

Hidden Donors: The Censoring Problem in U.S. Federal Campaign Finance Data

R. Michael Alvarez

Jonathan N. Katz

Seo-young Silvia Kim¹

California Institute of Technology

January 8, 2019

¹Kim is a Ph.D. candidate in the Division of Humanities and Social Science, Caltech, and is the corresponding author (sskim@caltech.edu); Alvarez is Professor of Political Science, Division of Humanities and Social Sciences, Caltech; Katz is Kay Sugahara Professor of Social Science and Statistic, Division of Humanities and Social Sciences, Caltech. An earlier version of this paper was presented as a poster at the Polmeth 2017 and MPSA 2018; we thank conference participants for their comments and suggestions.

Abstract

Inferences about individual campaign contributors are limited by how the Federal Election Commission (FEC) collects and reports data. Only transactions that exceed a cycle-to-date total of \$200 are individually disclosed, so the contribution histories of many donors are unobserved. We contrast visible donors and “hidden donors,” i.e., small donors who become invisible due to censoring and routinely ignored in existing research. We use the Sanders presidential campaign in 2016, whose unique campaign structure received money only through an intermediary committee, which is governed by stricter disclosure statutes, allowing us to study hidden donors. We show that there are many hidden donors, and they are more likely to be social and ethnic minorities. Hidden donors start giving relatively later, with contributions concentrated around early primaries, suggesting different interests or contribution incentives. We conclude that because of this censoring problem, the donor population may be quite different than found in previous research.

keywords: campaign finance | 2016 election | data censoring

The United States has established one of the world’s most comprehensive federal campaign disclosure processes. The Federal Election Commission (FEC), the federal entity which collects and disseminates these data, produces data that are transparent, accessible, and up-to-date. Consequently, there has been an enormous amount of academic research on campaign finance in the U.S. for the last few decades—see, for example, the recent review by Dawood (2015). Yet as we argue below, the FEC’s data is not a complete record of all dollars given and received in federal elections, and so it is likely that our understanding of campaign finance in the United States is incomplete.

One key feature of disclosure regulations in the U.S. is that currently, each federal campaign committee only has to report to the FEC contributions from an individual *who has already given \$200 in aggregate to that campaign committee*, either within a year or a two-year election cycle according to the committee type. The \$200 threshold has been in place for decades, and many donors’ first few—in some cases all—contributions are censored.¹ We call this the *censoring problem* in campaign finance data, and the campaign contributors whose contributions are below this threshold we call *hidden donors*.

Although a crucial part of the data generating process, the censoring problem and hidden donors have not been adequately addressed in the campaign finance literature (Key et al., 1955; Barber, 2016; Magleby, Goodliffe and Olsen, 2018). There is, after all, little that could be done about the censoring, short of changing the law. However, in aggregate, hidden donors are a nontrivial force that the politicians must cater to when deciding ideological positions and campaign strategies. Moreover, the undisclosed contributing behavior may be interacting with other political behavior, such as issue formation, turnout, voting, and future transitions into more serious donors.

A preliminary analysis of the censoring problem is possible due to an unexpected turn of events in the 2016 presidential election brought on by the campaign of Bernard “Bernie”

¹Prior to 1989, however, the data entry threshold of the Federal Election Commission was set at \$500. See the FEC’s [Thirty Year Report](#) published in 2005. Also, because the \$200 restriction is with nominal dollars, the data has to be filtered using inflation adjustments.

Sanders. For the first time, we have a chance to observe *all* of a major presidential candidate’s individual contribution records, without any of them being censored. This is because the Sanders campaign chose to receive money only through an intermediary committee as earmarked contributions: these contributions are governed by a different, stricter set of regulations, resulting in full disclosure.

Through the Sanders presidential campaign, we can compare datasets in which the censoring problem is and is not present. While the “law of available data” has driven us to investigate visible donors almost exclusively, this paper, using the rare and valuable opportunity presented by the Sanders campaign, ventures into the seldom-explored world of “hidden” donors.

We show that hidden donors differ substantially from the visible donors. Hidden donors are more likely to be female, non-white, and younger. They may also have different political goals, as they display different donation strategies. They are more likely to donate later in the election cycle than visible donors, on average entering the race three months later, with their giving concentrated when the race is heated. Most importantly, there are seven times more hidden donors than visible donors—suggesting that we may be observing only the tip of the iceberg in individual donor behavior.

Past Research

Many papers that study individual campaign contributors do so using survey data. Brown, Powell and Wilcox (1995) surveyed campaign contributors in the 1988 presidential election to uncover the dynamics of “serious money” and how campaigns choose their solicitation strategies. Francia et al. (2003) provided an analysis of the motivations and demographics of individual contributors to the 1996 congressional elections. More recently, Barber, Canes-Wrone and Thrower (2017) conducted a survey of Senatorial contributors in the 2012 election, concluding that individuals respond to Senators’ policy positions.

However, such studies inevitably rely on the FEC records to sample respondents, meaning that they are restricted to “visible” contributors, who contributed \$200 or more to at least one candidate. This is because Title 11 of the Code of Federal Regulations declare in 11 CFR 104.3(4)(i) *Content of Reports* that only contributions that exceed \$200 in aggregate will be *itemized* and reported. The records that do not meet the threshold will be deemed *unitemized*, and reported only as an aggregate, thus discarding information on who gave how much on what date. We offer more legal details below.

The comparisons that studies often make (e.g., Francia et al. (2003)) is between a visible contributor and an average citizen, or between different types of visible contributors. These studies have not examined how hidden contributors differ from those who are visible, nor do they compare hidden contributors and non-contributors. Unfortunately, small, hidden donors could be very different from the donor population that we have witnessed and analyzed. Verba, Schlozman and Brady (1995) show that even among campaign contributors, mean dollars given to campaigns vary greatly by family income, and only when the family income exceeds \$75,000 dollars does the mean dollars given exceed \$200. This implies that lower-income donors might be more likely to be hidden, and their exclusion from survey studies of contributors could be problematic if their motives for contributing are systematically different from higher-income, visible donors.

It is not that political scientists are altogether unaware of the censoring problem in individual contribution records. Key et al. (1955) had already pointed out how we do not know about “little givers.” Francia et al. (2003) recognizes the threshold by labeling donations after the \$200 threshold as “significant” donations, and Heerwig (2016) by labeling the given donors as “elite” donors. Barber (2016), while discussing how his survey does not include donors who gave less than \$200 in a footnote, notes that his picture of individual donors’ motivations may be incomplete, if unitemized (hidden) donors have different motivations from those who are itemized (visible).

Although this censoring problem is well-known, and there is no remedy, short of changing

federal campaign finance law, or asking campaigns to voluntarily report every small donation. The paucity of donors in the adult American population still means that it might be difficult in a random sample of citizens to sample many of those making small donations. There have been a few papers using the American National Election Studies (ANES) surveys (Panagopoulos and Bergan, 2006; Johnson, 2013), but the sample size that report contributing is very small—for instance, in the 2008 election, only 9.9% of survey respondents reported to have contributed money to a specific candidate campaign out of 2,100 respondents.² This makes it difficult to compare and contrast small and large donors.

