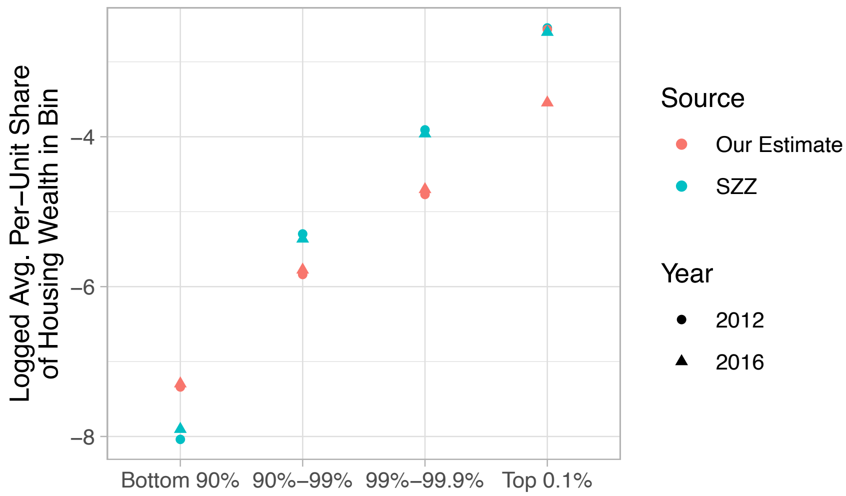
A.6.3 Comparison to Smith, Zidar, and Zwick (2023) Distributions

We compare the distribution of our wealth measure to summary statistics on housing and total wealth from aggregated individual-level tax filings, sourced from the U.S. Treasury and reported in Smith, Zidar, and Zwick (2022). Smith, Zidar, and Zwick (2022) find that Americans' housing wealth is monotonically increasing in overall wealth. This further justifies our use of rank-ordering in our measure construction. We need not accurately estimate absolute total wealth values for individuals, only their rank ordering between one another.

Similarly, the percentage of total housing wealth belonging to each percentile bin, as shown in Figure A5, is similar as measured by our measure and Smith, Zidar, and Zwick (2022)'s, even though our bins are constructed using housing wealth itself, whereas Smith, Zidar, and Zwick (2022) uses total wealth.

For our purposes, housing wealth derives from the assessed value of the properties owned by an individual, then ranked against the population; their measure estimates property equity, so that it can be compared to other investments as a portion of a broader investment portfolio. This may explain the difference in relative ordering between our measure Smith, Zidar, and Zwick's among the bottom 90% (as compared to other wealth bins), as the non-wealthy are more likely to have mortgages.

Figure A5: Proportion of per-unit housing wealth held in each bin. For comparability across bins, a unit is defined here as each 0.1% share of the population.



A.7 Omitted Campaign Dollars

Across all of our analyses, we omit money given to and spent by dark money groups, such as 501(c)(4)s, as we cannot systematically identify the individuals that contribute to these groups (see, e.g., Oklobdzija 2023). This threatens our analysis if we systematically underestimate contribution outcomes for certain individuals (here, almost exclusively the most wealthy) or certain campaigns. With respect to the former, dark money groups generally rely on a very small number of donors for funding, and therefore are unlikely to affect our results where the outcome reflects the *number* of donors. With respect to the latter, our finding that Trump underperforms among wealthy donors in 2016 would be threatened if these donors attempted to hide their support for Trump in 2016 (above any hidden support for Romney in 2012) by channeling previously-observed dollars into dark money groups. However, far more dark money was spent against Obama/for Romney in 2012 than against Clinton/for Trump in 2016, an aggregate trend that precludes this possibility.[6]

Our dataset excludes two additional sources of funds. The first is money itemized by campaigns, but unmatched to any individual we identify in both the L2 and CoreLogic data. Rates of inclusion of total itemized dollars by campaign are presented in the fourth column of TableA3 (incl. JFCs and super PACs). Rates of inclusion of total dollars from individual contributors (i.e., including the lump sums that committees report raising from donors who give below the disclosure thresholld) are reported in the rightmost column. The inclusion of itemized campaign dollars varies somewhat by campaign. We are able to match a low of 38.4% (Romney's) itemized dollars from all sources (PCCs, JFCs, and super PACs), and a high of 46% (both Biden and Trump 2020) itemized dollars. Because itemized contributions are by definition larger than unitemized contributions (and therefore originate from wealthier individuals on average), we are therefore likely to understate Romney's performance with wealthier donors. Trump's underperformance with this group, then, may be even worse than we find in our main results.

Our overall match rates, which account for unitemized dollars, are more consistent, with the exception of Trump 2016 and Trump 2020, for whom we match only 20.5% and 24.6% of dollars, respectively. However, these low rates are clearly driven by Trump's relatively high rates of unitemized totals—small contributions which come disproportionately from the less wealthy. Therefore, our results are likely to *understate* Trump's performance with the less wealthy. However, because he raised so little overall in 2016, this does not alter our interpretation of our finding that Trump's activation of the non-wealthy did not begin in earnest until the 2020 cycle.

In Tables A4 and A5 below, we report aggregate match rates by dataset, by contribution size for each presidential campaign. Our main analyses include CoreLogic–L2 merged data only, for which the match rates are given in the rightmost column for each campaign.

[6] https://www.opensecrets.org/dark-money/top-elections?cycle=2016

Table A3: Rates of Inclusion for Relevant Campaign Totals (incl. JFCs and Super PACs)

Campaign	Unitemized	Itemized	%Match of Item.	%Match of Raised
Obama 12	329,746,400	738,106,694	43.9	30.4
Romney 12	125,601,382	819,279,665	38.4	33.3
Clinton 16	202,569,767	880,005,611	41.4	33.6
Trump 16	295,704,465	270,936,543	42.9	20.5
Biden 20	472,001,140	1,180,599,830	46.4	33.1
Trump 20	1,014,116,247	1,140,052,575	46.5	24.6

Unadjusted for inflation for comparability to FEC summary statistics.

In general, match rates are somewhat higher for Republican candidates compared to each candidate's contemporaneous opponent (except in 2020). These are largely driven by higher match rates between contributions and the CoreLogic data, suggesting a higher homeownership rate among Republican donors. The magnitude of this difference, however, is far smaller than the magnitude of the partisan differences we find in our main results.

