How Hillary May Have Lost the White House:
The Electoral Effects of Presidential Campaign Visits in 2016

Logan S. Peretz
Honors Thesis
Department of Political Science
Northwestern University
Advisor: Professor McGrath
April 16, 2018

Abstract

Following her loss in the 2016 election, Hillary Clinton came under fire from political pundits who attributed her defeat to a failure to visit key swing states like Wisconsin and Michigan. Did Secretary Clinton's strategy on the ground cost her the election? This thesis investigates the electoral effects of campaign visits by Donald Trump, Hillary Clinton, and Gary Johnson and identifies if visits to a county during the 2016 Presidential Election increased a candidate's share of the vote in that county. The analysis is based on an original dataset of all visits by the major candidates between September 1st and Election Day 2016, disaggregated by the county level. A series of Ordinary Least Squares (OLS) regression models provide a foundation for analysis, while a Nearest Neighbor matching model provides a robustness check. The results imply that Secretary Clinton's visits in the later part of the campaign had a positive effect on her vote share in places she visited, while no significant effect was observed for Trump or Johnson. To this extent, the pundits' criticism appears to have been well founded. Given the incredibly large amount of limited resources that presidential candidates dedicate to holding such campaign stops, this research shows that campaign visits can be well worth the investment, depending on the candidate.

Contents

List of Figures and Tables

FIGURES

TABLES

Acknowledgements

Though my name is attached to my thesis, there is a collection of incredible people who, in one way or another, have made this possible; their names deserve to be here as much as mine does.

First, my love and thanks go to my parents, Joanne and Andy Peretz. This thesis, one of the most rewarding experiences I could have hoped for, is possible due to the culmination of years of support you have given me. I can't possibly put into words how much that has meant.

To my sister, Lauren, thank you for being my best friend.

My Grandfather, Dr. Harvey Peretz, continually makes my dreams come true, and I feel incredibly lucky and humbled to have you in my life at each step of this journey. "Thank you" truly is not adequate to fully express my gratitude.

To my Nana, Doris Minerley, you embody the American Dream, and I thank you for everything you do for me.

None of this research would be possible without my incredible advisor, Professor Mary C. McGrath, whose guidance and expertise has been unparalleled, even while she welcomes a new family member into the world. I'm immensely grateful for the opportunity to have worked with you for well over a year as part of my undergraduate experience.

Professor Andrew Roberts has led each of us honors students through this daunting process from the beginning, and without him we would truly be lost – thank you for your sagacious advice, for your time, and for your mentorship.

For my late Grandmother, Barbara Peretz, whose love, lessons, and passion for her family have served as a continuing source of inspiration in all aspects of my life.

Introduction

The ways in which US presidential candidates have attempted to gain support from the public have evolved tremendously since our nation's founding. Indeed, openly campaigning for the presidency fell strictly outside the realm of what was considered acceptable political behavior early in the United States' history. This paradigm, of course, has shifted massively over the centuries that followed, with campaigns now lasting years and existing synonymously with the elections they precede. Today, the act of crisscrossing the country to visit citizens in their own backyard to gain their votes represents a large piece of the process for candidates. While deciding which particular locales that candidates visit has long been central to campaign strategy, the choice of where to visit has received heightened attention in the wake of the 2016 election given Hillary Clinton's decision to largely pass over Michigan and Wisconsin.

With many pundits making pointed claims that Secretary Clinton would have won the election had she decided to instead spend more time in those states, an important puzzle has materialized: do these usually brief in-person candidate visits have a strong enough influence to translate to receiving more votes on election day? Though conventional wisdom suggests that campaigns, campaign strategists, and journalists alike generally classify campaign visits as a vital part of the election process, how much do these trips actually matter in the grand scheme? This thesis examines the effect that in-person campaign visits had on electoral support, as defined by percentage of vote share, in the areas which Donald Trump, Hillary Clinton, and Gary Johnson brought their physical presence during the 2016 election.

Given the massive scale of resources that campaigns allocate to their respective candidates' travel plans, the extent to which visits may or may not correspond to gaining votes carries significant weight in the discussion of campaign strategy. Naturally, if a candidate's

physical presence tends to have an empirical effect, campaigns can continue to leverage in-person visits to target areas that represent particular importance to the outcome of the election. In the event that campaign visits have a relatively weak effect on electoral support, however, this finding may significantly disrupt the way in which the conventional presidential campaign operates.

The results of my work suggest that Secretary Clinton's visits had a positive effect on her on share of the vote, while visits by President Trump and Governor Johnson did not have a positive effect. As such, depending on the candidate, visits can help tip the scales at the polls on election day.

Further, the finding that Secretary Clinton's visits had a positive effect lends support to critics who suggested her lack of campaign stops in swing states such as Wisconsin, Michigan, and Pennsylvania cost her the election.

While the effects of both broader advertising segments aired on TV and more narrowly targeted ads on social media are well documented as being statistically significant, a comprehensive data analysis of campaign visits' electoral effects has not been conducted in almost a decade. For this reason, this work fills a wide void in the existing literature and has significant consequences for how the zero-sum game of campaign spending is treated.

Chapter 1

Review of Relevant Literature

While typically considered central to the campaign process, the effect of in-person campaign visits and measuring the impact thereof has received perhaps a relatively low level of coverage by political scientists over the past twenty years. Potentially most important to caveat at the onset of this review of relevant literature is that though plenty of prior work exists regarding the effects of campaigns in general, few studies have explicitly set out to identify the relationship between visits and electoral support. Rather, much more frequently, past research has tended to focus on the more general "campaign effects," or what matters in campaigns at the macro level (i.e. national scandals, nominating conventions, etc.). That said, it is possible to divide research on campaign visits into multiple camps.

First, as previously mentioned, there are studies which provide helpful insight into campaign effects in general and touch upon campaign visits but for one reason or another, as will be elaborated upon below, fall short in allowing us to make meaningful conclusions regarding the relationship between visits and electoral support.

Next, there exists completed work which more directly addresses campaign visits and their effect on electoral support and tend to conclude that such an effect exists.

As a helpful visual aid, Table 1.1, located at the end of this section, summarizes the attributes of each piece of literature included for review.

Beginning with the general campaign effects literature, Holbrook and McClurg (2005) use a regression analysis, aggregating information at the state level, to examine the effect of campaign visits on voter mobilization as measured by turnout rates. Ultimately, the two conclude that there is no statistically significant effect of visits on turnout in the 1992, 1996, and 2000

elections. Importantly, the dependent variable here is mobilization, which is not the only means through which visits influence the polls.

Similarly, Holbrook (1994) conducted an earlier regression analysis, looking at the 1984, 1988, and 1992 presidential elections. Holbrook's most noticeable conclusion is that national conditions, such as the economy, carry as much as five times more impact on voters' decisions than do "campaign events," where these events are defined primarily as debates and national nominating conventions. Here, the independent variable "campaign events" does not include visits and, moreover, even though national conditions were more influential, "campaign events" still were statistically significant.

Closely related is Shaw (1999), wherein the author utilizes multiple regressions to examine the effect of campaign events, defined as occurrences that can alter the course of the election, such as scandals, major policy announcements, and nominating conventions. While Shaw ultimately concludes that some of the many campaign events in question have a statistically significant effect on voter preferences, the variable in question is not closely related enough to visits to draw any specific conclusions on that front. Interestingly, these studies are just a small set of cases often cited in literature reviews of newer work - on the topic of campaign visits in particular - as examples of research which concludes there is no effect of campaign visits.

An excellent example of this exists in the work done by Wood, (2016), in which the author cites many of the aforementioned studies as literature disproving an effect of campaign visits on voters' decision making. Wood's paper examines two things in particular: the effects of campaign visits on news coverage, and the effect of visits on mass attitudes towards candidates, as measured by a bespoke survey. Wood's survey found that visits can have a small effect on

mass attitudes, depending on where the visits take place and who the candidate is. Interestingly, the survey also found that likely voters from both parties were prone to reporting a visit by their party's candidate had occurred in their area recently, despite the fact that such visits had not taken place. Most relevant to point out at the onset of this study, however, is that Wood's blanket manner of citing past work provides a case exemplifying the current state of the misclassification of literature covering election visits.

There exists a clear pattern among this first group of research: because these authors set out to examine slightly different campaign phenomena than this paper, their findings cannot appropriately play the role of drawing meaningful conclusions regarding the effect of campaign visits in particular. Indeed, either the dependent variable doesn't measure electoral support, or the independent variables - such as "campaign events" - similarly are not useful to reach conclusions regarding visits and their effect on gaining votes.

To this extent, it is unsurprising that Druckman (2004) comments in his work on media priming in presidential elections that researchers are oftentimes not in alignment over the appropriate dependent variable for measurement when studying electoral phenomena.

Finally, Hillygus (2004) decided before embarking upon research on campaign effects that looking at campaign visits in particular proved too difficult and instead the work focused on televised debates, which may be indicative of the potential issues faced by past researchers which have elected to look more broadly at "campaign events."

On the other hand, a second group of research exists that more specifically targets campaign visits rather than general campaign effects and, equally importantly, measures electoral support in some form as the dependent variable.

Herr (2002) examines the 1996 presidential election and conducts a regression analysis at a state-level of aggregation similar to that of Holbrook, except with the explicit intention of investigating the effect of campaign visits on electoral support, measured by vote share. The author includes statistical controls such as the past voting history of each state in previous presidential elections, demographic information, and TV advertising by each campaign for each state. Additionally, Herr includes as a dependent variable in one of his regressions the percent increase of voter turnout relative to the previous two elections.

Indeed, Herr concludes that campaign visits between October 1st and Election Day had a statistically significant effect on vote share for Clinton in a positive direction, but not for Dole, presenting an interesting perspective that visits by different candidates are not functionally the same when it comes to gaining votes. Herr also found that visits did not have a statistically significant effect on voter turnout. This is of particular interest, as Herr's conclusion regarding turnout is in line with that of Holbrook and McClurg's (2005) aforementioned finding – highlighting that the lack of an effect on turnout does necessarily not tell us anything about a potential effect on electoral support. Indeed, Herr's work represents a useful model to emulate as I execute the research behind my thesis. That said, there also exist limitations to Herr's work.

Though Herr includes a variety of control variables in his regression analysis, this type of analysis still ultimately cannot control for the endogeneity problems which campaign visits present. Despite this, he draws the ultimate conclusion that there exists a statistically significant effect of Clinton's visits on his own vote share. Rather, a more appropriate conclusion to draw would be that a statistically significant relationship exists between President Clinton's visits and his own vote share. It could, perhaps, be the case that Clinton visited states that he felt he was doing better in - we simply do not know in which direction the causation works given a

regression analysis. Considering advances in statistical methods since the publication of Herr's work, it is apparent that better techniques can now be used to more adequately control for endogeneity.

In conducting similar research, Hill, Rodriguez, and Wooden (2010) additionally find that visits have a statistically significant effect for some candidates in various presidential elections between 2000 and 2008. As does Herr, the authors utilize a multinomial regression model to examine specifically the effect of in-person campaign visits, instead of general "campaign effects." Rather than observing the outcome as a product of votes received on election day, however, Hill, et al. instead utilize the results of weekly tracking polls as their dependent variable. This presents some problems for them, given the incomplete nature of the datasets they retrieved for elections between 2000 and 2008, causing them to drop states from their analysis, and use extrapolation and interpolation to fill missing data points in some other cases. Additionally, the methodological implications of using tracking polls differ from those of using post-election vote shares, as will be elaborated upon in *Methodology*.