Despite this, the literature has worked to make stylized facts about the small, hidden donors. Most obviously, the hidden donors are relatively poor, because giving is related to disposable income. Verba, Schlozman and Brady (1995) show that percentage of family income contributed to political campaigns increases with the family income level, sharply rising at \$50,000 or more at the time of the survey. Other researchers directly surveyed small donors (Graf, 2006; Malbin, 2013; Wilcox, 2008) to conclude similarly that small donors are more economically representative of the general population, and discuss the normative concerns regarding expansion of small donor population. Graf (2006) also show the small and large donors' difference by gender, occupation, and education. Papers such as Johnson (2010) and Culbertson, McDonald and Robbins (2018) tackle the small donor problem by using aggregate amounts reported at the campaign level, concluding that small donors are linked to more ideologically extreme candidates, as has been previously suspected (Bonica, 2011). Christenson, Smidt and Panagopoulos (2014) use candidate-level regression to show that the web presence and Internet-drawn small donors propel candidates forward after early-primary victories.

The recentmost work that has systematically compared visible and hidden campaign contributors is Magleby, Goodliffe and Olsen (2018). They were able to cooperate with major presidential candidates of 2008 and 2012 to receive a random sample of contributors

²2008 Time Series Study, ANES, May 19, 2015 version, subsample of post-election survey respondents: question V085033.

who gave donations of all sizes. With Obama, Romney, and McCain contributors, they generally do not find important ideological differences between visible and hidden donors.³ In terms of their demographic profiles, they found that visible donors are older and wealthier than hidden donors. Hidden donors, meanwhile, were more likely to have been solicited to contribute online.

Magleby, Goodliffe and Olsen (2018) makes it clear that campaigns prioritize large donors over the small, especially due to changes introduced by the Bipartisan Campaign Reform Act of 2002. However, appeals to small donors are becoming more prevalent in American campaigns, especially in recent elections as the Internet is becoming central to campaign fundraising (Malbin, 2013; Karpf, 2013) and lowering the cost of participation. The small donor population has been subjected to intense media scrutiny starting 2008 and especially during the recentmost 2016 and 2018 elections (Mider, 2015; Hamburger, 2016; Goldmacher, 2016; Burns, Shorey and Patel, 2018). They are also likely to increase as campaigns realize their potential and importance, especially having a set of voters that have not maxed out their individual contribution limits, and hence can be re-solicited (Nelson, 2010). In that case, it is crucial to understand who these small donors are and what they are doing, in order to understand how the campaigns must be perceiving and catering to their needs to raise more money.

This paper contributes to the literature on small donors by analyzing an entire donor population for a presidential candidate and on an individual-level data. While our data does not have the more in-depth questions and breadth that are available in surveys of donors like that of Magleby, Goodliffe and Olsen (2018) or Graf (2006), we are not limited to a few thousand survey respondents—we can bring the entire population of visible and hidden Sanders contributors to study. This allows us to show the counterfactual scenarios and compare various summary statistics, including or excluding hidden donors. Hence we help document the biases that the censored data brings to our analyses of campaign contributions.

³They refer to this comparison as itemized vs. small donors.

The Censoring Problem in Campaign Finance Data

Using the U.S. federal campaign finance data is much more complex than it appears. Disclosure statutes from the Code of Federal Regulations have many important and complex details that dictate how campaigns should collect and report the data. Moreover, these regulations, rooted in the The Federal Election Campaign Act of 1971 (FECA, Pub.L. 92–225), have changed numerous times over the years. Therefore, it is crucial to understand the legal details of campaign finance in the U.S. to get an accurate picture of even the simplest descriptive statistics.

In this Section we state the censoring problem in more detail. We show examples of hypothetical contributor behavior, and demonstrate what happens when the researcher tries to infer contributor behavior using just the censored data. We also show what makes the Sanders 2016 campaign different by explaining the legal details entailing intermediary committees and the disclosure of earmarked contributions.

Code of Federal Regulations, Title 11

The Federal Election Commission administers federal campaign finance law, under Title 11 of the Code of Federal Regulations. A federal campaign committee that meets the conditions will be registered with the FEC and will regularly file reports that disclose funds that are raised and spent. 11 CFR 104.3 *Contents of Reports* dictates this, and how the information on receipts are censored is stated in 11 CFR 104.3(4)(i):⁴

(4)(i) Each person, other than any political committee, who makes a contribution to the reporting political committee during the reporting period, whose contribution or contributions aggregate in excess of \$200 per calendar year (or per

⁴We immediately see that aside from the arbitrary \$200, there are two additional problems: one is that the \$200 is in nominal dollars, unadjusted for inflation, and the other is that whether a campaign committee is authorized or not—which is, in the campaign finance jargon, another name for candidate-affiliated committee, especially the principal campaign committees. For an exact definition, see 11 CFR 9032.1. We largely avoid the first problem as we only use a single cycle’s observation, and we will for the moment ignore the second problem.

election cycle in the case of an authorized committee), together with the date of receipt and amount of any such contributions, except that the reporting political committee may elect to report such information for contributors of lesser amount(s) on a separate schedule;

When a contribution must be recorded by meeting this threshold, this contribution is ‘itemized.’ Unitemized contributions are aggregated into a lump-sum and reported as a single number, and thus does not present the same amount of details as an itemized contribution. Naturally, the donors whose contributions never meet the threshold become entirely absent from the filed paperwork.

As aforementioned, intermediary committees are the primary organization that we utilize to investigate the censoring effect on the data. 11 CFR 110.6 *Earmarked Contributions* provide details as to what they are, and what they should disclose. While earmarked contributions are money designated to a clearly identified candidate/committee, intermediary committees are “anyone who receives and forwards an earmarked contribution to a candidate or candidate’s authorized committee.”⁵ The following excerpt from the FEC Candidate Guidebook summarizes the special disclosure requirements:

A political committee that serves as a conduit of an earmarked contribution must disclose the earmarked contribution, regardless of amount, on two separate reports: the committee’s next regularly scheduled FEC report, and a special transmittal report sent to the recipient authorized committee. 110.6(c)(1).

As we can see, intermediary committees have stricter disclosure requirements than other campaign committees. They must report all contributions, not just those that cross the \$200 threshold—see Appendix A for details. The issue, which we present with some detailed hypothetical examples in Appendix B, is that for campaigns that use the typical contributions process, we will not observe contributors to that campaign who donate less than \$200.

⁵While conduit or intermediary committees are interchangeably used and the FEC seems to prefer the former, we use the terminology ‘intermediary committees.’

The Sanders Campaign, 2016

The Sanders campaign was unique in many ways, including his approach to campaign finance. In his 2016 presidential campaign, one of the important issues for Sanders was his stance against the influence of “big” money and special interests. He claimed early on that his campaign would not be allied with any super PACs (Lee, 2016; Qiu, 2015). In addition, the Sanders campaign was reported by the media in late 2015 to have organized only seven traditional fundraisers, while the Clinton campaign had by then already organized more than 110 (Associated Press, 2015). In March 2016 it was further reported that the campaign had only two more traditional fundraisers (Gaudiano, 2016).