Within each party, match rates are uniformly higher for Trump 2016 than for Romney 2012 for each range of contribution amounts. If anything, this ought to attenuate the significant fundraising disadvantages we find for Trump as compared to Romney. Likewise, we find that Trump 2020's contribution match rates are *lower* than Romney's (except among the largest contributions, which likely originate disproportionately from the wealthiest Americans), which would again only *attenuate* the fundraising advantages we observe for Trump 2020 as compared to Romney.

Table A4: Match Rates, by Candidacy and Contribution Amount (Democrats)

	Obama			Clinton			Biden		
Contribution Amount	CL	L2	CL-L2	CL	L2	CL-L2	CL	L2	CL-L2
($0, $50]	0.48	0.58	0.41	0.49	0.68	0.46	0.49	0.69	0.47
($50, $100]	0.51	0.60	0.44	0.54	0.70	0.50	0.51	0.70	0.49
($100, $200]	0.54	0.62	0.46	0.55	0.71	0.51	0.54	0.71	0.51
($200, $500]	0.55	0.62	0.46	0.56	0.70	0.51	0.55	0.71	0.52
($500, $1000]	0.57	0.62	0.48	0.57	0.70	0.53	0.57	0.72	0.54
($1000, max]	0.58	0.62	0.49	0.57	0.66	0.51	0.58	0.72	0.55

Unadjusted for inflation for comparability given fixed FEC itemization threshold.

Match rate also generally increase with contribution size (except among the largest contributions, perhaps because they come disproportionately from donors intentionally reporting non-residential addresses). This suggests a general pattern of greater homeownership among (likely wealthier) donors giving larger sums. However, yet again, we find the lowest match rates in groups where we observe the best performance in our data—evidence against concerns that variation in match rates drive our findings. For example, the lowest match rates in our dataset are found among the largest contributions to Romney and the smallest contributions to Democrats (and to Trump 2020)—all cases where we observe better performance among the non-wealthy than for Romney or Trump 2016.

Table A5: Match Rates, by Candidacy and Contribution Amount (Republicans)

	Romney			Trump 2016			Trump 2020		
Contribution Amount	CL	L2	CL-L2	CL	L2	CL-L2	CL	L2	CL-L2
($\$0, \$50$]	0.63	0.67	0.54	0.61	0.73	0.57	0.53	0.69	0.49
($\$50, \$100$]	0.66	0.69	0.56	0.63	0.73	0.58	0.55	0.70	0.51
($\$100, \$200$]	0.66	0.67	0.56	0.63	0.71	0.57	0.56	0.70	0.53
($\$200, \$500$]	0.64	0.63	0.53	0.63	0.68	0.57	0.57	0.70	0.54
($\$500, \$1000$]	0.64	0.62	0.52	0.63	0.67	0.57	0.59	0.70	0.54
($\$1000, max$]	0.59	0.55	0.46	0.61	0.63	0.54	0.59	0.68	0.54

Unadjusted for inflation for comparability given fixed FEC itemization threshold.

While the match rate between campaign contributions and parcel records is approximately 50-65% depending on the campaign and contribution amount, it does not necessarily follow that only this proportion of contributors are homeowners. When making contributions, donors may report non-residential addresses, or report their name and address in such a way that we are unable to successfully match them to a name and address pair in (either) dataset. However, assuming that a very high proportion of donors are registered to vote, we can estimate that the percentage of contributions that could reasonably match to a residential record is approximately the "L2" match rate column reported in Tables A4 and A5 (approximately 60-70%). The fact that nearly all contributions we successfully match to a parcel record (50-65%) also match to a registered voter record (comparing the "CL" columns in Tables A4 and A5 to the "CL-L2" columns) suggests a very high rate of homeownership among donors making itemized contributions (i.e., 50-65% divided by 60-70%).

B Descriptive Statistics and Variable Details

B.1 Demographic Profiles of Campaign Donors

Our earliest available state-level snapshots begin in late 2013. As a result, for 2012 our voter file data is restricted only to those individuals who are recorded as having registered to vote prior to the 2012 general election. L2 contains voter-level, self-reported demographic information from state administrative sources (in the case of age, sex, and in some states ethnorace and party registration (Hersh 2015)) and imputed, merged, or proprietary measures of income, net worth, homeownership status, estimated value of the primary residence, and, if not self-reported, ethnorace and partisan identification).

We use individual-level covariates for observations that merged to at least one row in the nationwide voter file from L2, Inc. These covariates include ethnicity, gender, and education:

- *Education* - As do other analyses of L2 data (e.g. Enamorado and Imai, 2019; Lalani et al., 2020; and Bonica et al., 2021), we use this variable to explore demographic variation. L2's description states that L2 uses both modeled and self-reported data. L2 separates education into 11 categories, including "Unknown" and both "Extremely Likely" and "Likely" versions of 5 ordered classifications: Less than HS Diploma, HS Diploma, Some College, Bachelor's Degree, and Graduate Degree. We dichotomize the variable into "College Degree" and "No College Degree". The proprietary nature of this variable is certainly a limitation. However, our main analyses do not use this variable (or other modeled L2 data). Our key predictor is our own wealth measure. In addition, by dichotomizing the variable into a binary (college or non-college), we reduce measurement error in the original variable.

- *Ethnorace* - Ethnorace is self-reported by registrants in several states (Hersh 2015), and modeled by L2 based on individual-level attributes and geography where not self-reported.

- *Gender* - Gender is coded by L2 as a binary variable, with data drawn from state voter registration files where possible, and modeled on first name otherwise.

B.2 Demographic Distributions for Matched Donors

In Tables A6 to A9, we provide the demographic breakdowns of matched donations for each presidential candidate in our sample, by wealth (Bottom 90% and Top 10%), for donors and dollars.[7] In nearly every election cycle, the Republican candidate raised more of their money from men (both in number and amount) and from white donors (number and amount), and less from donors with college degrees (number and amount). This is true both in the top 10% of our distribution, and in the bottom 90%.