Further, an experiment carried out by Shaw and Gimpel (2012) entailed actually randomizing Texas governor Rick Perry's gubernatorial campaign schedule over the course of a three-day period and found that visits had a significant effect on voter support which lasted beyond a secondary survey conducted 7 days after the visit. While, of course, this thesis will not be able to employ similar methods of randomization and looks instead at a presidential scale, Shaw and Gimpel's work provides reason to be optimistic about investigating the effect through alternative means such as those employed by Herr as well as Hill, et al.

Indeed, Shaw and Gimpel provide a sound, experimental example that in-person visits have an effect on electoral support, and show there is good reason to investigate this phenomenon on the presidential level.

When examining the pattern prevalent in this second group of research, it appears as if research that explicitly focuses on campaign *visits* rather than campaign *effects* (visits categorically lumped with a variety of other factors) is far more able to reach meaningful conclusions regarding the topic this thesis concerns. On a similar note, studies that explicitly target electoral support as the main outcome - primarily as measured by share of votes at the polls in a presidential election - rather than mechanisms such as mobilization, yield useful results.

This in mind, my research fits very nicely into the landscape of the pre-existing research done on this topic. While previous authors have had success in running analysis of the effect of visits on share of votes, no research has been carried out on the 2016 election in particular. Additionally, the methodological limitations of prior studies represent a chance to now use newer statistical techniques to explore this topic. Further, the lack of pre-existing studies on the most recent election provides me a unique opportunity to move this area of political science research forward.

Table 1.1: Summary of Relevant Literature

Study	Dependent Variable	Independent Variables	Year/Election Covered	Level of Aggregation	Main Results
Holbrook and McClurg (2005)	Turnout rates	Campaign Visits	1992, 1996, 2000 (Presidential)	State	No substantial effect on overall turnout, but may affect turnout of partisans
Holbrook (1994)	Point average in public opinion polls	"Campaign events" (treated as a single variable)	1984, 1988, 1992 (Presidential)	National	National conditions matter significantly more than "campaign events"
Shaw (1999)	Point average in public opinion polls	"Campaign events" (treated as separate variables)	1952 to 1999 (Presidential)	National	Some "campaign events" do have a statistically significant effect
Wood (2016)	Bespoke survey results	Campaign Visits	2012 (Presidential)	Individual (in 4 battleground states)	Visits can have a small effect
Herr (2002)	Vote share	Campaign Visits	1996 (Presidential)	State	Visits in October had a significant effect for Bill Clinton
Hill, Rodriguez, and Wooden (2010)	Changes in weekly state opinion polls	Campaign Visits	2000, 2004, 2008 (Presidential)	State	Visits can have an impact, but varies by candidate and location
Shaw and Gimpel (2012)	Public opinion, media coverage, contribution, and volunteer data	Campaign Visits	2006 (Texas Gubernatorial)	City	Governor Perry's visits had a positive impact on his own public opinion data, but also for his opponent

Chapter 2

The 2016 Election: Background

Though the 2016 Presidential Election took place on Tuesday, November 8, 2016, the campaign began long before then when Ted Cruz declared his candidacy in March of 2015. Following Mitt Romney's 2012 defeat to incumbent President Barack Obama, no clear frontrunner for the Republican nominee for president existed, resulting in an unusual Republican field, both in size and composition.

A total of 17 candidates had tossed their hats into the ring, including early favorites Jeb Bush, Marco Rubio, and John Kasich. These seasoned politicians were joined by more unorthodox candidates such as neurosurgeon Ben Carson and celebrity businessman Donald Trump.

Though Trump's campaign was first written off as a publicity stunt, his unusual "speak my mind" attitude, populist policy platform, and image as an outsider firmly against corruption and political correctness helped to quickly thrust him to the top of the most crowded field in presidential primary history. Though establishment Republicans had hoped a contested nominating convention would provide a route to awarding the nomination to someone other than Trump - who was viewed as hijacking the party - unexpectedly large electoral victories by Trump's campaign during the primaries led to each of his competitors withdrawing from the race before the convention itself had occurred during the summer of 2016.

The race for the Democratic nomination, however, represented a vastly different state of affairs. Given her political experience, previous presidential runs, and recent stint as Secretary of State, Hillary Clinton had long been considered the only serious candidate for the Democratic nomination and, indeed, represented the favorite for the presidency itself. Though Senator Bernie

Sanders, with his grassroots campaign efforts and self-described democratic-socialist platform looked to offer an alternative for voters, Clinton easily won the nomination.

Meanwhile, former two term Governor of New Mexico, Gary Johnson had been awarded the Libertarian Party's nomination for the presidency.

Both Trump and Secretary Clinton brought with them into the general election unusually low favorability ratings due to a variety of controversies on each of their behalves. As a result, Johnson represented an unusually popular third-party candidate, at times drawing comparisons to Ross Perot's presidential runs in 1992 and 1996.

After the dust had settled from each major party's nominating conventions, the stage was set for what was then considered to be a relatively easy election for Secretary Clinton to win (Wicks 2016). While public opinion polling had consistently shown her as the clear victor over Trump from the beginning, a slew of scandals for "The Donald," ranging from leaked sexist remarks to allegations of sexual misconduct had further entrenched the prevailing thought that Clinton would win the election in a landslide.

During the period between the conventions and Election Day, the candidates took approaches towards campaigning which differed substantially. Indeed, Clinton's campaign, while clearly more organized, had made television advertising much more of a priority than the Trump campaign, having outspent him by a factor of over 5-to-1 by the end of September (Migliozzi, et al. 2016).

In terms of organization, Clinton followed the Obama campaign's template of using high amounts of local organization with many staffers on the ground, as well as a particular emphasis on Get Out the Vote (GOTV) efforts, more so than their opponents did. The campaign would

later receive criticism after the election, though, as it appears these voter mobilization efforts may have actually driven Trump supporters to the polls, as well (Smith 2017).

On the other hand, Trump relied on Twitter and bombastic public statements to create free media attention for himself. On top of this, Trump's campaign sought, to much more of an extent than Clinton, to use social media sites like Facebook as a means through which fundraising efforts are brought directly to loyal supporters (Lapowsky 2018).

Finally, in regard to the ground game, Trump had also elected to make a considerably larger number of visits than did Secretary Clinton between September 1st and Election Day. While Trump visited 99 locations, Clinton held only 62 campaign visits. In addition to this tendency to visit more places, Trump displayed a preference for large, boisterous crowds, and consistently bragged about the historically large attendance at his events.

In comparison to previous elections, Trump's rallies often played the role of a spectacle: he engaged in arguments with members of the crowd and press, while violence both inside and outside his events was common. This served the role of gaining Trump unusually large amounts of free media attention. On the other hand, Clinton's events played a much more traditional role as part of her larger campaign.

While Governor Johnson had captured more attention than typically received by third-party candidates, his campaign efforts were dwarfed by those of the two major parties. This is reflected by the fact that his fundraising, ad expenditure, and campaign visit totals were nowhere near the scale of his opponents'. On top of this, his exclusion from the major presidential debates prevented him from gaining a level of legitimacy needed to seriously contend in the election.

Though early results on Election Day 2016 looked promising for Secretary Clinton, a string of rather unexpected victories for Trump in swing states such as Florida, Pennsylvania,

Wisconsin, and Michigan quickly turned the tables as the night progressed and resulted in perhaps the most surprising electoral victory in modern American History.

Chapter 3

Discussion of Campaign Visits

As this research focuses on the effects of campaign visits, a discussion of visits themselves, as well as potential mechanisms through which visits can affect the outcome of an election, becomes helpful at the onset.

The specific motivation for candidate visits may stem from a variety of places (Wood 2016). First, perhaps most commonly thought of, is that candidates hope to cause an increase in media coverage through their appearances, on two levels: local and national. While it is no difficult task for any candidate to say heartfelt things directed towards the residents of a particular area, it is a visit to the area in question itself that represents a far more meaningful event than merely pandering in a press conference or talk show appearance.

The usual pomp and circumstance of a high-profile visit, such as landing a campaign aircraft at a local airport and the use of local roads by a candidate's motorcade help frame the candidate in a way that is far more salient to local voters than would otherwise be possible. Additionally, local press homes in on these appearances, projecting images of candidates to those in the area. In particular, the candidate is given a platform through which they can direct their statements on local issues rather than more nationally relevant talking points.

Beyond local press coverage, a candidate may hope that visits produce more appearances in national news coverage. In addition to the fact that moving a candidate around the nation to a variety of campaign events helps fuel the horse-race style coverage of employed by national press outlets, there also exist less obvious press benefits on the national level.

The long process of traveling the nation to different campaign appearances helps foster a friendly relationship between the candidate and the press corps that accompany the campaign,

which may be beneficial to the candidate. Timothy Crouse's 1973 book *Boys on the Bus* and Alexandra Pelosi's 2003 documentary *Journeys with George* both highlight the type of camaraderie that develops between the traveling press and the candidate, as well as their staff.

Finally, candidates may hope that their campaign appearances will prove effective in tapping into the variety of beneficial resources that exist in each location. For example, rallies provide a prime location for a campaign to recruit volunteers en masse. These volunteers serve as useful manpower by canvassing door-to-door, conducting phone calls on the candidate's behalf, and operating field offices - all important aspects of a local campaign effort. In addition to gaining the support of volunteers, campaign visits also represent a useful opportunity for the candidate to connect with local politicians who have the ability to serve as advocates on the nominee's behalf in the community during the larger campaign.

Further, while private fundraising events represent a natural source of donations, rallies and other public appearances may similarly have effects that spillover into the realm of raising funds. Visiting a location may serve as a helpful means of raising money to line the campaign's war chest by showing potential donors that the candidate cares about issues relevant to the local community.

Visits are, of course, not made randomly. Rather, a strategic process lies behind the decision to visit particular locations. Perhaps obviously, states identified as "battleground" or "swing" states receive a particularly high level of attention during the campaign. Indeed, given the structure of the electoral college, it makes little sense for a candidate to visit a state that they are virtually guaranteed to win. If campaigns believe that visits carry with them some sort of an effect on voters, it becomes natural that their limited funds are spent on visiting locations that impact their ultimate success in being elected.

Empirically, this is reflected in the fact that 94% of candidate visits in the 2016 presidential campaign took place in 12 states. Of these 12, 11 had been previously identified as battleground states (Stanage 2016). For this reason, it is unsurprising that Scott Walker, former candidate for the presidency in 2016 campaign, remarked that "the nation as a whole is not going to elect the next president. Twelve states are."

Chapter 3

Hypotheses

Myriad factors contribute to an individual's decision of whom to cast a vote for. As such, a variety of independent variables are at play when examining what leads to a candidate performing better at the polls. First, I hypothesize that the number of times a candidate visits a given geographic area has a positive effect on that candidate's vote share.

H1: The number of times a candidate visits a given county has a positive effect on the share of votes that the candidate receives, when controlling for other factors.