This distaste for conventional fundraising meant that most of Sanders’ donations were digitally processed. His website noted on April 30, 2016 that 94% of its contributions were made online.⁶⁷ His fundraising appeals were mostly digital, using emails, texts, and social media (Corasaniti, 2016), and his website pointed to the ActBlue contributing page. The campaign used no other online platforms and had no offline fundraising staff. Hence, almost all individual donors who desired to contribute to Sanders would have had to donate through ActBlue, regardless of their wealth, connections, or intentions.

Because ActBlue is an intermediary committee, it is subject to the same regulations as explained in [Code of Federal Regulations, Title 11](#). We can track all Sanders donors, regardless of the size of their donations to his campaign, a first for any level of campaign. Regardless of whether they gave fifty cents or maxed out individual contribution limits, they

⁶The website address is <https://berniesanders.com/revolution/>. Our final estimate of how much Sanders campaign received online is 98.1% of his total individual contributions, which is an increased percentage from 94% reported in the website, the estimate which is calculated before mid-May 2016—hence the disparity is likely because the last months’ donations were made all online. That the 94% is calculated before mid-May can be inferred from the Wayback Machine’s snapshot of the website, the first snapshot of which is at May 21, 2016, and shows the 94% claim.

⁷5.1% of donors’ year-to-date contribution records do not match the actual sum of records, (1) potentially because they have not been sufficiently record linked due to the conservative linking of contribution records that constitute the same individual, (2) the committee inaccurately calculated the contributor year-to-date by mixing up different contributors, or (3) because these donors have given in the aforementioned few traditional fundraisers. The third possibility seems viable, since we find that 68.7% of those with record irregularities are visible. The second possibility is also high, because we often find that intermediate records of contribution sum year-to-date are grossly inaccurate. This speaks to the difficulty of working with campaign finance data.

would be captured in ActBlue reports. This means that the censoring problem disappears with respect to Sanders’ contributors, giving us an unprecedented opportunity to study both visible and hidden donors to a major presidential campaign. See Appendix C for details.

It is true that the data is restricted to a subset of presidential donors, who are ideologically liberal but attracted to a non-conventional, independent candidate. However, there are unmistakable advantages in utilizing the Sanders data. It is likely that we will never have another major candidate whose contributions are so transparently presented to the public and have all his donors exposed, both small and large.

Table 1 shows the proportion of unitemized contributions within all individual contributions for the four major presidential candidates of 2016, as well as House and Senate Democrats and Republicans in 2016. Again, this demonstrates that there are many hidden donors that we cannot observe, and that they play a substantial role in elections, especially in presidential races.

Data

The data is collected from the Federal Election Commission. While the data itself is public, there are several ways to collect the data. You can use the FEC FTP site to bulk download the individual contributions data, use third-party data, or build a database from scratch using the the OpenFEC API (<https://api.open.fec.gov/developers/>). While the bulk data is easy to download and robust to deduplication, this contains only the visible donors and excludes the small donor contributions by applying the \$200 filter. The third party data such as from the Center for Responsive Politics has some benefits, such as industry coding and standardization of occupation, but also misses intermediary contributions that we are looking for. Hence we built our own database using raw FEC data. Due to rate limit issues, this can take up to months using a single credential.

Using OpenFEC API, all the 2016 cycle’s raw individual contribution data was down-

loaded as text files in a 100-line batches, more than 1,133,000 in number and 165Gb in storage. They were then wrangled and record-linked to establish which donations were coming from the same individual. Most importantly, we use contributions to Bernie 2016, the Sanders presidential committee, and to ActBlue, and create a union of contribution records to Sanders while filtering out duplicate contributions, if any. This way, we can uncover donations ranging from a dollar to the \$2,700 individual limit, and determine the hidden donors. We perform a similar operation for other committees as well, as other Democratic campaign committees and PACs also extensively use ActBlue.

The campaign reports do not contain gender or racial/ethnic information per se, and only carry name, address, the money's destination, date, occupation, and employer. To buttress our analysis, we use contributor names and address geo-locations to infer gender and race/ethnicity. For gender inferences, we use the R package `gender` (Mullen, 2018; Blevins and Mullen, 2015), which uses first names and the Social Security Administration's yearly dataset of first names to infer the gender of the donor. If the donor has voluntarily entered a prefix such as a Mr., we override the gender inference with the gender inferred from the self-reported prefix. For race/ethnicity, we use the R package `wru` (Khanna, Imai and Jin, 2017; Imai and Khanna, 2016). That is, we utilize the Census Bureau's surname list and use Bayes rule to infer the race. In particular, we use a census block level inference, after geocoding each address using the Census Geocoder and obtaining the latitude and longitude of the address.

Who Are the Hidden Donors?

We define a Sanders contributor as anyone who donated to Sanders' presidential campaign committee during the 2016 presidential election cycle. The committee launched on April 30, 2015⁸ and the cycle ends on December 31, 2016.

⁸Statement of Organization, Bernie 2016. Available at <http://docquery.fec.gov/pdf/537/15031422537/15031422537.pdf>, last accessed November 19, 2018.

The total estimated number of Sanders contributors is over 2 million, including both visible and hidden donors—this demonstrates that he received a large number of contributions to his presidential bid from a large base of donors. However, among the 2 million Sanders donors, only 12.4% (250,352 out of 2,017,638) would have been visible if the Sanders campaign did not receive individual contributions only through intermediary committees. In other words, there were seven times more contributors than would have been actually visible. This simple fact shows the potential magnitude of the censoring problem in the U.S. campaign contributions data, and similar to the numbers reported in Magleby, Goodliffe and Olsen (2018), where they estimated that 82% of Romney donors and 88% of Obama donors in 2012 were small donors.

Figure 1 shows the donor base comparative to the population estimated by the 2016 American Community Survey, geographically. Naturally, we can see that the donor population is strong in Vermont, with the town of Newfane with the largest proportion of Sanders donors (32% of residents). 74% of the top 100 zip codes in Sanders-donor proportion are Vermont zip codes, with California and Massachusetts trailing respectively at 12% and 5%.

Visible vs. Hidden Donors: Demographics

Table 2 displays the demographic difference between visible and hidden Sanders contributors. A hidden contributor is more likely to be male and a racial or ethnic minority, whether black, Hispanic, or Asian. However, it is noticeable that while the percentages of black or an Asian contributors do not differ greatly between visible and hidden Sanders donors, a hidden contributor is much more likely to be Hispanic than a visible contributor.⁹ This is an interesting observation, and could reflect the fact that Hispanic voters are younger than other racial groups (Patten, 2016), and therefore unable to afford donating more than \$200.