Table A6: Matched Donor Number - Bottom 90%

	Obama '12	Romney '12	Clinton '16	Trump '16	Biden '20	Trump '20
Gender						
Male	46.8%	70.1%	39.1%	67.7%	46.5%	59.9%
Female	53.2%	29.9%	60.9%	32.3%	53.5%	40.1%
Ethnorace						
White	77.6%	93.3%	81.9%	92.6%	84.2%	89.2%
Black	13.7%	0.7%	6.4%	0.8%	5.8%	0.9%
Hispanic	4.2%	3.0%	5.9%	3.4%	4.8%	5.3%
Asian	2.4%	1.4%	3.0%	1.4%	2.7%	2.6%
Other	2.2%	1.6%	2.9%	1.8%	2.4%	2.0%
Education						
CollegeDegree	72.0%	71.7%	72.4%	63.3%	70.8%	57.3%
NoCollegeDegree	28.0%	28.3%	27.6%	36.7%	29.2%	42.7%
Education Within White Donors						
CollegeWhite	72.6%	71.4%	72.5%	63.1%	70.8%	57.1%
NoCollegeWhite	27.4%	28.6%	27.5%	36.9%	29.2%	42.9%

Table A7: Matched Donor Amounts - Bottom 90%

	Obama '12	Romney '12	Clinton '16	Trump '16	Biden '20	Trump '20
Gender						
Male	51.8%	71.0%	44.7%	69.1%	55.1%	63.6%
Female	48.2%	29.0%	55.3%	30.9%	44.9%	36.4%
Ethnorace						
White	79.2%	93.2%	81.2%	91.9%	85.6%	89.4%
Black	11.5%	0.6%	5.1%	0.7%	4.3%	0.8%
Hispanic	3.7%	2.6%	5.6%	2.8%	3.8%	4.7%
Asian	2.9%	1.7%	3.8%	2.9%	3.3%	3.0%
Other	2.8%	1.8%	4.3%	1.7%	3.0%	2.2%
Education						
CollegeDegree	77.2%	77.2%	77.6%	66.7%	76.0%	62.3%
NoCollegeDegree	22.8%	22.8%	22.4%	33.3%	24.0%	37.7%
Education Within White Donors						
CollegeWhite	77.2%	76.7%	77.6%	65.6%	75.6%	62.2%
NoCollegeWhite	22.8%	23.3%	22.4%	34.4%	24.4%	37.8%

[7] Observations with unknown values of a given attribute are dropped within that attribute.

This better contextualizes the wealth similarities between Obama in 2012 and Trump in 2020. While these two candidates both drew support from the non-wealthy, that support came from different race, gender and education groups: Non-wealthy Trump donors were far more likely than non-wealthy Obama donors to be male, white, and lack college degrees (Tables A6 and A7).

Table A8: Matched Donor Number - Top 10%

	Obama '12	Romney '12	Clinton '16	Trump '16	Biden '20	Trump '20
Gender						
Male	51.9%	70.0%	44.4%	67.8%	52.1%	63.3%
Female	48.1%	30.0%	55.6%	32.2%	47.9%	36.7%
Ethnorace						
White	82.3%	92.2%	82.2%	91.5%	82.7%	86.0%
Black	5.8%	0.5%	3.2%	0.4%	2.9%	0.4%
Hispanic	3.7%	3.2%	4.6%	3.5%	4.3%	5.1%
Asian	4.6%	2.0%	5.8%	2.2%	6.2%	5.4%
Other	3.6%	2.1%	4.3%	2.4%	3.8%	3.1%
Education						
CollegeDegree	84.3%	81.5%	84.9%	75.7%	83.4%	73.4%
NoCollegeDegree	15.7%	18.5%	15.1%	24.3%	16.6%	26.6%
Education Within White Donors						
CollegeWhite	84.2%	81.3%	84.8%	75.6%	83.3%	73.1%
NoCollegeWhite	15.8%	18.7%	15.2%	24.4%	16.7%	26.9%

Table A9: Matched Donor Amounts - Top 10%

	Obama '12	Romney '12	Clinton '16	Trump '16	Biden '20	Trump '20
Gender						
Male	59.1%	72.2%	55.0%	73.3%	62.0%	69.2%
Female	40.9%	27.8%	45.0%	26.7%	38.0%	30.8%
Ethnorace						
White	82.7%	92.7%	77.0%	93.2%	82.6%	88.5%
Black	4.4%	0.4%	2.1%	0.3%	2.0%	0.7%
Hispanic	3.1%	2.9%	4.2%	2.6%	2.8%	3.8%
Asian	5.2%	1.7%	5.8%	1.8%	6.7%	3.9%
Other	4.7%	2.4%	11.1%	2.2%	5.9%	3.1%
Education						
CollegeDegree	85.6%	83.3%	85.3%	73.5%	87.2%	77.1%
NoCollegeDegree	14.4%	16.7%	14.7%	26.5%	12.8%	22.9%
Education Within White Donors						
CollegeWhite	85.5%	83.5%	83.1%	75.3%	86.8%	76.6%
NoCollegeWhite	14.5%	16.5%	16.9%	24.7%	13.2%	23.4%

In Table A10, we offer a comparison to Table A6, drawn from the CES election surveys in 2012, 2016, and 2020. We use these surveys as a validation of the demographic estimates from our sample, made up exclusively of homeowners - a restriction not found in the CES. Importantly, the patterns we observe in our own data are replicated here, including both the traditional partisan splits and the specific differences across the Obama 2012 and Trump 2020 candidacies.

Table A10: CES Bottom 90 Distribution of Self-Reported Donors

	Obama '12	Romney '12	Clinton '16	Trump '16	Biden '16	Trump '20
Gender						
Male	44.92%	65.83%	44.88%	58.77%	43.68%	57.20%
Female	55.08%	34.17%	55.12%	41.23%	56.32%	42.80%
Ethnorace						
White	75.41%	90.17%	76.50%	87.36%	78.55%	85.63%
Black	15.03%	0.52%	11.03%	1.45%	8.71%	1.66%
Hispanic	4.17%	2.52%	6.41%	4.11%	6.00%	5.78%
Asian	0.78%	0.55%	1.88%	1.29%	2.15%	0.95%
Other	4.61%	6.24%	4.18%	5.78%	4.59%	5.97%
Education						
CollegeDegree	62.95%	54.55%	63.76%	46.43%	67.97%	49.04%
NoCollegeDegree	37.05%	45.45%	36.24%	53.57%	32.03%	50.96%
Education Within White Donors						
CollegeWhite	64.82%	54.23%	64.07%	44.62%	68.01%	46.31%
NoCollegeWhite	35.18%	45.77%	35.93%	55.38%	31.99%	53.69%

C Supplemental Figures: Results

C.1 Full Distributions of Donation and of Donor Retention by Wealth

Figure A6 presents the full distribution of contribution rates and per-capita dollars, broken down into deciles, with the highest decile further broken down into more granular bins. This is the disaggregated version of Figure 1 from the main paper. Figure A7 does the same for presidential donor retention rates, serving as the disaggregated version of Figure 5 from the main paper. Finally, Figure A8 disaggregates the bottom deciles to replicate Figure 6 from the main paper.