As discussed in the prior chapter, the channels through which visits can affect electoral support are plentiful. For example, visits could lead to voters believing the candidate cares about their locale, and that helping to elect a candidate that visits often could be personally beneficial to a voter in the long run once the candidate takes office. To this point, how a candidate appears to a voter is known to matter significantly - generating emotional appeal has traditionally been critical for past candidates, and it is often the person, rather than the fine details of their policy that matter. Indeed, perhaps it is that visits generate a positive emotional response and humanizes the candidate in the eyes of the voter. Additionally, it is important to recognize that a variety of other factors are at play. For example, income, age, race, and how the location voted in previous elections are all extremely predictive of how an individual will vote, and there is a clear causal pathway between demographic information and voters' decisions (younger people tend to be more liberal due to the generally liberal nature of higher learning institutions, black voters tend to vote for democrats as democratic policy tends to be more inclusive towards black citizens, and

wealthier people tend to vote republican as conservative fiscal policy typically benefits wealthier individuals). As will be discussed in later sections, controlling for these variables helps me to estimate the effect of visits.

H2: Visits between October 1st and Election Day will have more of positive impact on the share of votes the candidate receives than visits which take place between the nominating convention and September 30th.

As discussed above, generating an emotional response from a voter is critical for a candidate who wants to receive their support on election day. Indeed, I hypothesize that visits closer to election day matter more for two reasons. First, it is quite possible that the average American voter simply does not pay enough attention during a seemingly endless campaign process until the election is imminent and media coverage is at a peak. Second, it is also possible that the memory of a voter is short and that the effect of a visit fades over time. For example, a visit in early September may not resonate very much with a voter in November, as the emotions and excitement will have largely worn off after such a gap.

It is important to note that this is perhaps counter to one prevailing theory that stipulates voters have made up their mind by the late stages of the election and are less prone to having their opinion swayed (Bowman 1996).

Chapter 4

Methodology

Case Selection / Scope

For the sake of my research, I looked at candidate visits which take place after the two major parties' nominating conventions, beginning with September 1, as this represents the primary time in which presidential campaigns begin to significantly ramp up their ground game. I chose to not include data before this time to isolate the effects of visits by candidates during the general election campaign.

Indeed, tracking visits from before September, while primary season is still underway and visits are less frequent, may add unnecessary noise and complications. The inclusion of visits before nominating conventions would entail an entire gamut of new factors to control for, such as visits by a plethora of competitors in the primaries. Ultimately, this would detract from my hope to examine the effect of in-person campaign visits on the general election in particular.

Additionally, the analysis is limited to visits that were conducted by candidates for the purpose of appealing to the voters of the area where the visit takes place. These visits, for example, take the form of campaign rallies, appearances at voter registration events or other similar occasions. A specific form of visit that was deliberately excluded was private fundraising events, as these visits aim less to have a direct impact on the voter in the area, and the focus is not outwardly facing such as with the category of visits being included for analysis.

While there may be overlap between fundraising events and campaign events, as rallies serve as a source of fundraising, these events still have an outward appeal and are included. What is left out, however, are events such as Trump's private Mar-a-Lago fundraisers which are held with far less of an intent to appeal to the general public.

Specifically, analyses of the Clinton, Trump, and Johnson campaigns were conducted. While third party candidates typically have a negligible effect on the overall dynamic of the election, the high levels of unpopularity surrounding candidates Clinton and Trump led to an unusually high number of voters to ultimately cast a ballot for the Johnson campaign as a third-party option. A similar approach was taken by JP Herr in his analysis of the 1996 election due to the popularity of Ross Perot. To this extent, it is necessary to include the Johnson campaign as part of my analysis to inform a holistic understanding of the dynamics at play in the 2016 election.

In regard to the scope of analysis being run, counties were used as the level of aggregation. While looking at the state level represents an attractive option, as used in past work on the subject, the availability of data from the 2016 election allows for a more specific analysis that is preferable for several reasons.

First, zooming in on counties allows us to make more specific conclusions regarding the effects of visits. On the state level, explanatory power is lost due to the size of states and the differing demographic nature of population centers within them. Indeed, in investigating electoral behavior, control variables are critically important, as they allow us to make comparisons across similar geographic areas and examine the effect of varying the number of visits to that area. As an example, lumping all of Illinois' very demographically different counties together would produce a different statistical representation than if each of the counties were separated out. Similarly, visits in one area of a state may only have localized effects on voting behavior and including all state visits together statistically may make such effects harder to identify.

Second, while there are 50 states, plus the District of Columbia, that are relevant when conducting state-level analysis, there are over 3100 counties in the nation. Increasing the sample size of our analysis by such a large quantity gives us a level of statistical power that is much higher than if aggregation took place on the state level.

Dependent Variables

The dependent variable I use for measurement is the amount of electoral support a 2016 presidential candidate received in that election, as measured by the percentage of total votes that the candidate in question received in each county.

Though some past authors in this field have examined electoral support through public opinion polling, this method can be less helpful for a variety of reasons relating to the scope of the polls themselves. Specific polls typically will either focus on the nation itself, or a specific location (usually a single state), rather than multiple locations, such as all battleground states. When attempting to paint a holistic picture of voters' behavior across the nation, it becomes inevitable that combining the results from different polls must be employed to create a sufficient data set, as have previous studies employing polling data to investigate this phenomenon. Indeed, using the results of different polls with varying methodologies and combining to create a single outcome to measure is often not ideal from a methodological point of view. Additionally, such polls are uniformly not specific enough to perform analysis at the county level - which this thesis aims to conduct - due to the geographic spread of the sample used in the polling process. Instead, to track electoral support, the results of the election itself provides geographically narrow, concrete, observable data of how voters behaved, rather than a macro level projection as to how they will behave.

Independent Variables

Regarding independent variables, there are two important classes of information needed to be included, as mentioned above in *Hypotheses*. First, it is necessary to include variables pertaining to candidate visits. Second, a variety of controls are important, as with any statistical analysis. For candidate visits, variables were generated for each candidate using data obtained through the candidate visit tracker built by journalists Libby Isenstein, Andrew McGill, Kimberly Wailey, and Adam Wollner, wherein the location and date of all campaign visits which took place in the 2016 cycle could be found (Isenstein, et al. 2016). This information was self-coded into the number of visits by county for each candidate, broken down by September and October visits, with the former representing visits taking place in a given county in the month of September 2016 and the latter representing visits taking place between the beginning of October 2016 and Election Day. This setup gives us 6 variables: Clinton September Visits, Clinton October Visits, Trump September visits, Trump October visits, Johnson September visits, and Johnson October visits. This chronological breakdown allows for the separation of visits by closeness to Election Day and, thus, gives us a means of testing the second hypothesis that visits occurring later in the election cycle generate more of an impact on electoral support.

Next, I utilize a variety of control variables, each representing a predictor of vote choice. First, a particularly strong predictor of future vote choice is past voting behavior. In this regard, I look to the level of support a county has given either the Republican Party or Democratic Party in past elections as a control, using data from the New York Times ("Presidential Election Results" 2017). From this I build a control into the right-hand side of the equation by setting up a variable representing the average of the 2012 and 2008 election in a given county for one of the parties (having averages for both parties would present collinearity problems). This takes the

form of an independent variable defined as Base Republican= ([Romney 2012 vote share + McCain 2008 vote share]/2), similar to the approach used by Herr in 2002

Another option to control for a candidate's endogenous population, employed by Wood in his 2016 paper, exists in setting up the dependent variable as the difference between partisan share of votes in 2016 and the average of the previous two presidential elections, thus showing a positive or negative effect relative to previous years. As an example: (Trump 2016 vote share - [(Romney 2012 vote share + McCain 2008 vote share)/2] would exist as the Trump dependent variable.

I elected against this as to avoid defining a key outcome variable in relative terms, especially when considering the rise of third party candidates in 2016 meant that the two major parties received a slightly smaller percentage of the vote across the board. For this reason, the variable remains on the right-hand side of the equation and operates as a traditional independent variable, rather than introduce unnecessary noise to the dependent variable, especially when there is no clear benefit to doing so in this case. Indeed, Herr's successful work in 2002 included this variable as an independent variable rather than manipulating the dependent variable.

Next, a variety of demographic data are included as controls, split into racial and economic variables, which come from a 2016 paper entitled *Inequality across US counties: an S factor analysis* by Emil O. W. Kirkegaard (Kirkegaard 2016). For race, several variables are included, each representing the proportion of the population in a given county that identifies as a particular race: White, Black, Asian, Hispanic, and Native. The economic control takes the form of the median income of the population in a given county, expressed as a unit of thousands of dollars.

These demographic and party base controls are especially important, as such information is a strong predictor of the voting behavior in a county. By controlling for these factors, the effect of campaign visits on vote share can be isolated more successfully through the elimination of potential confounding variables.

The final control reflects the most common way in which a candidate seeks to influence voters: advertising. As county-by-county data on campaign advertising is not available, this variable applies to all counties in a given state, the level at which campaign advertising is tracked. Additionally, the advertising variables represent the amount each campaign spent on advertising expressed in terms of millions of dollars, as reported by Ad Age, which compiled a "Presidential Campaign Ad Score Card" (Dumenco 2016). While the relationship between campaign visits and electoral support is perhaps under-studied academically, there exists a more proven link between a campaign's advertising spend and their success at the polls. By controlling for advertising spending, we can examine fluctuations in vote share by county that are unrelated to a campaign's advertising efforts.

Statistical Procedure

For the statistical approach itself, three separate ordinary least square (OLS) regression models were used with the dependent variables in each case representing the share of votes received by each candidate. All three models include visit variables for each candidate, broken into September and October visits. With each successive model, however, additional controls are successively added to show that some relationships between visits and electoral support hold even when including increasing amounts of controls. The composition of each model is reflected below in Table 4.1.

Table 4.1: Constituent Independent Variables for Each Regression Model

Model 1	Model 2	Model 3
Campaign Visit Variables	Campaign Visit Variables	Campaign Visit Variables
Republican Base Variable	Republican Base Variable	Republican Base Variable
	Demographic Controls	Demographic Controls
		Advertising Control

Additionally, all models are iterated three times, once per candidate, for a total of 9 regressions (i.e. a Clinton Vote Share run was completed, followed by a Trump Vote Share run, then a Johnson Vote Share run, all following the pattern in Table 1).

Separating out the regressions by candidate allows for more explanatory power than simply running one regression, as the dependent variable in the latter case becomes difficult to appropriately select, especially given a 3-person race. On the other hand, the chosen 3-model setup allows for the measurement of the effect of one candidate's visits on another candidate's vote share (i.e. in the regression looking at Trump's share of votes, observing if the coefficient on the "Clinton Visit" independent variable is significant one way or the other, which helps to determine who each candidate is taking votes from).

While using a regression analysis provides a great overview of potential correlations between visits and vote share, simply using OLS models does not allow us to make determinations as to whether or not the relationship is causal. Indeed, campaign visits are planned strategically, rather than assigned to locations randomly, potentially biasing the results generated by a regression, as there are certain, possibly unobservable, factors that cause specific locations to be chosen for candidate visits.

In this case, a matching technique is helpful to more accurately compare counties that do receive visits with counties that do not receive visits from candidates. Matching, as a technique, looks to weed out endogeneity by finding statistically similar observations between the treatment and control groups through analysis of the covariates at play. Though there is no way to guarantee endogeneity has been eliminated, as there still may exist unaccounted for or unobservable variables at play, the key is that the weighting of observations in such a model help provide a more robust estimation of the relationships being measured.

Specifically, the model matches those similar observations across control and treatment groups, and measures the difference in outcome between each pair, as a means of producing an average effect.