Table 3 shows top 10 types of employment of all 2016 contributors and their percentage

⁹These are still statistically significant differences. We do not offer p- or t-statistics separately for any of our descriptive statistics. Because the sample size is so large, *every* difference that we present in this paper is statistically significant ($p < 0.001$).

for the subset of Sanders supporters. Noticeable is the presence of students in the hidden donor population while the likelihood of being retired is much higher for visible donors. This signals a difference in age groups, i.e., that the hidden donors are likely to be younger in age. In addition, while visible donors are more likely to report being attorneys and physicians, hidden donors are slightly more likely to be teachers. Considering that the average annual salary of teachers in public schools was \$56,383 in 2012¹⁰ while a physician’s lowest pay was \$189,000¹¹, this hints at wealth differentials, again, which is to be expected. However it is also interesting to note that the proportion visible donors reporting unemployment is actually slightly higher than among hidden donors.

Predicting Hidden Donor Status

In this subsection we estimate generalized additive models (GAMs) to predict hidden donor status, using state dummies, gender interacted with race, and smoothed covariates for wealth. Because we do not have individual-level wealth, we use zip-code level median household income and housing prices of 2015, both of which are fitted as smooth terms. We also generate an occupation variable by classifying the following: student, unemployed, homemaker, retired, and employed. The reason is that, relative to more specific occupational labels that we have shown above, these are less likely to be exaggerated or be ambiguous.¹²

Using the fast stable restricted maximum likelihood method (REML) of Wood (2011), Figure 2 the summed effects of parametric terms in our GAM model.¹³ A full summary and graphical representations of other effects such as state effects or occupational effects are in

¹⁰National Center for Education Statistics, [Digest of Education Statistics: 2013](#)

¹¹Forbes, [The Best- And Worst-Paying Jobs For Doctors](#), July 20, 2012.

¹²For instance, a donor declaring herself as ‘president’ may mean very different things, ranging being the head of a charitable organization to a more traditional corporate executive. The size and power of the organization they lead—and subsequently the donor’s socioeconomic status—can be very different.

¹³Summed effects, as opposed to partial effects, are the predicted value of the outcome from a certain combination of covariates. For example, the summed effects in the first panel of Figure 2 is calculated by assuming a California employed resident with a median zip-code house price in 2015 of \$263,879 and a median zip-code household income of \$68,831, and intercepts added. Because of the lack of individual-specific covariates and the simplicity of the model, the fit is poor, resulting in summed effects that are quite off the scale. Not much should be made of the x-axis numbers other than their relative order.

Appendix D.

We see that, as suspected from the descriptive statistics, Hispanics of either gender are likely to be hidden donors. Asians follow, and then others. The most interesting point from the first panel of Figure 2 is that black men are least likely to be hidden donors. Although the standard errors are larger on estimates of the black population due to their small number, a t-test on the subset of black or white men shows that black men are less likely to be hidden donors relative to white men ($p < 0.001$), although the coefficients are not substantively different. In all races, women are more likely to be hidden donors, and Hispanic women disproportionately so (interaction effect $p < 0.05$).

Occupation-wise, a student is likely to be a hidden donor, while a retired person is likely to be visible, which is as expected. Incredibly, those who claim to be unemployed are more likely to be visible than those who entered regular occupations. This could be indicative of Sanders' popularity among the unemployed. However, whether the self-proclaimed unemployment always reflects the true occupational status is unknown.

Clearly, our results are consistent with the demographic comparisons made in Graf (2006) and Magleby, Goodliffe and Olsen (2018). Hidden donors are more likely to be women, young, and low in income than visible donors, which can be inferred by the distribution of "retired" and "student." Two interesting results were that black men were more likely to be visible donors than white men, and that the unemployed were likely to be more visible than those who reported regular job titles. The donation differences we observe in the data for black and white men need further study. On the other hand, that unemployed were more likely to be visible Sanders donors could be an indication of his popularity among the unemployed, with his economic policy positions.

Differences in Contribution Pattern

Given the demographics differences that we have documented earlier, here we investigate whether hidden and visible donors vary in their contribution patterns.

Amount and Frequency

Table 4 displays the basic descriptive statistics for frequency and sum of all Sanders contributions of the 2016 cycle. Note that there are—obviously—huge differences between visible and hidden contributors. When the median hidden donor gives a dollar to Sanders, the median visible donor will give close to fifteen. Figure 3 shows the density plot of how the sum of contributions to Sanders are distributed, capped at \$500 to clearly highlight the difference. We see that hidden donors' money has a different distribution, and with peaks at various salient contribution amounts such as \$10 (12.0% of hidden donors), \$15 (8.5%), \$50 (7.6%), \$27 (6.1%), and so on. There are no such extreme peaks for visible donors' contributions. 4.8% of visible donors gave \$250, 3.1% gave \$300, and so on.

The \$27 contribution amount is particularly interesting. This is an amount that the Sanders campaign deliberately made salient mid-election, emphasizing that the campaign relied on small donors and not the fat-cats of Wall Street—this emphasis has been extensively noted by the media (Bump, 2016; Foran, 2016; Mehta et al., 2016). It was also echoed in Sanders's later campaign strategies, as they added a \$27 donation option in their suggested amounts. And indeed the campaign had successfully mobilized small donors enough that its median donor was a one-time giver of \$27. However, \$27-giving is more indicative of a visible donor compared to the benchmark distribution. While 20% of Sanders donors have given \$27 at least once, conditional on a \$27 giving record, a donor is 26.5% likely to be visible. This may suggest that such compliance with the campaign message—enough to choose such a non-salient amount—is indicative of donor loyalty, who ends up contributing enough to earn the visibility.

Contributions to other committees also show differences.¹⁴ Note that the majority of Sanders donors are likely not to show up on other committees' itemized contributions, visible and hidden alike. The mean sum to other committees from visible donors are smaller compared to amount to Sanders ($\$531.0 < \632.0). The trend is reversed for hidden donors ($\$57.7 > \45.7), which is quite impressive, and combined with the quantiles, hints at an extremely skewed distribution of giving to non-Sanders committees.

Timing

Timing is a rarely studied aspect of campaign donations, except in the literature on PAC contributions—for example, see Stratmann (1998). However, timing may provide some insight into donor motivations. If we view contributions as a rational act in which a donor weighs her marginal benefits and marginal costs, whether her gift is expressively or instrumentally motivated, the first contribution date reveals precisely when the marginal benefits first exceed marginal costs of the person. If donations are driven by solicitation, timing may show how an equilibrium is struck between demand and supply.

Table 5 shows various summaries of when the donors choose to give. Notice that visible contributors on average enter earlier in the race with their first contribution. While visible donors give long before the primaries, hidden donors enter right before the primaries start. Visible donors also leave later than hidden contributors, again by a difference of two to three months, a substantial amount of time. Note how half of hidden donors started to leave from March 9th or later, although it was the day after the Sanders' surprise win in Michigan. In fact, by April 26 (the time of the so-called Acela primaries), less than one in six Sanders

¹⁴Note that contributions to other committees should be interpreted immediately at face value, and the above descriptives are naive. That not many hidden donors give to other committees can be a mechanical result from the data generating process, because we are not able to detect under-\$200 donations of either visible or hidden donors from other committees unless they come through intermediary committees. We counter this by deliberately keeping in our records the under-\$200 donations via intermediaries for all committees, both visible and hidden contributions. As we will discuss in Appendix C, there is a substantial evidence that intermediaries played a huge role in the 2016 election not just in the Sanders campaign, and we argue that keeping all contributions is an informative, consistent, and conservative method in highlighting the differences in contributor types that we argue for.

hidden donors remained (16.9%). For visible donors, a majority of them stopped only after May 18.