C.2 Main Results Using L2 Measures

Our use of wealth rank derived from property values in our main analyses requires that we restrict focus to property owners. While this is not a substantial threat to our behavioral findings for the wealthiest Americans, nearly all (more than 95%) of whom own property,[8] it potentially threatens our findings for the less wealthy, particularly if campaigns rely on contributions from non-homeowners at different rates. We find some evidence of this in Appendix A.7. Further, though we *rank order* property values, it may be that even *ranked* property wealth fails to properly order the relative *total* wealth of individuals. If this is the case for certain groups that are represented in the campaigns' donorates at different rates, this would threaten our analysis.

To address these concerns, we show that two of our main findings (Figure 1 and Figure 5) replicate using alternative measures of wealth, across most of the distribution of wealth for which we are able to observe variation in rank ordering with these alternative measures.

Namely, Figure A9 replicates Figure 1 using (A) L2's (binned) income estimates for *all registrants for whom L2 provides such an estimate*, and (B) L2's (binned) net worth estimates for the same. That is, the denominator for panels that present L2's measure is registered voters, rather than registered property owners as in Figure 1.

[8] https://www.federalreserve.gov/econres/scfindex.htm

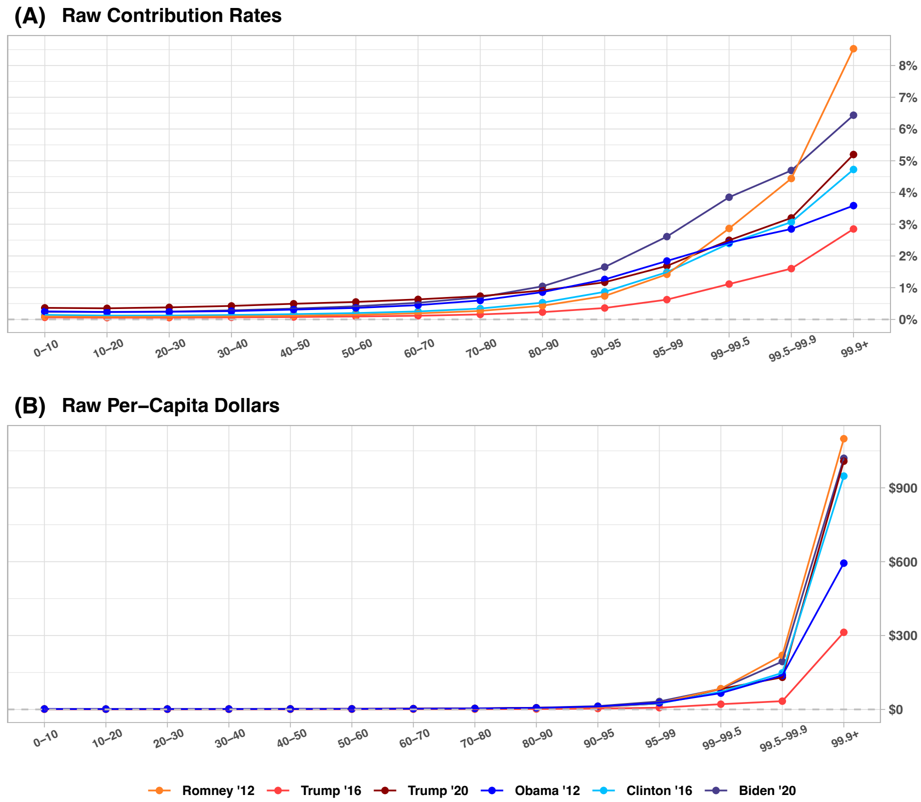


Figure A6: Association between percentile wealth bins (x-axis) and: (A) contribution rate, (B) per capita contribution amount.

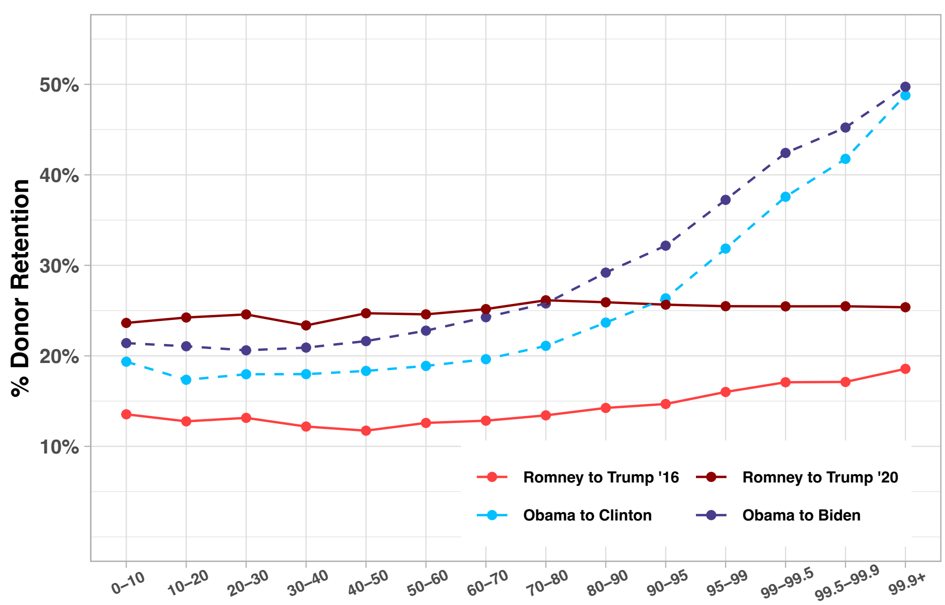


Figure A7: Cross-cycle donor retention by party, wealth bins.