In this case, a 1:1 Nearest Neighbor Matching technique was used to identify the Average Treatment Effect on the Treated (ATT) of each of the visit variables to more accurately identify if there is an effect of visits on vote share.

As the model require a binary treatment variable, a 1 is assigned to counties in which any visit took place for each candidate in each month, while a 0 is assigned to counties not visited by that candidate for the given month. As such, the treatment variables are (Trump September Dummy, Trump October Dummy, Clinton September Dummy, and so forth).

This particular technique entails calculating a Mahalanobis distance - a geometric measure of difference - for each observation in the treatment and control group, which is helpful for identifying statistical similarity.

Each observation in the group of counties that received a visit during the election is matched 1:1 with the non-visited county that is most statistically similar based on Mahalanobis distance (hence, nearest neighbor).

Additionally, this analysis is run with replacement, meaning that two treatment observations can be matched with the same control treatment. This is particularly helpful here, as 14 different covariates are at play and we are most interested in finding the best statistical similarity between counties that have been visited with counties that have not been visited. Indeed, allowing for replacement makes this easier than if replacement were not used.

The control observations not matched with treatment observations are not included in the analysis.

Finally, it is important to note that the estimand being used here, the Average Treatment on the Treated (ATT), is slightly different than that of the regression models. While the regression models effectively measured the Average Treatment Effect (ATE), the ATT measures the effect of visits for counties that were actually visited, rather than estimating an effect for the population as a whole (as the ATE measures).

Chapter 5

Results

Descriptive Statistics Regarding Visits

Before delving into results of the aforementioned statistical models, it is first helpful to provide an overview of the visits which each candidate took during the campaign to foster a better understanding of the data being analyzed.

Regarding the total number of visits by each candidate, Mr. Trump took the most with 99 during the campaign. Secretary Clinton took considerably fewer than Mr. Trump, having taken 62 visits. With the fewest number of visits, Governor Johnson took 27.

This discrepancy between the number of visits by the two major parties' nominees is especially prominent in September, wherein Trump made nearly double the number of campaign stops that Clinton did: 32 as compared to Clinton's 17.

Interestingly, all but 28 of Mr. Trump's visits were in unique counties, having visited 71 different counties across his 99 visits. On the other hand, a smaller percentage of Secretary Clinton's visits were in a unique county, with 38 of her 62 visits taking place in unique counties. Finally, Governor Johnson visited a unique county in 25 instances from his 27 total visits.

The specific, unique, counties which each candidate visited are reflected below in Figures 5.1 - 5.3.

Many of the counties in question played host to a candidate on multiple occasions, with 37 counties receiving 2 or more visits. Unsurprisingly, many of the counties receiving the most visits are located in the most contentious states: Philadelphia County, Pennsylvania and Cuyahoga County, Ohio received 7 visits, while Allegheny County, Pennsylvania hosted 6 visits. Each of these counties contains a major metro area such as Philadelphia, Cleveland, and

Pittsburgh, respectively. An outlier exists, however, in the case of Washington, D.C. which hosted 6 different visits despite having among the strongest Democratic voting tendencies in past elections.

Among the counties with more than one visit, 22 of those 37 hosted both Secretary Clinton and President Trump at some point between September 1st and Election Day. With his considerably smaller campaign footprint, Governor Johnson visited 8 counties also visited by either of the major party campaigns.

Figure 5.1: Unique Counties Visited By Trump

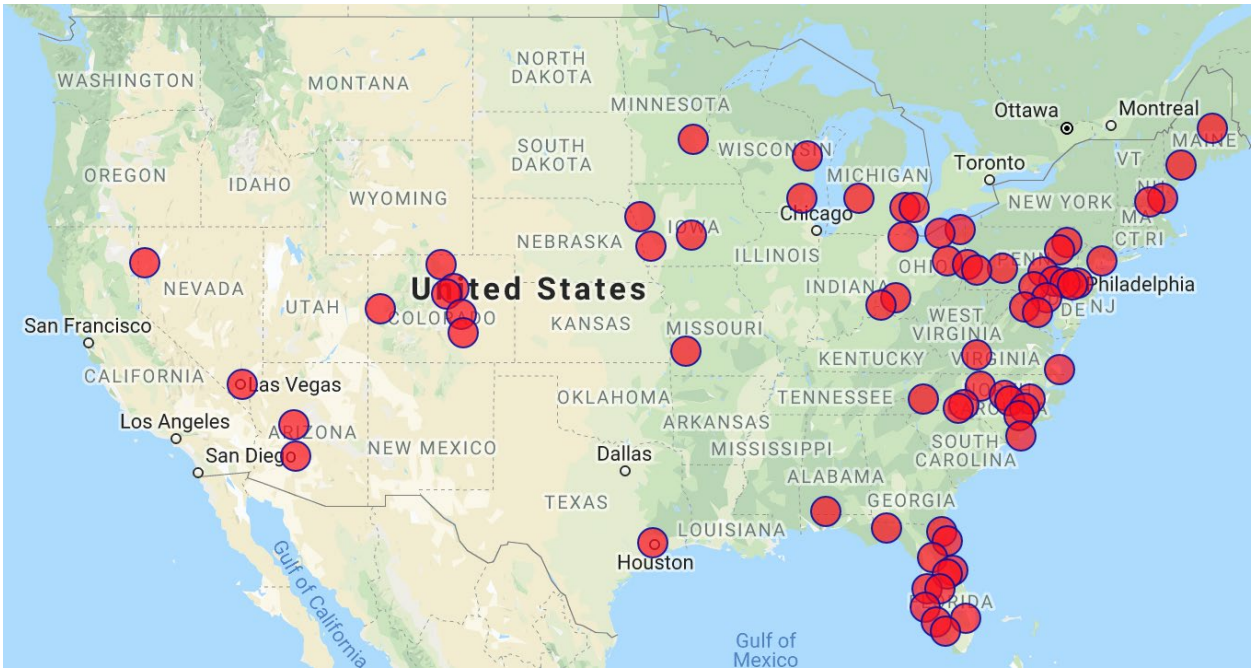


Figure 5.2: Unique Counties Visited by Clinton

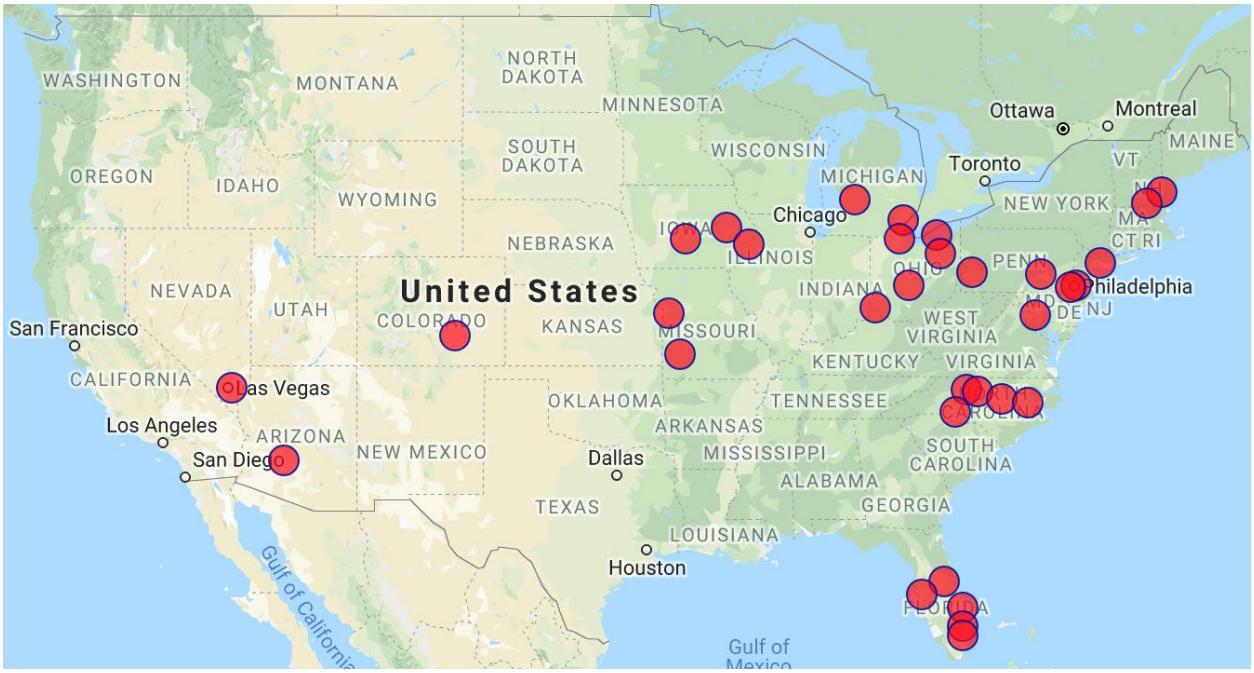
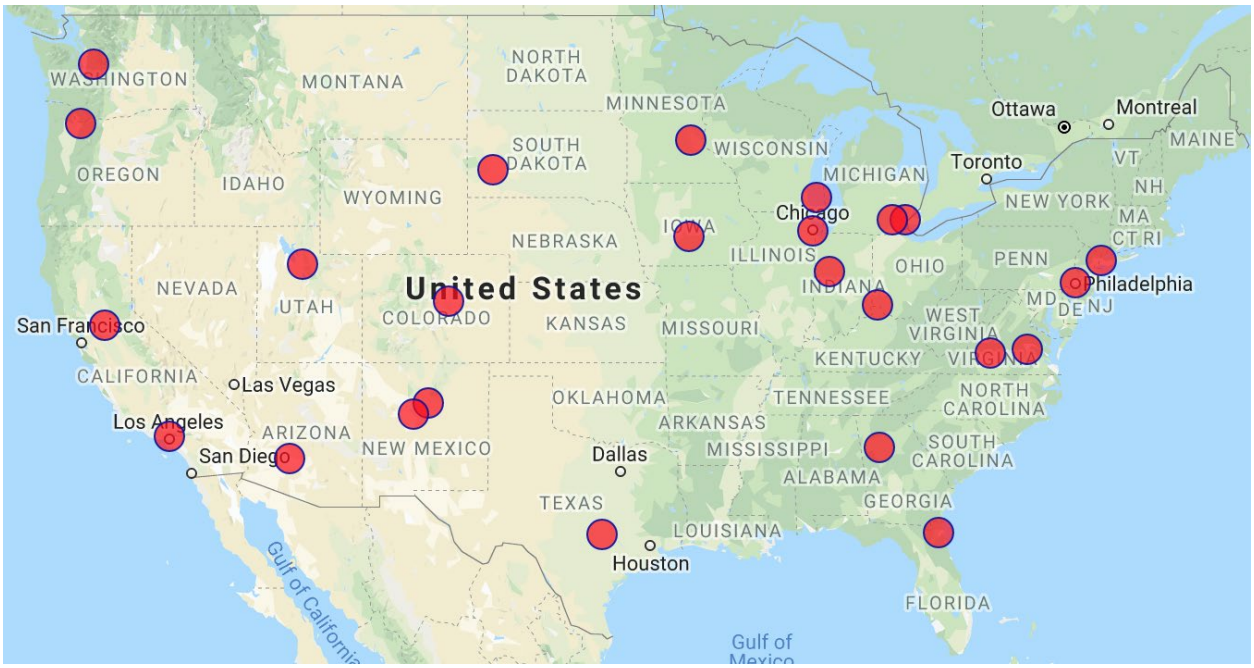


Figure 5.3: Unique Counties Visited by Johnson



Regression Models

The regression models produced several statistically significant findings. As a reminder, 3 regression models were run, with each successive model adding in more control variables as a means of checking for robustness.