Table 5 shows the percentage of donors who gave to Sanders or non-Sanders committee *after* some key dates in which Sanders' prospect declined. We include Super Tuesday (March 15), Acela primaries (April 26), when Clinton became a presumptive nominee (June 6), when Sanders officially endorsed Clinton (July 12nd), and when the Democratic Convention officially closed (July 28). Note how, while the proportion of hidden donors that give to other committees stays consistent, hidden donors who give to Sanders rapidly decline as the election progresses. This vividly contrasts with visible donors. By the time that Clinton is the presumptive nominee, while almost 30% of visible donors still opened their wallets for further campaign contributions to Sanders, while only 5% of hidden donors did. Hidden donors seem to be quicker to abandon Sanders while steadily supporting other committees, if any.

It is surprising to see that so many visible donors contributed to Sanders after he had effectively lost in his quest for the Democratic nomination. Even if we account for the possibility of campaigns mislabeling donations with different dates than they actually received, the percentage seems substantial—see for example that one out of twenty visible donors gave to Sanders even after Sanders endorsed Clinton. Less than one percent of hidden donors donated after the endorsement in July. Figure 4 graphically shows this by weekly trends of contribution frequency and sum. Note that hidden donors' giving to Sanders—altogether—are much more concentrated in the earlier weeks of primaries, especially in February and March, so

It could seem that a class of donors, who were rapidly disillusioned by the candidate as the primaries progressed, stop giving mid-campaign and as a result became hidden. However, again note that the first contribution dates also differ meaningfully—which is an important piece of evidence. That hidden donors are late entrants to the market signal that they were not die-hard supporters to begin with, and that they may have been persuaded in to give

only after Sanders gained sufficient momentum.

The difference in enthusiasm could be attributed to many factors, such as ideological differences (Ensley, 2009), candidate valence, information costs, or allocable budget. For example, given that donors are instrumental, the ones who are further away from Sanders ideologically may have been attracted to give by his rising prospects, but as his viability fell, they withdrew as their marginal benefit of giving fell below the costs. It could be that these are donors who are generally not interested in politics enough to consider giving, and only when Sanders became a serious figure, were attracted just enough to give once or twice. There could be a donor class with limited budget that can be spared on political donations, for instance due to lower income, and thus cannot afford to give expressive donations such as those made after Sanders loses the primary race. Unfortunately, with the existing data we cannot immediately conclude whether either of these two explanations hold true.

Differentiating Visible Donors

Given the results above, there is a possibility that visible donors become visible simply because they have patiently given many times, and eventually ran over the \$200 threshold. These donors, which we will call *eventually visible* donors, may be different from the ones who have given over \$200 in their first contribution, who we will call *immediately visible* donors. If for instance a donor gives \$50 monthly starting January 2016 and she opts out during or before April, she will be hidden, while if she remains to give once more until May, she becomes visible. Distinguishing the different types of visible donors may be important in understanding differences in their political behavior from hidden donors.

Table 6 shows summary statistics for contribution frequency, averages, and sum across the three types of donors. Eventually visible donors constitute only 2.7% of the donor population, and hence immediately visible donors 9.7%.

By calculating the average donation size, we can see the difference between hidden and visible donors again. While there may exist a subpopulation of visible donors who are more

homogeneous at the high end of hidden donor distribution, it is not universal—the visible donors’ median donation average is more than twice that of hidden donors. More striking is how frequently these eventually visible donors give—the median gives 11 times to the Sanders campaign, seemingly an indication of a consistent support base. In fact, the sum of money from this class of eventually visible donors make up the 44.2% of Sanders’ individual contributions, followed by the sum from hidden donors (38.7%).

What really sets the eventually visible donors apart from hidden donors is seen in Table 7. Note how the eventually visible class of donors persistently gave to non-Sanders committees, in even larger proportions than the immediately visible donors. Immediately visible donors donated larger amounts to Sanders as can be seen from Table 5—we may have expected them to be not only richer but more politically involved than those eventually visible. However, eventually visible donors entered earlier and rode out the election season for much longer than either of the other donor types. Combined, it looks like there is a stark difference between hidden donors and the class of donors who have given enough times to earn themselves the \$200-threshold visibility.

One interesting point from Table 7 is that the median date of first donation for hidden donors and for immediately visible donors are very similar—only four days apart, a week to go before the Iowa caucuses. This again shows that what sets the eventually visible donors apart is *when* they started to give. We cannot make any causal claims here, but given that the eventually visible donors are both early givers and givers to other committees, it will be interesting to see whether an early engagement with a campaign committee evolves into more comprehensive political participation, or those who would have been involved in politics would have started giving early at any rate.

Conclusion

How is our understanding of campaign finance restricted by the Federal Election Commission (FEC)'s data generating process? Because only transactions that exceed cycle-to-date aggregate of \$200 are disclosed, many donors' first few—or all—contributions are censored. We have shown that small donors, who become hidden with censoring (*hidden donors*), are very different from large donors who are usually the only donors visible to campaign outsiders (*visible donors*). Hence we argue that our knowledge of individual campaign contributors is incomplete, and that there are many more people hidden beneath the tip of the iceberg.

Hidden donors, more than seven times in number than visible donors, are more likely to be students, female, and racial/ethnic minority, compared to visible donors. This reflects the economic inequality that may exist in the populations of campaign donors, and hints that we usually only see a limited picture of much wealthier, privileged donors' political behavior. Note that the demographics of the hidden donors are still quite distinct from the average American.

Hidden donors tend to contribute later in the election than visible donors, and their giving is concentrated when the race is heated. The median hidden donor starts to give to Sanders a week before the primaries start, and does not give more than once or twice—whether it be a lack of enthusiasm, Sanders' diminished viability, or the donor's constricted campaign budget, is difficult to identify. Whichever is the case, the hidden donors will have largely cleared out around March 2016, while a substantial portion of visible donors continue to give even with their candidate's dwindled prospects.

The hidden donors, in aggregate, make up 38.5% of the contributions to Sanders from individual donors. While this proportion may be slightly overestimated due to the uniquely popular, small-donor powered nature of the Sanders campaign, the hidden donors are still a significant subset of the financial constituents that a politician must be responsive to. If hidden donors truly reflect the popular vote better than the elite, visible donors, their entry or attrition may be important numbers that a campaign is paying attention to, and

tailoring their campaigns accordingly. If these are younger, less well-off donors who will grow to become wealthier in the years to come, the number and behavior of hidden donors will impact a politician’s long-term election strategies.