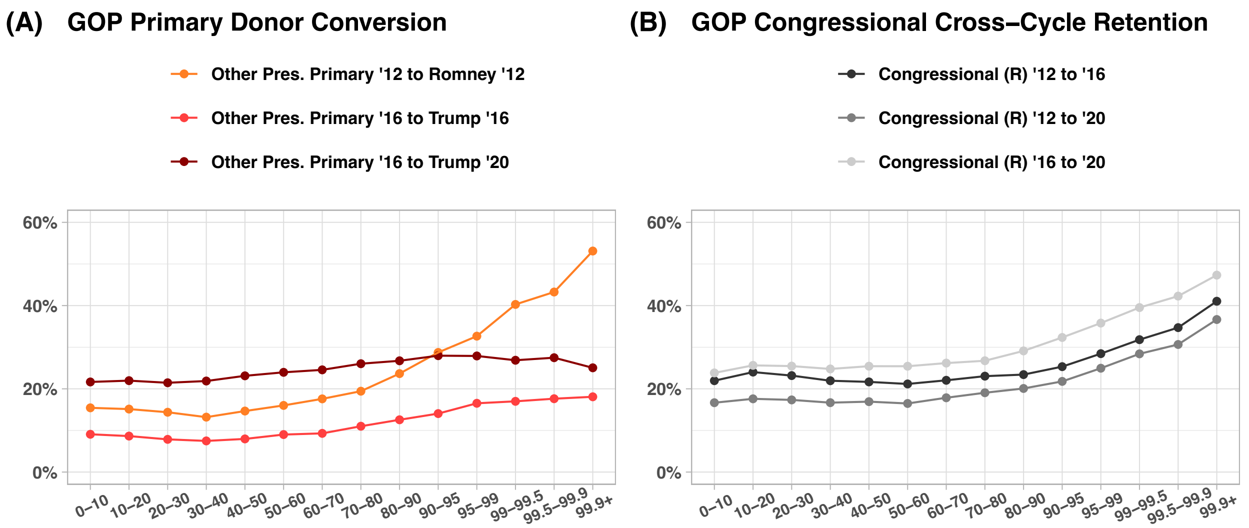
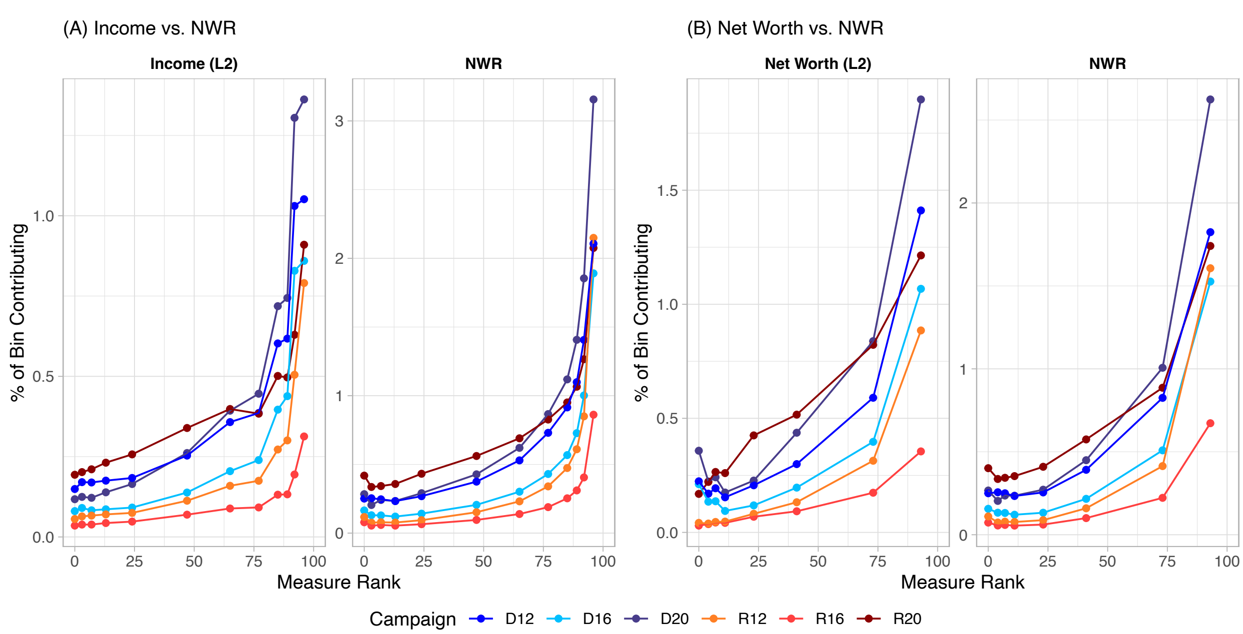


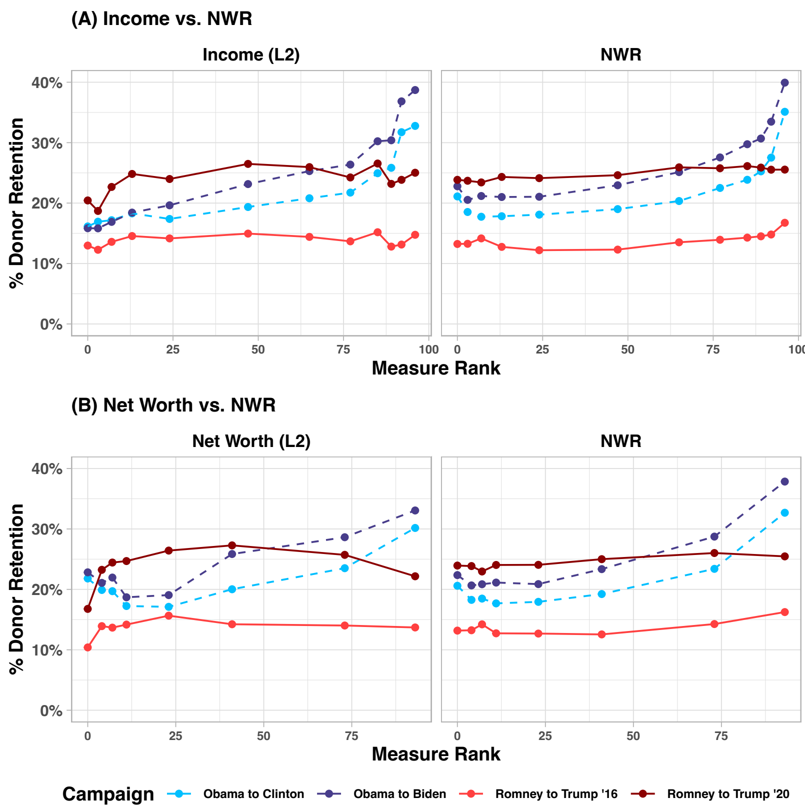
Figure A8: Panel (A): rates at which donors to Republican primary losers convert to general election nominees. Panel (B): cross-cycle donor retention rates for Republican congressional donors.