Trump Vote Share Regressions

In regard to Trump's vote share, regression results are shown on the following page in Table 5.1.

Table 5.1: OLS Estimates of Relationship Between Candidate Visits and Trump Vote Share

Dependent Variable: Trump Vote Share

VARIABLES	(1) Model 1	(2) Model 2	(3) Model 3
Clinton Sept. Visits	-1.795*	-1.365**	-1.511**
	(0.935)	(0.674)	(0.690)
Clinton Oct. Visits	-0.164	0.00827	-0.0776
	(0.647)	(0.443)	(0.492)
Trump Sept. Visits	-1.940**	-0.847	-0.818
	(0.840)	(0.616)	(0.622)
Trump Oct. Visits	-2.197***	-1.240**	-1.467***
	(0.716)	(0.490)	(0.434)
Johnson Sept. Visits	-4.870**	-1.299	-1.378
	(1.966)	(1.502)	(1.499)
Johnson Oct. Visits	-7.876***	-4.475***	-4.425***
	(1.587)	(1.668)	(1.645)
Base Republican	1.003***	0.881***	0.882***
	(0.0316)	(0.0284)	(0.0280)
Median Income (In Thousands)		-0.267***	-0.267***
		(0.0513)	(0.0515)
White		0.165	0.162
		(0.298)	(0.298)
Black		-0.0106	-0.0117
		(0.298)	(0.298)
Asian		-0.544	-0.547
		(0.340)	(0.339)
Hispanic		-0.00611	-0.00909
		(0.300)	(0.300)
Native		0.0653	0.0645
		(0.318)	(0.318)
Clinton Ads (In Millions)			0.108
			(0.0834)
Trump Ads (In Millions)			-0.0456**
			(0.0192)
Constant	5.354***	6.698	6.927
	(1.949)	(29.88)	(29.87)
Observations	3,111	3,111	3,111
R-squared	0.857	0.911	0.912

Robust standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

The first model indicated that Trump's visits in September had a statistically significant negative relationship with his vote share, having a coefficient of -1.94 and a standard error of .84 (significant at the 95% level). This indicates a Trump visit in September corresponds with an estimated decrease in Trump's vote share of nearly 2 percentage points. In models 2 and 3, however, after adding in more controls, this relationship loses its statistical significance with coefficient estimates of -0.85 and -0.82, respectively for models 2 and 3. Trump's October visits also had a statistically significant negative relationship with his own vote share; a relationship that held its significance across all three models. Model 3, the most robust of the models, yielded a coefficient of -1.47 on Trump's October Visits, indicating a correspondence between a Trump visit in October and a decrease in his share of the vote by roughly 1 and a half percentage points. The model produced a standard error of .43 on the coefficient, indicating an estimate significant at the 99% level.

Generally, adding more demographic controls in Model 2 decreased the size of the relationship between Trump's October visits relative to that produced in Model 1; adding controls for advertising in Model 3 increased the coefficient slightly compared to Model 2.

Next, there existed a generally negative relationship between Clinton's September visits and Trump's vote share. Indeed, Model 2 indicates that a visit in September by Clinton is associated with a decrease in Trump's share of the vote by roughly 1.3 percentage points, an estimate significant at the 95% level. This is similarly found by the more robust third model, which yielded a coefficient of -1.51 (suggesting a visit by Clinton in September decreases Trump's share of the vote by an estimated point and a half) and a standard error of .69, also significant at the 95% level.

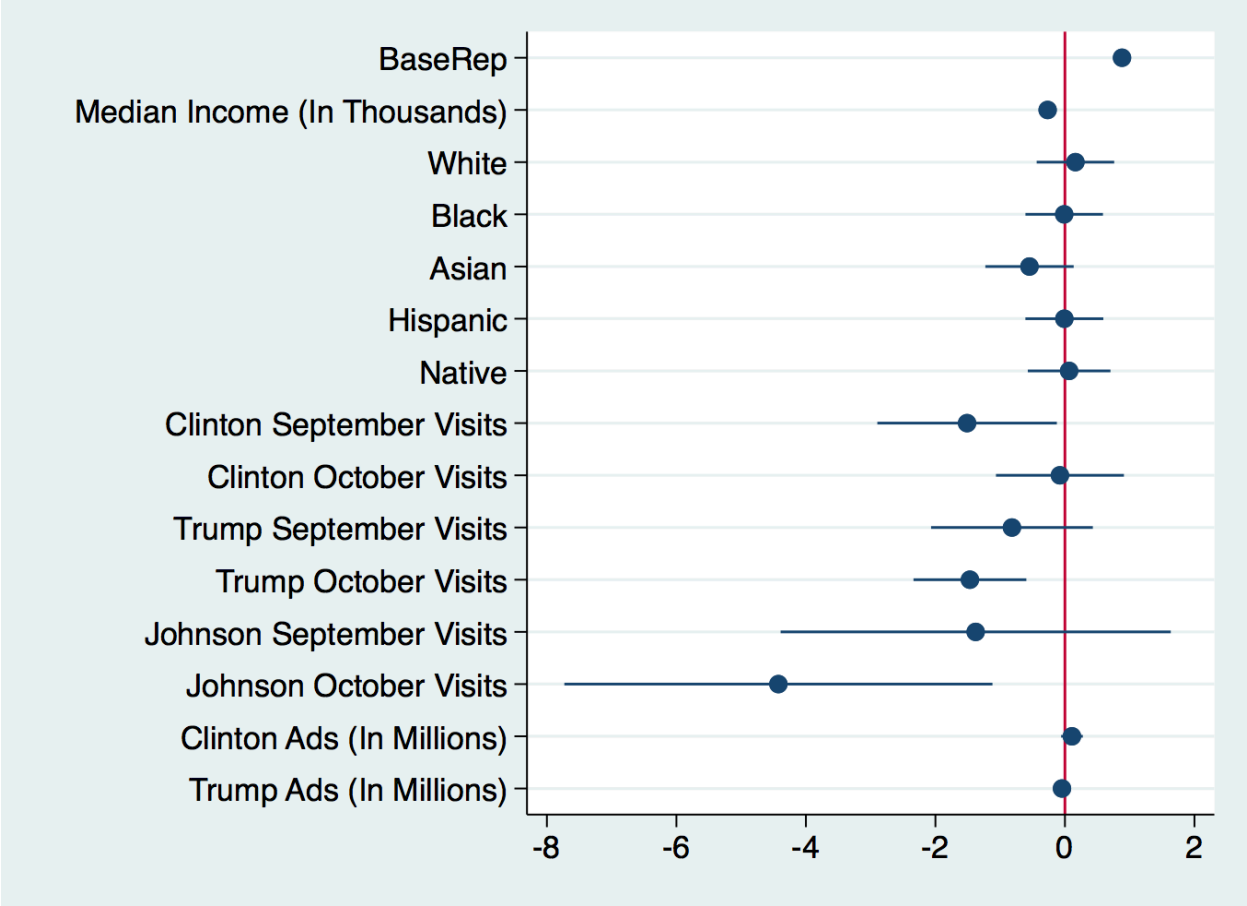
Adding more demographic controls in Model 2 and advertising controls in Model 3 caused some slight variation in the coefficients, though Models 2 and 3 produced estimates of the relationship between Clinton's September Visits and Trump's share of the vote that were more statistically significant than that shown by Model 1.

While Clinton's September visits seemed to have an effect on Trump's vote share, her October visits had no statistically significant effect across all three models.

Regarding Governor Johnson's visits, his September visits had a statistically significant relationship with Trump's vote share in the first model, though after adding in robustness through the subsequent models, this relationship lost its significance. His October visits were shown to have a statistically significant relationship with Trump's vote share, though, across all three models. Model 3 ultimately produced a coefficient of -4.425, with a standard error of 1.65, indicating significance at the 99% confidence level. This suggests a visit by Governor Johnson in October is associated with a decrease in Trump's vote share of nearly 4.5 percentage points.

The results of Model 3 in particular are illustrated in Figure 5.4, with each point plotted representing the coefficient on each variable, while each horizontal line plotted represents the 95% confidence interval which corresponds to those coefficients.

Figure 5.4: Visualization of Coefficients and Confidence Intervals Produced by Trump Model 3



Clinton Vote Share Regressions

Moving to the models run on Clinton's vote share, the general statistical output is shown below in Table 5.2.

Table 5.2: OLS Estimates of Relationship Between Candidate Visits and Clinton Vote Share

Dependent Variable: Clinton Vote Share			
VARIABLES	(1) Model 1	(2) Model 2	(3) Model 3
Clinton Sept. Visits	2.165**	1.361**	1.384**
	(0.928)	(0.627)	(0.590)
Clinton Oct. Visits	0.639	0.258	0.216
	(0.563)	(0.345)	(0.368)
Trump Sept. Visits	2.485***	1.293**	1.171**
	(0.840)	(0.562)	(0.487)
Trump Oct. Visits	2.130***	1.451***	1.280***
	(0.751)	(0.504)	(0.402)
Johnson Sept. Visits	4.420**	1.104	1.260
	(1.767)	(1.341)	(1.309)
Johnson Oct. Visits	5.788***	2.367***	2.493***
	(0.920)	(0.756)	(0.751)
Base Republican	-0.986***	-0.834***	-0.833***
	(0.0319)	(0.0199)	(0.0200)
Median Income (In Thousands)		0.254***	0.254***
		(0.0414)	(0.0414)
White		0.00917	0.00859
		(0.257)	(0.258)
Black		0.280	0.276
		(0.256)	(0.257)
Asian		0.731**	0.734**
		(0.284)	(0.286)
Hispanic		0.184	0.184
		(0.257)	(0.258)
Native		0.0968	0.0961
		(0.274)	(0.275)
Clinton Ads (In Millions)			-0.0274
			(0.0654)
Trump Ads (In Millions)			0.0803***
			(0.0152)
Constant	88.98***	68.12**	68.07**
	(2.092)	(25.59)	(25.72)
Observations	3,111	3,111	3,111
R-squared	0.861	0.942	0.943

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Each of the three models indicated a statistically significant relationship between Clinton's September visits and her own vote share. Specifically, the coefficient on Model 1 shows an estimated relationship between a Clinton September visit and an increase in her own vote share of about 2.2 percentage points, significant at the 95% level with a standard error of .93. Model 2 indicated a coefficient of 1.36 and standard error of .63 with the same level of significance as the prior model, showing an association between a Clinton September visit and an increase of her share of the vote of over 2 percentage points. Finally, the most robust of the models, Model 3, yielded a coefficient of 1.38 with a standard error of .59, showing significance at the 95% level and estimating a relationship between a Clinton visit in September and an increase of her vote share by around 1.4 percentage points.

Adding robustness through demographic controls in Model 2 decreased the size of estimated relationship relative to Model 1, while adding controls for advertising in Model 3 caused only a very small deviation of the coefficient in upward direction in comparison to Model 2.

Moving to Trump's visits and their relationship with Clinton's vote share, there was a statistically significant relationship across the board for both September and October visits produced by all 3 models.