If we can find a way to extract the information that is censored, we may be able to answer a more variety of questions in campaign finance and political participation. While we may never be as lucky as in the 2016 presidential campaign to have this much data on a major politician, the ever-increasing reliance of the Democratic campaigns on ActBlue, and that the \$200 requirement stands unadjusted by inflation, will help us uncover the small donors as election cycles progress.¹⁵ How to use this data to counter the “law of available data” will require discretion and attention to details.

¹⁵For a robustness check with different threshold other than \$200, again, which can seem arbitrary, we provide the same Tables and Figures using different thresholds in Appendix F. The same conclusions hold for thresholds \$100, \$500, and \$1,000.

References

- Associated Press. 2015. “Bernie Sanders pulls in nearly as much cash this quarter as Hillary Clinton.” *Chicago Tribune*. October 1, 2015. <http://www.chicagotribune.com/news/nationworld/politics/ct-clinton-sanders-fundraising-20150930-story.html>. (accessed November 19, 2018).
- Barber, Michael. 2016. “Donation motivations: Testing theories of access and ideology.” *Political Research Quarterly* 69(1):148–159.
- Barber, Michael J, Brandice Canes-Wrone and Sharece Thrower. 2017. “Ideologically sophisticated donors: Which candidates do individual contributors finance?” *American Journal of Political Science* 61(2):271–288.
- Blevins, Cameron and Lincoln Mullen. 2015. “Jane, John... Leslie? A Historical Method for Algorithmic Gender Prediction.” *DHQ: Digital Humanities Quarterly* 9(3).
- Bonica, Adam. 2011. “Small Donors and Polarization.” *Boston Review*. July 22, 2011. <http://bostonreview.net/bonica-small-donors-polarization>. (accessed November 19, 2018).
- Brown, Clifford W, Lynda W Powell and Clyde Wilcox. 1995. *Serious money: Fundraising and contributing in presidential nomination campaigns*. Cambridge University Press.
- Bump, Philip. 2016. “Bernie Sanders keeps saying his average donation is \$27, but his own numbers contradict that.” *Washington Post*. April 18, 2016. https://www.washingtonpost.com/news/the-fix/wp/2016/04/18/bernie-sanders-keeps-saying-his-average-donation-is-27-but-it-really-isnt/?utm_term=.15cbffe56b6a. (accessed November 19, 2018).
- Burns, Alexander, Rachel Shorey and Jugal Patel. 2018. “Small Donors Fuel a Big Democratic Lead in 2018 Fund-Raising.” *The New York Times*. Octo-

- ber 16, 2018. <https://www.nytimes.com/interactive/2018/10/16/us/politics/campaign-finance-small-donors.html>. (accessed November 19, 2018).
- Christenson, Dino P, Corwin D Smidt and Costas Panagopoulos. 2014. "Deus ex Machina: Candidate web presence and the presidential nomination campaign." *Political Research Quarterly* 67(1):108–122.
- Corasaniti, Nick. 2016. "Bernie Sanders Campaign Showed How to Turn Viral Moments Into Money." *The New York Times*. June 24, 2016. <https://www.nytimes.com/2016/06/25/us/politics/bernie-sanders-digital-strategy.html>. (accessed November 19, 2018).
- Culberson, Tyler, Michael P McDonald and Suzanne M Robbins. 2018. "Small Donors in Congressional Elections." *American Politics Research* p. 1532673X18763918.
- Dawood, Yasmin. 2015. "Campaign Finance and American Democracy." *Annual Review of Political Science* 18:329–348.
- Ensley, Michael J. 2009. "Individual campaign contributions and candidate ideology." *Public Choice* 138(1-2):221–238.
- Foran, Clare. 2016. "Bernie Sanders's Big Money." *Washington Post*. March 1, 2016. <https://www.theatlantic.com/politics/archive/2016/03/bernie-sanders-fundraising/471648/>. (accessed November 19, 2018).
- Francia, Peter L, John C Green, Paul S Herrnson, Clyde Wilcox, Lynda W Powell et al. 2003. *The financiers of congressional elections: Investors, ideologues, and intimates*. Columbia University Press.
- Gaudiano, Nicole. 2016. "Revolution Messaging helps drive Sanders' 'political revolution'." *USA Today*. March 18, 2016. <https://www.usatoday.com/story/news/politics/elections/2016/2016/03/18/>

[revolution-messaging-helps-drive-sanders-political-revolution/81977160/](https://www.washingtonpost.com/news/energy-environment/wp/2018/11/19/revolution-messaging-helps-drive-sanders-political-revolution/81977160/).

(accessed November 19, 2018).

Goldmacher, Shane. 2016. “Trump shatters GOP records with small donors.” *Politico*. September 19, 2016. <https://www.politico.com/story/2016/09/trump-shatters-gop-records-with-small-donors-228338>. (accessed November 19, 2018).

Graf, Joseph. 2006. *Small donors and online giving: A study of donors to the 2004 presidential campaigns*. Institute for Politics, Democracy & the Internet, Graduate School of Political Management, George Washington University.

Hamburger, Tom. 2016. “Clinton campaign says it has experienced its own small donor boom.” *Washington Post*. February 16, 2016. https://www.washingtonpost.com/news/post-politics/wp/2016/02/16/clinton-campaign-says-it-is-attracting-small-donors-online/?utm_term=.4101d94bce35. (accessed November 19, 2018).

Heerwig, Jennifer A. 2016. “Donations and dependence: Individual contributor strategies in house elections.” *Social science research* 60:181–198.

Imai, Kosuke and Kabir Khanna. 2016. “Improving ecological inference by predicting individual ethnicity from voter registration records.” *Political Analysis* 24(2):263–272.

Johnson, Bertram. 2010. “Individual contributions: A fundraising advantage for the ideologically extreme?” *American Politics Research* 38(5):890–908.

Johnson, Bertram N. 2013. *Political giving: Making sense of individual campaign contributions*. FirstForumPress, inc.

Karpf, David. 2013. The Internet and American political campaigns. In *The Forum*. Vol. 11 De Gruyter pp. 413–428.

- Key, Valdimer Orlando et al. 1955. “Politics, parties, and pressure groups.”.
- Khanna, Kabir, Kosuke Imai and Hubert Jin. 2017. *wru: Who are You? Bayesian Prediction of Racial Category Using Surname and Geolocation*. R package version 0.1-7. <https://CRAN.R-project.org/package=wru>
- Lee, Michelle Ye Hee. 2016. “Fact Checker: Sanders’s claim that he ‘does not have a super PAC’.” *Washington Post*. February 11, 2016. https://www.washingtonpost.com/news/fact-checker/wp/2016/02/11/sanderss-claim-that-he-does-not-have-a-super-pac/?utm_term=.3b597c7129cc. (accessed November 19, 2018).
- 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.
- Malbin, Michael J. 2013. Small donors: Incentives, economies of scale, and effects. In *The Forum*. Vol. 11 De Gruyter pp. 385–411.
- Mehta, Seema, Anthony Pesce, Maloy Moore and Christine Zhang. 2016. “Bernie Sanders’s Big Money.” *Los Angeles Times*. June 3, 2016. <https://www.latimes.com/projects/la-na-pol-sanders-donors/>. (accessed November 19, 2018).
- Mider, Zachary. 2015. “Campaign 2016’s Unlikely Tycoons: Small Donors Fuel Insurgents.” *Bloomberg*. October 16, 2015. <https://www.bloomberg.com/news/articles/2015-10-16/campaign-2016-s-unlikely-tycoons-small-donors-fuel-insurgents>. (accessed November 19, 2018).
- Mullen, Lincoln. 2018. *gender: Predict Gender from Names Using Historical Data*. R package version 0.5.2. <https://github.com/ropensci/gender>