Figure A9: Fig.1a using L2 income bins and comparable National Wealth Rank bins



For comparability between results generated from this measure and our National Wealth Rank (NWR) presented in the main body, we calculate the percentage of registrants in each of the L2 measures' bins, and re-bin our NWR measure accordingly. For example, approximately 25% of registrants with an income estimate in the L2 data fall in the 5th greatest bin, which represents roughly the 20th to the 45th percentile (represented by the span between the 5th and 6th points on each of the lines in the leftmost panels of (A)). We then re-cut our NWR measure at these percentiles and present comparable contribution rates in adjoining panels, by party. (y-axes vary between results using our measure and L2's measure because contribution rates are, in general, higher among homeowners – but we are only concerned with relative relationships within and across campaigns, rather than absolute levels of support).

Figure A10: Fig.5 using L2 income bins and comparable National Wealth Rank bins



L2's binned income and net worth estimates are both top-coded, such that approximately the top 5% are binned together in each of the two measures. Therefore, this exercise only allows us to compare results between measures for the top 5% and down.

We find that estimates of relative contribution rates across wealth and between campaigns are substantively similar between measures. For example, across both measures as compared to our own, Trump's 2020 activation of donors strictly dominates both his prior 2016 performance, and Romney's 2012 performance. (Of course, here we lose the ability to examine variation *within* the top 5% wealthiest individuals, so campaign rank orders in the top-most bin vary from our main results.) Similarly, among Democrats, Biden in 2020 outperforms both prior Democratic general election candidates among the wealthiest donors, with Clinton in 2016 activating donors at the lowest rates across the range of latent affluence.

Figure A10 presents results from Figure 5 using the same strategy as above. Again, though we cannot disaggregate donor behavior *within* the top 5% of wealth using L2's top-coded measures, results are substantively similar between these measures (left panels of (A) and (B)) and our own (right). This time, retention outcomes are similar across measures both in terms of *relative* ordering within and between campaigns, *and absolute* levels. After conditioning on prior donor status, the expansion of the analysis group to include non-property owners no longer substantially increases only the denominator.

Table A11: Percent of Donors Giving in One, Two, or Three Cycles, by Party

Party	2012	2016	2020	% of Donors	n
Republican Donors	Donor	Donor	Donor	2%	26,216
	Donor	Donor	Non-Donor	1%	17,898
	Donor	Non-Donor	Donor	3%	45,179
	Donor	Non-Donor	Non-Donor	20%	283,629
	Non-Donor	Donor	Donor	4%	64,379
	Non-Donor	Donor	Non-Donor	7%	101,412
	Non-Donor	Non-Donor	Donor	62%	894,529
Democratic Donors	Donor	Donor	Donor	4%	72,810
	Donor	Donor	Non-Donor	3%	66,414
	Donor	Non-Donor	Donor	4%	85,187
	Donor	Non-Donor	Non-Donor	28%	541,265
	Non-Donor	Donor	Donor	5%	94,920
	Non-Donor	Donor	Non-Donor	15%	285,095
	Non-Donor	Non-Donor	Donor	41%	789,221

Each of these exercises suggests that our focus on property owners does not obscure a different overall story that would be told if we were able to include non-property owners as well in our analysis. They also confirm the value of having a measure that is continuous across a wide range, allowing for more granular analysis at the very highest reaches of the range.

C.3 Donor Flows by Party

Table A11 tracks donors from 2012 to 2016 to 2020, separately by party. The data include any donor making an itemized contribution (above $200) to at least one campaign. It includes individuals who did not match to L2 and/or CoreLogic.[9] We exclude cross-party donations, as they are very few. As the table shows, new Republican donors—–those who gave to Trump in 2020 and not Romney—compose the largest of all the cohorts in *either* party.

C.4 Dollar Retention by Campaign

As noted in the main text, Trump heavily under-performed among Romney's wealthiest donors (Figure 5). Here, in Figure A11, we replicate this under-performance with cross-cycle *dollars* retained. As Figure A11 shows, in 2016, Trump retained Romney donors' dollars at a steady rate across wealth bins. While he recovered somewhat in 2020, he did so evenly across donor wealth. For Democrats, unlike Trump, the rate sharply increases with wealth and is comparable from 2016 to 2020.

C.5 Supplemental Cross-Sectional Demographic Results

C.6 Donor Retention by Demographics

[9]Excluding unmatched donors yields similar results.

Figure A11: Cross-cycle dollar retention by party, wealth bins.

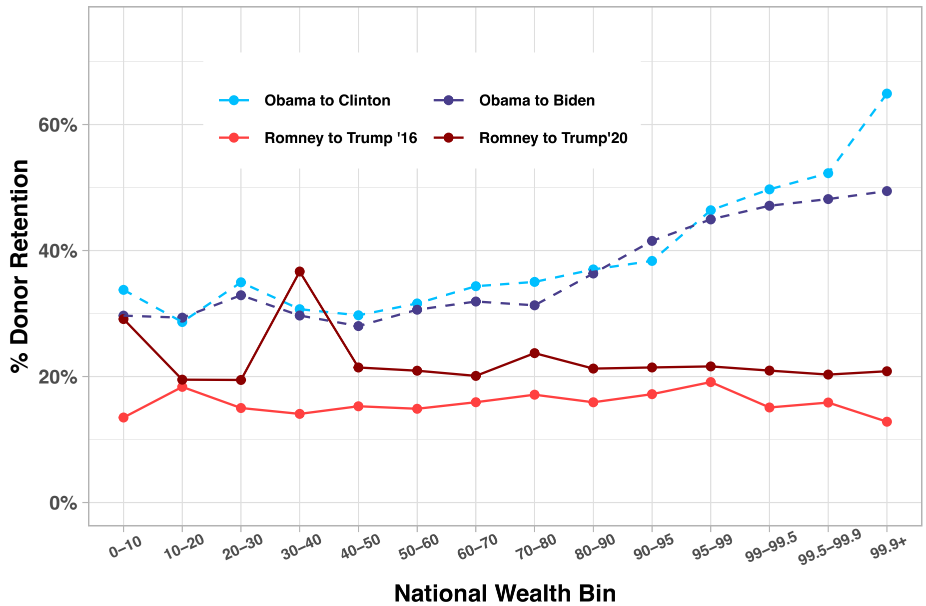


Figure A12: Percent change in donors as compared to 2012 copartisan candidate by sex, education and ethnoracial subgroups.