Looking at Trump's September visits, Models 1 and 2 indicated a significant relationship at the 95% and 99% level, respectively. Model 3 yielded a coefficient on Trump September visits of 1.17 and a t-value of .49, significant at the 95% level and indicating that a Trump visit in September corresponds with an estimated increase in Clinton's share of the vote by slightly above 1 percentage point.

The size of the coefficient on the Trump September Visit variable decreased each time controls were added: there was a decrease of well over 1 percentage point between Models 1 and 2 and a slight decrease between Models 2 and 3.

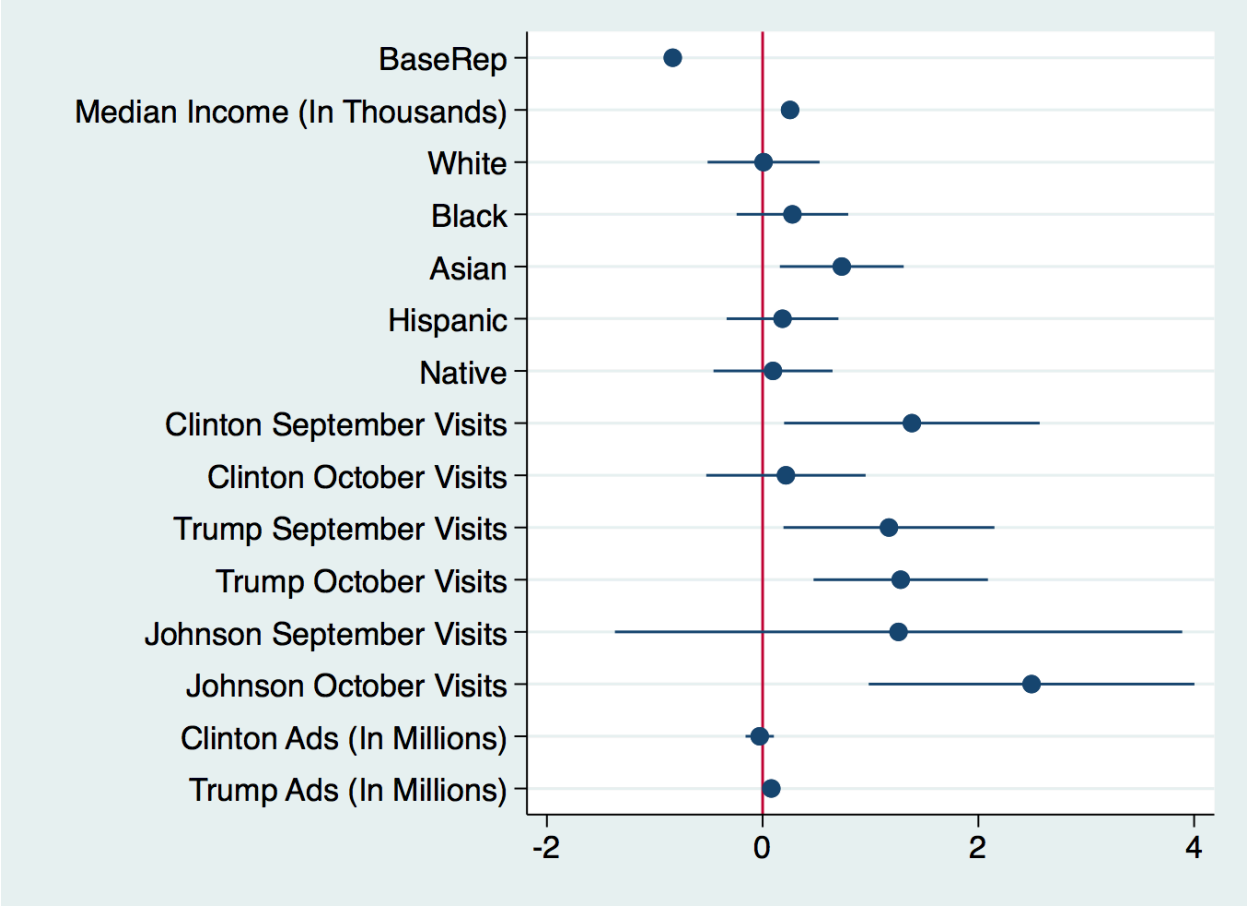
In regard to his October visits, Models 1 and 2 both showed statistically significant relationships, with confidence at the 99% level. This was similarly shown in Model 3, which yielded a coefficient of 1.28 and standard error of .40, implying a relationship between a Trump visit in October and an increase in Clinton's vote share by 1.28 percentage points.

As with Trump's September visits, the coefficient on his October visit variable decreased as controls were added in each model.

Finally, Johnson's September visits were shown to have no statistically significant relationship with Clinton's vote share other than in Model 2, where significance was found at the 95% level. Model 3, however, failed to corroborate this relationship, instead producing a standard error of 1.31, indicating no statistical significance.

The results of Model 3 in particular are shown below in Figure 5.5.

Figure 5.5: Visualization of Coefficients and Confidence Intervals Produced by Clinton Model 3



Johnson Vote Share Regressions

The regressions focusing on Governor Johnson's vote share, as shown below in Table 5.3, provided outputs regarding visits that were generally statistically insignificant.

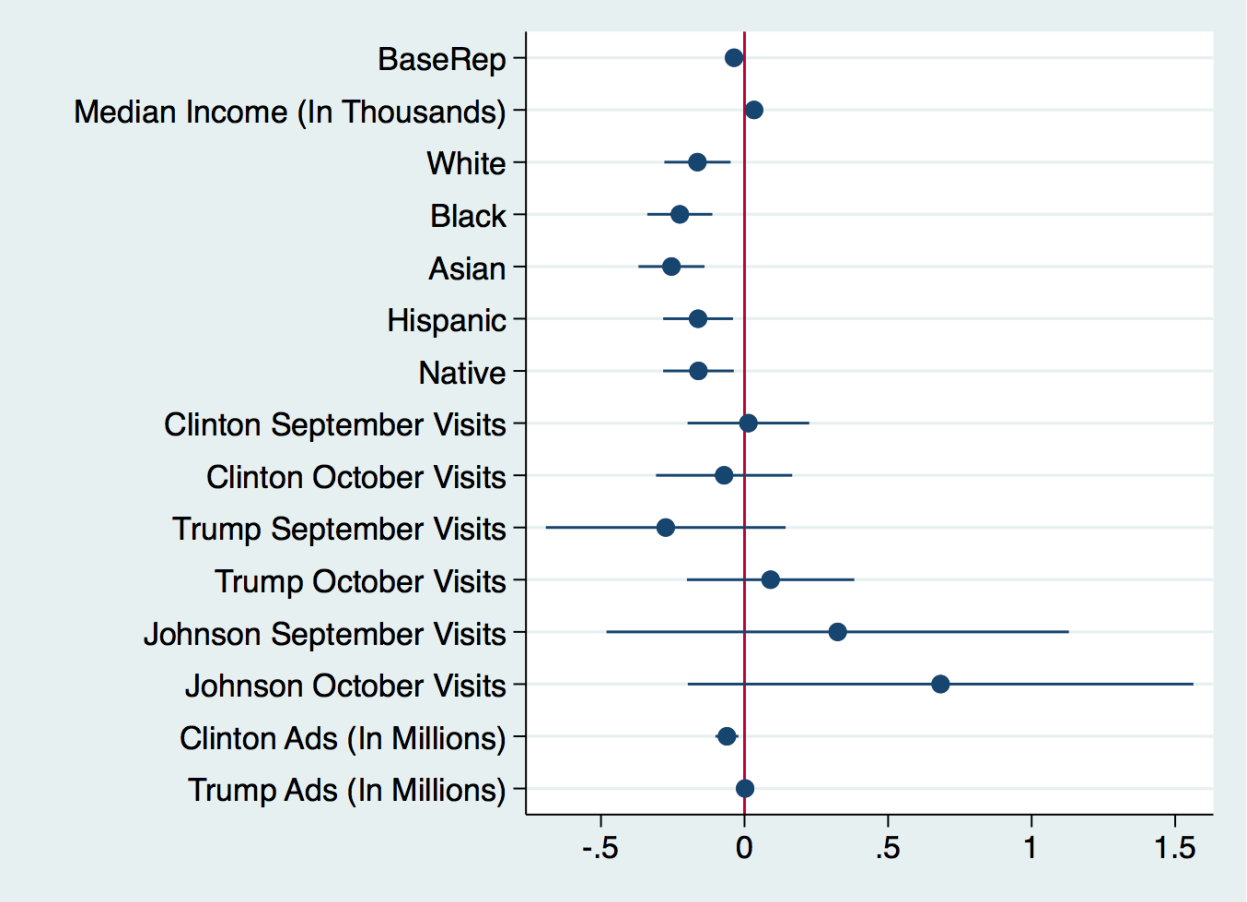
Table 5.3: OLS Estimates of Relationship Between Candidate Visits and Johnson Vote Share

Dependent Variable: Johnson Vote Share			
	(1)	(2)	(3)
VARIABLES	Model 1	Model 2	Model 3
Clinton Sept. Visits	-0.331*	-0.0744	0.0134
	(0.190)	(0.157)	(0.105)
Clinton Oct. Visits	-0.271**	-0.142	-0.0711
	(0.105)	(0.105)	(0.118)
Trump Sept. Visits	-0.357*	-0.298	-0.274
	(0.187)	(0.204)	(0.208)
Trump Oct. Visits	0.0895	-0.117	0.0908
	(0.230)	(0.194)	(0.145)
Johnson Sept. Visits	0.417	0.327	0.325
	(0.487)	(0.386)	(0.401)
Johnson Oct. Visits	0.680	0.760*	0.683
	(0.550)	(0.439)	(0.438)
Base Republican	-0.0138*	-0.0358***	-0.0365***
	(0.00733)	(0.00570)	(0.00572)
Median Income (In Thousands)		0.0328***	0.0332***
		(0.0119)	(0.0123)
White		-0.166***	-0.164***
		(0.0582)	(0.0574)
Black		-0.228***	-0.225***
		(0.0571)	(0.0564)
Asian		-0.255***	-0.254***
		(0.0583)	(0.0572)
Hispanic		-0.164***	-0.162**
		(0.0610)	(0.0605)
Native		-0.161**	-0.160**
		(0.0619)	(0.0612)
Clinton Ads (In Millions)			-0.0613***
			(0.0199)
Trump Ads (In Millions)			0.00203
			(0.00511)
Constant	3.971***	21.40***	21.27***
	(0.502)	(5.713)	(5.639)
Observations	3,111	3,111	3,111
R-squared	0.020	0.377	0.394

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Governor Johnson's September and October Visits yielded coefficients with standard errors of .40 and .38 in Model 3, respectively, indicating no statistical significance in both cases. Additionally, the model produced insignificant coefficients on Trump's September visits and on Trump's October visits. Finally, Secretary Clinton's visits in September and October were found to not have a significant relationship with Governor Johnson's vote share. The results of Model 3 are represented below in Figure 5.6.

Figure 5.6:Visualization of Coefficients and Confidence Intervals Produced by Johnson Model 3



Matching Model Results

As described in the Methodology, six Mahalanobis nearest-neighbor matching models (with replacement) were run, using a dummy variable for each candidate, broken down by month, as the treatment. The Average Treatment Effect on the Treated (ATT) produced by the model in each case are presented below in Table 5.4.