- Nelson, C. 2010. "Strategies and Tactics of Fundraising in 2008." *Campaigns and Elections American Style* pp. 93–104.
- Panagopoulos, Costas and Daniel Bergan. 2006. "Contributions and contributors in the 2004 presidential election cycle." *Presidential Studies Quarterly* 36(2):155–171.
- Patten, Eileen. 2016. "The Nation's Latino Population Is Defined by Its Youth." *Pew Research Center*. April 20, 2016. <http://www.pewhispanic.org/2016/04/20/the-nations-latino-population-is-defined-by-its-youth/>. (accessed November 19, 2018).
- Qiu, Linda. 2015. "Is Bernie Sanders the only presidential candidate without a super PAC?" *Politifact*. September 30, 2015. <http://www.politifact.com/truth-o-meter/statements/2015/sep/30/bernie-s/bernie-sanders-only-presidential-candidates-withou/>. (accessed November 19, 2018).
- Stratmann, Thomas. 1998. "The market for congressional votes: Is timing of contributions everything?" *The Journal of Law and Economics* 41(1):85–114.
- Verba, Sidney, Kay Lehman Schlozman and Henry E Brady. 1995. *Voice and equality: Civic voluntarism in American politics*. Harvard University Press.
- Wilcox, Clyde. 2008. Internet fundraising in 2008: A new model? In *The Forum*. Vol. 6 De Gruyter.
- Wood, Simon N. 2011. "Fast stable restricted maximum likelihood and marginal likelihood estimation of semiparametric generalized linear models." *Journal of the Royal Statistical Society: Series B (Statistical Methodology)* 73(1):3–36.

Tables

Office/Party	Unitemized Individual \$/ (Unitemized + Itemized) (%)
House Democrats	22.4
House Republicans	13.9
Senate Democrats	25.0
Senate Republicans	18.7
Hillary Clinton (Dem.)	25.8
Bernie Sanders (Dem.)	58.1
Donald Trump (Rep.)	64.9
Ted Cruz (Rep.)	38.5

Table 1: Proportion of Unitemized Contributions in Individual Contributions, 2016 Elections

Table 2: Demographics of Visible and Hidden Contributors, 2016 Sanders Campaign

	Visible	Hidden
% of Men	57.2%	53.2%
% of Whites	90.5%	86.7%
% of Blacks	2.8%	3.0%
% of Hispanics	4.2%	7.4%
% of Asians	2.5%	2.9%

Table 3: Occupations of Visible and Hidden Contributors, 2016 Sanders Campaign

	Visible	Hidden
% of Unemployed	26.7%	25.7%
% of Engineer	4.8%	3.1%
% of Teacher	4.1%	4.9%
% of Retired	3.9%	1.3%
% of Attorney	2.3%	1.1%
% of Professor	2.1%	1.2%
% of Physician	2.1%	0.8%
% of Consultant	1.9%	1.3%
% of Student	1.2%	4.0%
% of Homemaker	0.4%	0.3%

(Occupations are sorted by the visible donor percentage for a better contrast.)

Table 4: Quantiles of Contribution Amounts and Frequencies, Visible and Hidden Contributors, 2016 Sanders Campaign

		Mean	Quantiles		
			0.25	0.50	0.75
Visible, to Sanders	Frequency	13.2	5	10	17
	Sum	\$549.0	\$260.0	\$365.0	\$595.0
Hidden, to Sanders	Frequency	2.5	1	1	3
	Sum	\$46.1	\$15.0	\$28.0	\$60.0
Visible, to others	Frequency	19.1	1	4	15
	Sum	\$523.0	\$2.0	\$33.6	\$188.0
Hidden, to others	Frequency	3.8	0	0	2
	Sum	\$13.2	\$0.0	\$0.0	\$5.0

Table 5: Timing of Contributions, Visible and Hidden Contributors

	Visible	Hidden
Median first contribution date to Sanders	10/04/2015	01/29/2016
Median last contribution date to Sanders	05/18/2016	03/09/2016
% of donors who gave to Sanders, after		
— Super Tuesday	83.5%	44.9%
— Acela primaries (Super Tuesday III)	63.3%	16.9%
— Clinton became presumptive nominee	29.5%	4.6%
— Sanders endorsed Clinton	5.5%	0.8%
— End of Democratic convention	0.1%	0.0%
% of donors who gave to non-Sanders cmte., after		
— Super Tuesday	39.4%	13.8%
— Acela primaries (Super Tuesday III)	39.0%	13.3%
— Clinton became presumptive nominee	37.9%	12.6%
— Sanders endorsed Clinton	35.9%	11.7%
— End of Democratic convention	34.3%	11.2%

Table 6: Quantiles of Contribution Amounts and Frequencies to Sanders, Immediately Visible, Eventually Visible, and Hidden Donors

		Mean	Quantiles		
			0.25	0.50	0.75
Immediately Visible, to Sanders	Frequency	6.5	1	3	7
	Average	\$233.0	\$50.1	\$128.0	\$250.0
	Sum	\$745.0	\$275.0	\$481.0	\$857.0
Eventually Visible, to Sanders	Frequency	15.1	7	11	19
	Average	\$43.4	\$21.1	\$32.7	\$52.4
	Sum	\$494.0	\$258.0	\$350.0	\$550.0
Hidden, to Sanders	Frequency	2.5	1	1	3
	Average	\$22.5	\$10.0	\$15.0	\$27.0
	Sum	\$46.1	\$15.0	\$28.0	\$60.0

Table 7: Timing of Contributions, Immediately Visible, Eventually Visible, and Hidden Donors

	Immediately Visible	Eventually Visible	Hidden
Median first contribution date to Sanders	01/25/2016	09/23/2015	01/29/2016
Median last contribution date to Sanders	04/03/2015	05/25/2016	03/09/2016
% of donors who gave to Sanders, after			
— Super Tuesday	61.5%	89.6%	44.9%
— Acela primaries (Super Tuesday III)	38.7%	70.1%	16.9%
— Clinton became presumptive nominee	15.7%	33.3%	4.6%
— Sanders endorsed Clinton	3.0%	6.2%	0.8%
— End of Democratic convention	0.1%	0.0%	0.0%
% of donors who gave to non-Sanders cmte., after			
— Super Tuesday	72.3%	92.8%	13.8%
— Acela primaries (Super Tuesday III)	54.9%	79.2%	13.3%
— Clinton became presumptive nominee	37.6%	56.1%	12.6%
— Sanders endorsed Clinton	28.6%	41.5%	11.7%
— End of Democratic convention	25.9%	36.7%	11.2%