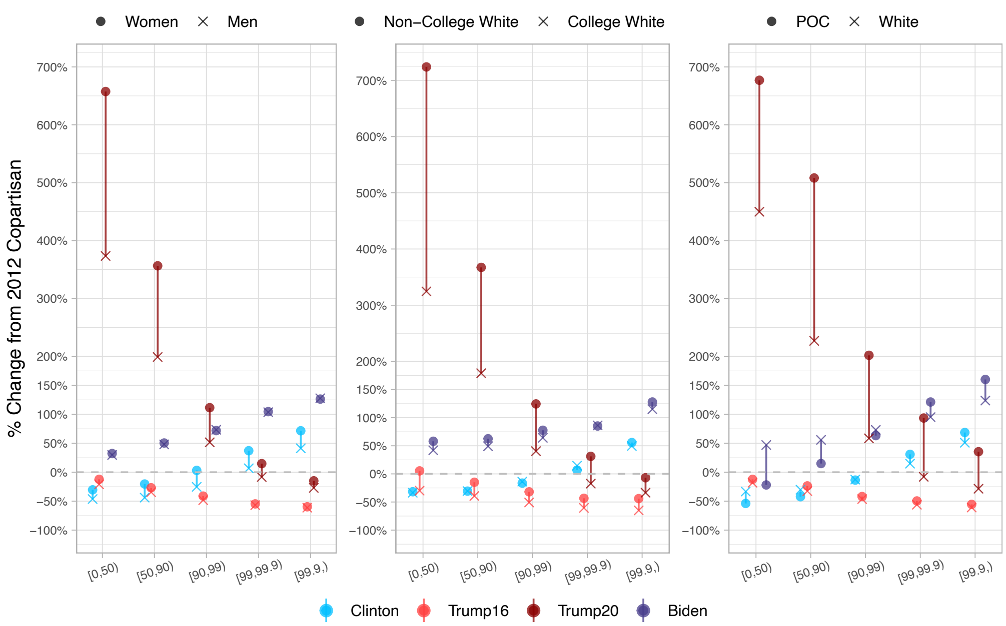


Figure A13: Change in Dollars Raised by Ethnoracial Group and Wealth Bin

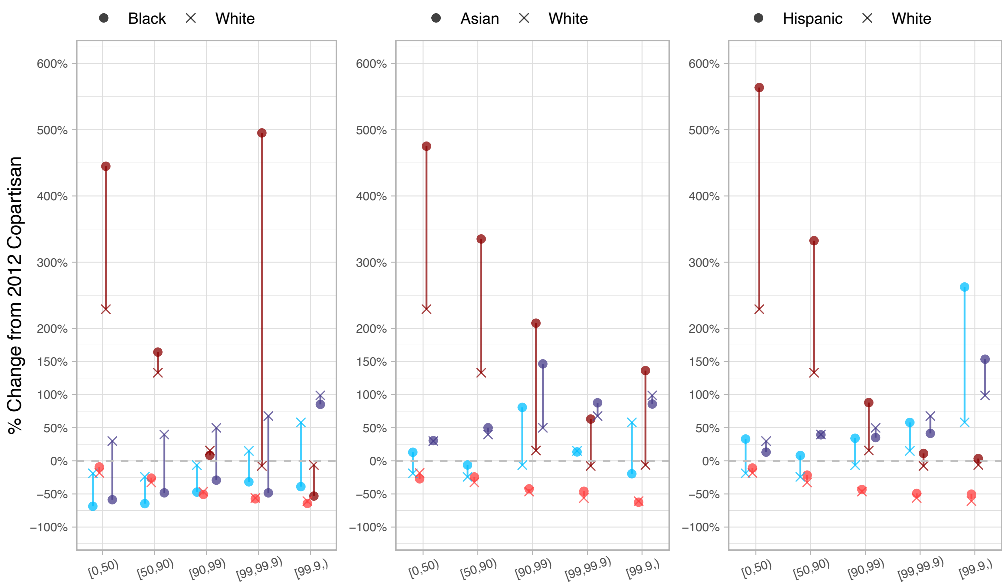


Figure A14: Change in N Donors by Ethnoracial Group, Wealth

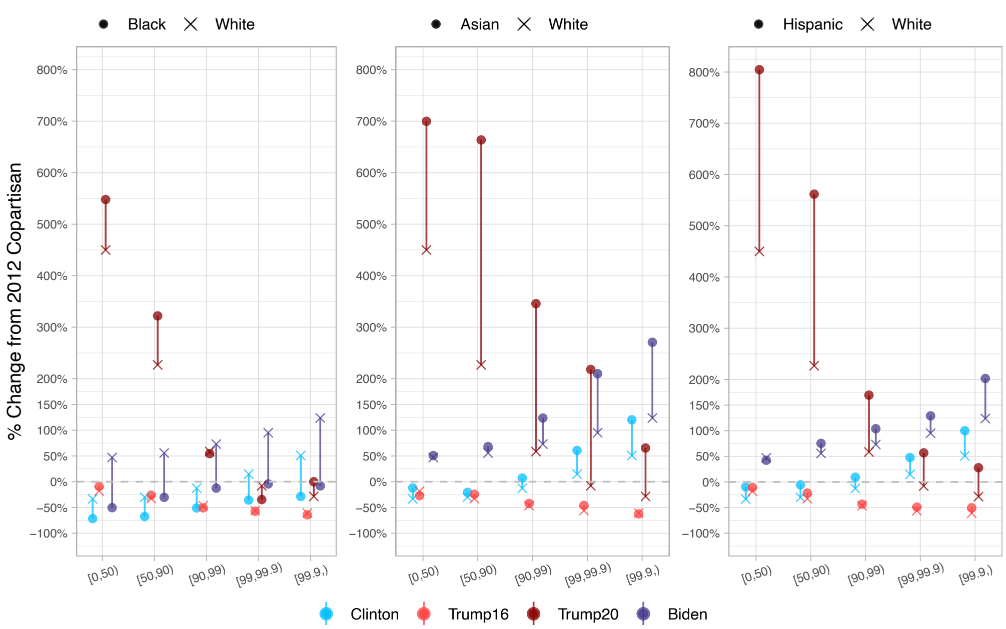


Figure A15: Differential donor retention by education, ethnorace and sex, and by wealth bin, across four cross-cycle general election campaign pairs, 2012–2016 and 2012–2020.