Table 5.4: Estimates of ATT for Visits on Vote Share

VARIABLES	(1) Trump Sept ATT	(2) Trump Oct ATT	(3) Clinton Sept ATT	(4) Clinton Oct ATT	(5) Johnson Sept ATT	(6) Johnson Oct ATT
Trump Vote Share						
Trump September Dummy	-0.909					
	(0.359)					
Trump October Dummy		0.959				
		(0.609)				
Clinton Vote Share						
Clinton September Dummy			10.45			
			(1.569)			
Clinton October Dummy				6.218**		
				(2.001)		
Johnson Vote Share						
Johnson September Dummy					0.0871	
					(0.157)	
Johnson October Dummy						0.334
						(0.672)
Observations	3,111	3,111	3,111	3,111	3,111	3,111

z-statistics in parentheses
*** p<0.01, ** p<0.05, * p<0.1

The first matching model, looking at the effect of the Trump September Visit Dummy, produces an ATT which is not statistically significant. Next, the estimated ATT of the Trump October Visit Dummy also is not statistically significant.

Moving on to Secretary Clinton's vote share, her September Visit Dummy was shown to have a statistically not-significant effect. Her October Visit Dummy, however, had significance at the 95% level, with an ATT of 6.22 percentage points on her share of the vote. The respective 95% confidence intervals of these estimates are [-2.6, 23.5] for September and [.13, 12.3] for October.

Finally, neither of the models pertaining to Governor Johnson's September and October models produced ATTs with statistical significance.

Methodological limitations

One possible limitation of this analysis is rooted in the fact that it treats counties as units isolated from the goings on in neighboring counties. A visit by a candidate in Miami-Dade County, Florida likely has some sort of spillover effects into neighboring Broward County, Florida, which did not receive any visits. The effect of this would be to bias downward the observed effect of visits. If both Broward and Dade counties receive an equal effect from a visit that took place in Dade, only Dade County is treated by the model as having actually received a visit. As such, the potential difference in vote share observed between counties that received visits and those that did not would be less than if this were controlled for. This may have had somewhat of an effect on the results of the regression models but would only have an effect on the matching models if treated counties were coincidentally matched with counties that neighbored visited counties.

Next, while the regression models took into account how many times a candidate visited a county, the matching model cannot similarly take this into account, as a product of the fact the treatment variable must be binary in the latter case. Rather than examining the effect of the

number of visits, the matching model examined the effect of whether or not a candidate visited a county at all. This could present some issues, as a county receiving 3 visits is treated the same, statistically, as a county receiving 1 visit. This does not represent a particularly large concern in terms of the results, however, as both the median and modal number of visits among visited counties was 1.

Another potential limitation is the fact that some independent variables, such as Base Republican, and the demographic controls are highly multicollinear, which introduces imprecision on those variables' coefficients. This does not represent much of a concern, however, as we are not very interested in precise coefficients on the control variables, but just care that they exist as controls when looking at the effect of visits on vote share.

Next, the set of advertising data used to control for other forms of a campaign's influence over voters did not correspond perfectly to the September 1 to Election Day timeframe that visits were examined over. Rather, the only comprehensive advertising dataset that could be found at the time the analysis was conducted covered the last few weeks of the election. The concern here is also not particularly large, as the vast majority of advertising takes place in the final days of the election and would also not represent an issue unless the proportion of spending in each state relative to another changed significantly between the beginning of September and the beginning of the dataset's timeframe - which is unlikely to have occurred. Additionally, the advertising data served to add robustness between Models 2 and 3, with the inclusion of the advertising variables being the sole difference. The regression models did not show results that significantly differed upon adding these variables in, diminishing concern.

Finally, while every effort was made to control for omitted variable bias through the inclusion of a variety of covariates, the complexity of measuring voting behavior is such that

there is always the possibility of having missed a very specific, yet important variable. In this vein, there also exists the possibility that some determinant of voting behavior or factor that influences a candidate's choice of where to visit is simply unobservable. While my results do not seem to show anything particularly unusual, there cannot be 100% certainty that there do not exist some unsolvable biases.

Chapter 6

Discussion of Results

Review of Results

Clinton's visits seemed to have helped her, while visits by Trump and Johnson did not help either of their respective numbers at the polls. The set of regressions suggested a negative relationship between Trump's October visits and his own vote share in a given county, a positive relationship between Clinton's September visits and her own share of a county's votes, and no relationship between Johnson's visits and his own share of the votes in a county. Additionally, there was a negative relationship between Trump's vote share and both Clinton's September Visits and Johnson's October visits. A positive relationship was observed between Clinton's vote share and both Trump's September and October visits, as well as between her share of the vote and Johnson's October visits.

The matching models showed results which differed slightly from the regression models in terms of statistical significance but did not deviate a concerning amount in terms of the output. A positive effect of Clinton's October visits on her share of the votes in counties which she visited was observed while no statistically significant effect of Trump's visits or Johnson's visits on their own shares of the vote in counties they visited was seen in the matching model.

Implications of Results

The findings of the matching model and the regression models provide support for Hypothesis 1, which predicts that number of visits has a positive effect on a candidate's share of the vote in a given county, in the case of Secretary Clinton's visits. The hypothesis is not supported by either model in the cases of visits by Trump and Johnson.

Additionally, while the output of the matching models imply Secretary Clinton's October visits had an effect on her share of the vote and her September visits did not, we cannot comfortably say that her October visits had a more significant effect than her September visits, statistically speaking, due to overlap in the 95% confidence intervals of the estimates produced by those models. As a result, support for Hypothesis 2, that visits later in the campaign season have more of an effect on a candidate's share of the vote, cannot be concluded.

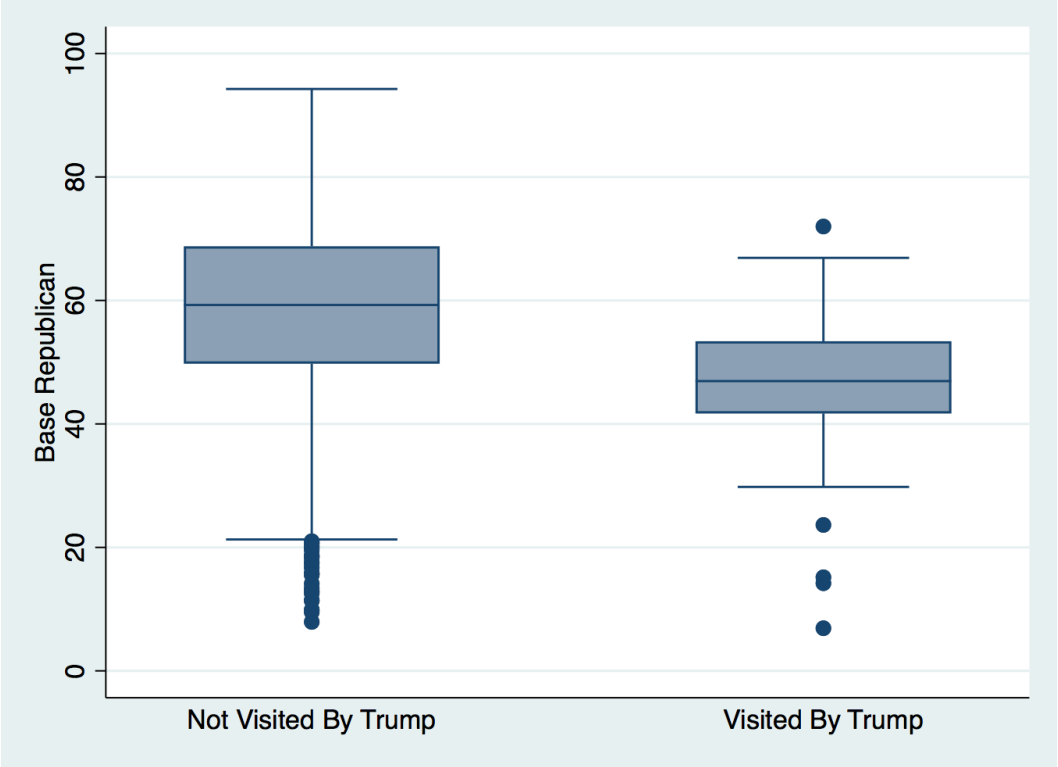
The discrepancies between the regression models and the matching models also deserve discussion, as these differences can provide insight as to the strategies of each campaign.

In the case of Trump, the regression models produced a negative relationship between his visits and his own vote share on the county level. Importantly, the matching models, which help control for endogeneity in the location of visits did show a negative effect for September and a positive effect for October, though not precisely estimated.

This is perhaps because Trump tended to visit locations that he was generally less popular in, producing the negative correlation in the regressions, but not showing an effect when taking measures to control for endogeneity. A trend can be seen in Figure 6.1, which shows the distribution of the Base Republican variable among the counties Trump visited, plotted next to the distribution for counties he did not visit. It is, in fact, the case that Trump tended to visit counties that were less Republican, on average, than the counties he did not visit.

There could be intentional reasons for this pattern. For example, perhaps the Trump campaign deliberately sought to visit counties where he needed to make up ground. Additionally, Trump seemed to be planning on launching a television network after the election and possibly visited locations that could have been useful for that purpose.

Figure 6.1: Distribution of Base Republican Variable for Trump, by County

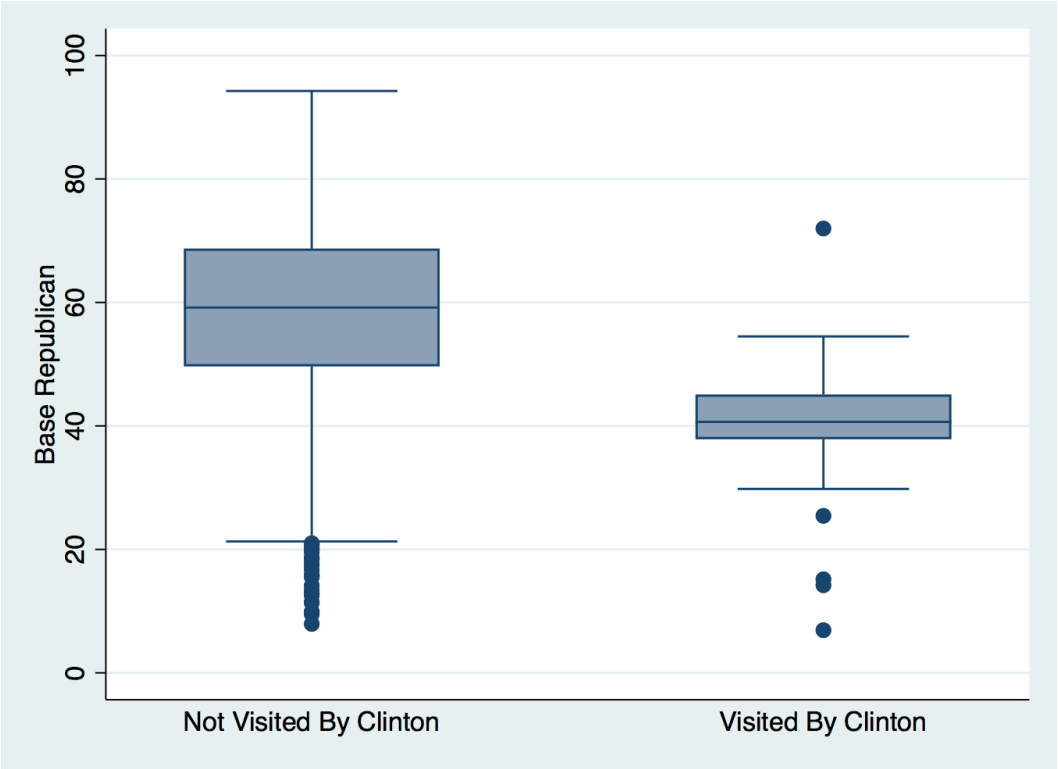


While the conventional wisdom may suggest that Trump liked to address large, adoring crowds, this is not mutually exclusive with hosting events in counties he was less popular in, as his events were still attended by thousands of supporters, who ultimately only represent a drop in the bucket of many counties' populations. Additionally, while it appears Trump may have visited locations he is less popular in, the territory is unlikely to be completed filled with his opponents' supporters and devoid of his own, as most of the visits took place in swing states, where levels of support between candidates do not differ as significantly as in a solid blue state like California, for example. That said, the Trump September model did indicate a negative coefficient, as the

Trump Regression did, though in this case the estimate was not precise enough to be statistically significant.