Figures

Figure 1: Proportion of Sanders Donors By Estimated Population, Zip Code Level

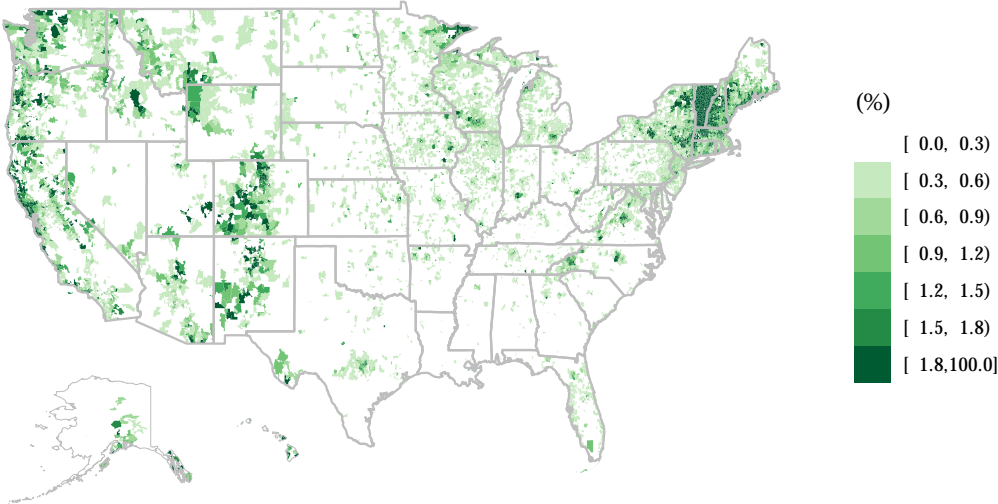


Figure 2: Demographics and Occupations' Effect on Hidden Donor Status

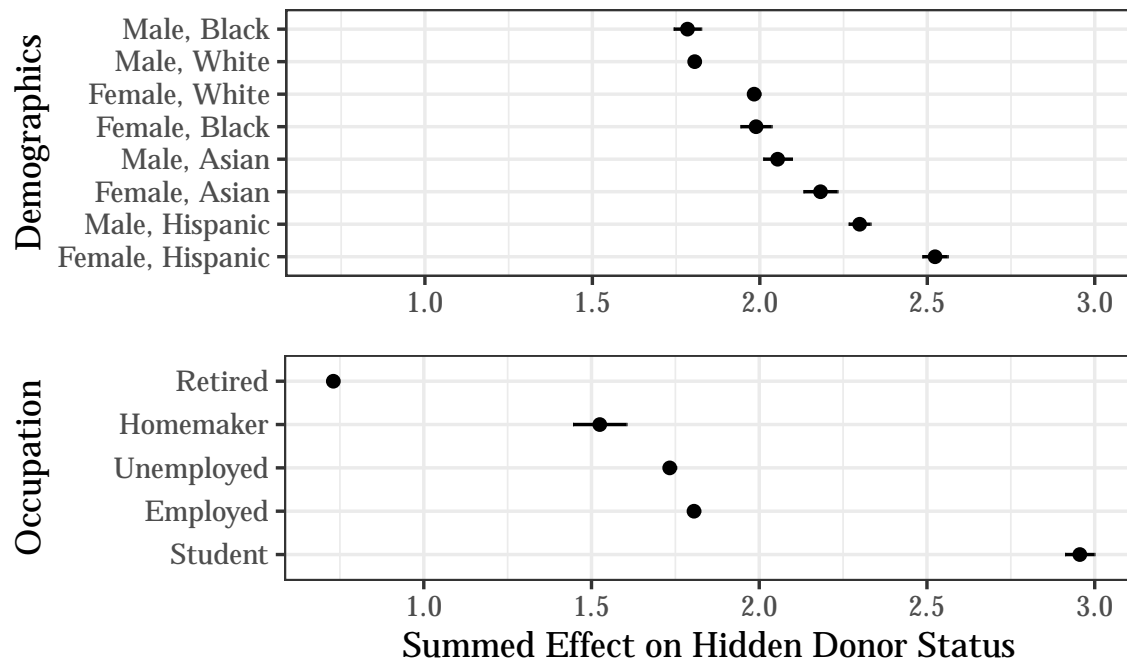


Figure 3: Distribution of Contribution Sum to Sanders, Visible and Hidden Contributors, Truncated at \$500

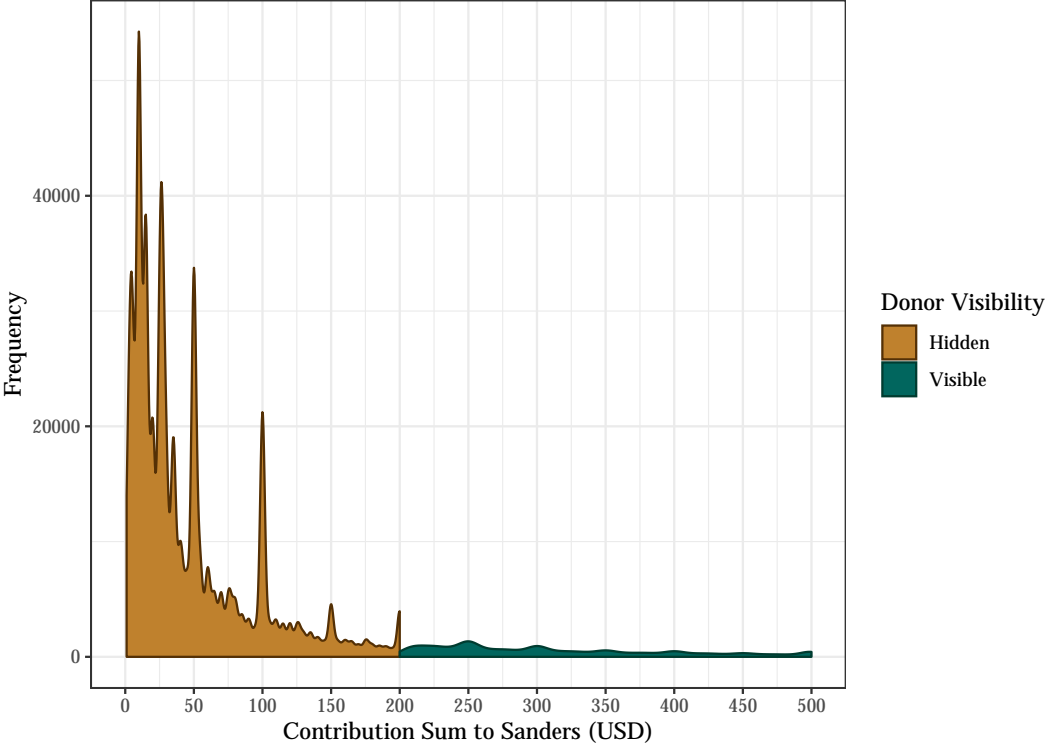


Figure 4: Weekly Distribution of Contribution Frequency and Amount Sum, by Donor Type and Destination