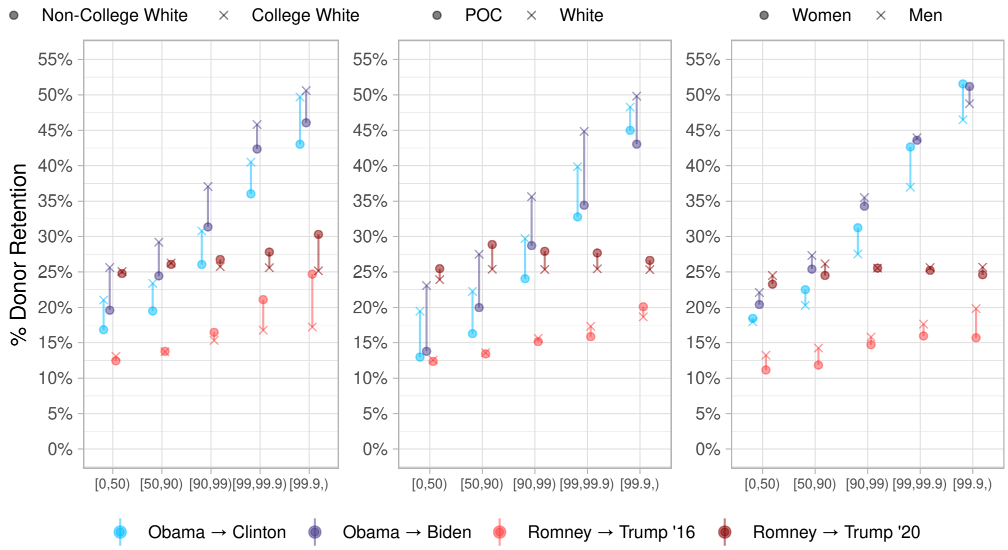
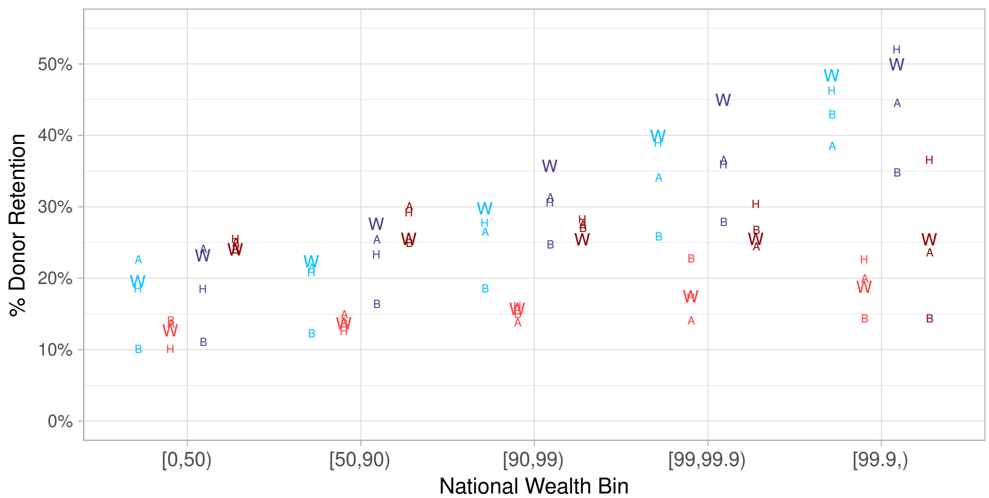


Figure A16: Prior donor retention by wealth and ethnorace. (A = Asian, B = Black, H = Hispanic, W = White.)



References

Bonica, Adam. 2023. *Database on Ideology, Money in Politics, and Elections v.3.*

Bouton, Laurent, Julia Cagé, Edgard Dewitte, and Vincent Pons. 2022. "Small Campaign Donors." *National Bureau of Economic Research Working Paper* 30050.

Enamorado, Ted, Benjamin Fifield, and Kosuke Imai. 2019. "Using a Probabilistic Model to Assist Merging of Large-Scale Administrative Records." *American Political Science Review* 113 (May): 353–371.

Giraud-Carrier, C., J. Goodliffe, B. M. Jones, and S. Cueva. 2015. "Effective Record Linkage for Mining Campaign Contribution Data." *Knowledge and Information Systems* 45 (November): 389–416.

Hall, Andrew B., and Jesse Yoder. 2022. "Does Homeownership Influence Political Behavior? Evidence from Administrative Data." *The Journal of Politics* 84 (January): 351–366.

Hersh, Eitan D. 2015. *Hacking the Electorate: How Campaigns Perceive Voters.* Cambridge University Press.

Linacre, Robin, Sam Lindsay, Theodore Manassis, Zoe Slade, Tom Hepworth, Ross Kennedy, and Andrew Bond. 2022. "Splink: Free Software for Probabilistic Record Linkage at Scale." *International Journal of Population Data Science* 7 (August).

Magleby, David B, Jay Goodliffe, and Joseph A Olsen. 2018. *Who Donates in Campaigns? The Importance of Message, Messenger, Medium, and Structure.* Cambridge University Press.

Oklobdzija, Stan. 2023. "Dark Parties: Unveiling Nonparty Communities in American Political Campaigns." *American Political Science Review*: 1–22.

Smith, Matthew, Owen Zidar, and Eric Zwick. 2022. "Top Wealth in America: New Estimates Under Heterogeneous Returns." *The Quarterly Journal of Economics* 138 (08): 515-573.

Yoder, Jesse. 2020. "Does Property Ownership Lead to Participation in Local Politics? Evidence from Property Records and Meeting Minutes." *American Political Science Review* 114 (November): 1213–1229.

Yoder, Jesse. 2021. "Essays on Elections and Voting in the United States." Ph.D. diss. Stanford University.

Zhang, Shiyu, James Wagner, Elisabeth Gerber, and Jeffrey Morenoff. 2022. "What Parcel Tax Records Tell Us About Homeownership Measurement in Surveys." *Survey Research Methods* 16 (August): 133–145.