A similar phenomenon was observed for Secretary Clinton, in the opposite direction. While her September visits were shown by the regressions to be positively correlated with her own vote share, the matching model did not show a statistically significant effect (though it did indicate a positive coefficient, imprecisely estimated). The Clinton October matching model did produce a positive ATT significant at the 95% level, however. Figure 6.2, on the following page, shows the distribution of the Base Republican variable based on Clinton's visits, indicating that she tended to visit counties that were more Democratic than the counties she did not visit, on average. This could reflect a desire to energize a Democratic voting base that may have seen the election as already won, in hopes that visits would increase turnout of Clinton's supporters.

Figure 6.2: Distribution of Base Republican Variable for Clinton, by County



More generally, it is important to note what these results mean. First, the coefficients produced by the regression models and the matching models represent slightly different things. With a regression analysis, each coefficient on the visit variables correspond to an association between increasing the number of that type of visits by 1 and a change in the vote share for one of the candidates (specified by the dependent variable).

The matching model, however, requires a binary treatment variable, rather than a continuous variable like in the regressions. Because of this, the visit variables take the form of a 1 for a county a candidate visited at all (regardless of if they visited 1 time or 4 times, the variable would still be a 1 in each case, for example), and a 0 in any county a candidate did not visit.

As the matching model examines the Average Treatment Effect on the Treated (ATT), the output simply helps us understand the effect of visits on counties that were actually visited. We cannot conclude, however, that if Clinton had visited more counties in general, that she would have been able to enjoy the same positive effects of her visits in all those counties. As was has been discussed throughout this thesis, the locations to visit are chosen for a reason and not randomly.

It could be the case, though, that if Clinton had visited more counties similar to those that she did visit that she might have won over voters in those counties. Indeed, the populations in swing states, as well as the counties that constitute those states, have some degree of similarity across them that causes campaigns to deem them worth visiting. This suggests that critics of her strategy to pass over vital swing states such as Michigan, Pennsylvania, and Wisconsin were possibly correct in claiming that this decision cost her the election, as the counties she passed over in these states likely had similarity to other counties she did visit.

Given how razor-thin the margins across these three states and considering that had Clinton won these three states she would have been elected, this becomes even more likely. As Clinton herself noted in her book *What Happened,* "if just 40,000 people across Wisconsin, Michigan, and Pennsylvania had changed their minds, I would have won," though Clinton, expectedly argued that the campaign was doing everything they can to win in those states (Smith 2017).

For example, Clinton lost Wisconsin by roughly 27,000 votes. As the matching model estimated, Clinton's October visits had a positive effect on her share of the vote of roughly 6.2 percentage points. This suggests that had Clinton visited Milwaukee County in October, holding all else equal, she could have increased the number of votes she received in that county by

around 27,000 votes. In addition, by visiting another, smaller county such as Jefferson County in October, she could have picked up approximately 2,600 more votes, putting her well into winning territory for the state as a whole.

This can also be seen in Michigan, where Clinton lost to Trump by around 11,600 votes. Based on the estimates produced by the matching model visiting Kent County, home to Grand Rapids, in October could have increased Clinton's vote total by roughly a predicted 30,000 ballots.

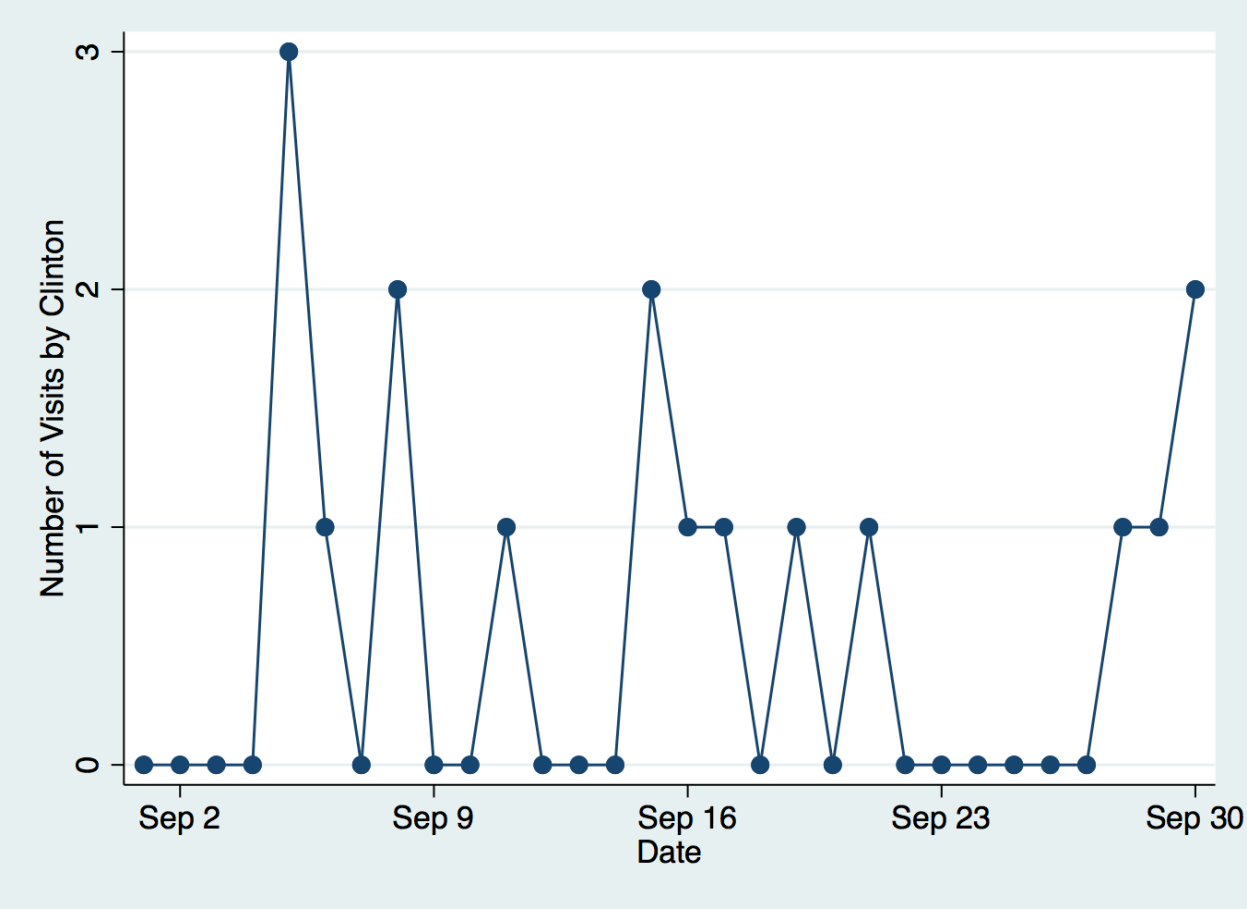
The fact that the models imply Clinton's visits had an effect, while Trump's did not, also helps shed light upon the fact that the effects of in-person visit vary based on a variety of factors. For example, Trump's off the cuff, anything goes style of speaking would have resonated with undecided voters in a different way than Clinton's more traditional manner of public appeal. Further, Trump's openly antagonistic relationship with the press also could have led to worse coverage of his visits in local media and negatively affected the perception of the visit among local voters. Additionally, at a more general level, the Clinton campaign was extremely well organized relative to the more chaotic Trump campaign, which could have also made a difference.

This lends support to Wood's conclusion that visits differ by candidate. Similarly, Herr's study of the 1996 election noted that differences in the candidates themselves led to different effects of their visits. While Bill Clinton enjoyed an effect of his visits on the state level and had a particularly charismatic presence, Bob Dole's visits did not have such an effect, and, in fact, was known for his lack of charisma.

Given the fact that Clinton's visits are observed to have a positive effect, it is somewhat paradoxical, then, that she elected to not take more visits. The discrepancy between Clinton and

Trump's number of campaign stops in September is almost certainly due, in part, to Clinton's case of pneumonia that prevented her from hitting the road ambitiously that month. This can be seen in Figure 6.1, which plots Clinton's September visits over time. After Clinton collapsed at Ground Zero on September 11th, she took a break from the campaign trail until September 15th, followed by a longer break - beginning on the 22nd and ending on 27th - as she recovered from her illness.

Figure 6.3: Number of Clinton September Visits by Date



Still, though, Clinton made 22 fewer visits in October than Trump did, even despite her health returning to normal well before the start of that month. Interestingly, according to

Democratic pollster Stan Greenberg, the Clinton campaign elected to wind down the most thorough of its polling operations, known as "message polling," which is helpful for pinpointing attitudes of voters in particular locations. Rather, the campaign instead relied solely on analytics polling, which helps create much simpler information such as horse race data and general candidate favorability. Further, the campaign elected to poll across swing states, treated as one single polling unit, rather than poll on a level of aggregation that was state-specific.
For this reason, "the Clinton camp may have missed the underlying warning signs in key states like Michigan, Pennsylvania, and Wisconsin" (Shepard, et al. 2017). Indeed, it is possible that the discrepancy in visits could be a result of a lack of complete information regarding the electoral situation.

On the other hand, Trump's visits appeared to have no effect, unlike Clinton's. It is interesting to consider, then, that perhaps some candidates are better off putting more effort into private fundraising events than campaign rallies, if there appears to be no effect for them. While tempting to engage in such thought puzzles, we are unfortunately stuck in the quagmire of an unobservable counterfactual.

It could be the case, for example, that Trump's visits did have an effect, but Clinton's visits cancelled out such effects. Had Trump taken significantly fewer visits, he likely would have been subject to criticism by the public and media alike for being out of touch with the average American, as shown by his lack of a desire to visit voters' locales. Unfortunately, these are things that models cannot control for. Similarly, if both candidates' visits largely cancel each other out in multiple counties, the result shown by our models is no or a diminished effect.

When observing a less extreme counterfactual, wherein a candidate who lacks charisma is deciding between slightly fewer campaign visits and instead devoting resources to slightly

more fundraisers and more advertising, this could represent an attractive option. This is especially true as campaigns move further into the virtual world, using specifically targeted, advanced advertising techniques that have proven to represent significant value on a cost basis. Indeed, Trump's campaign took advantage of micro-targeted Facebook ads, that high-level campaign officials claim won them the election (Stanage 2016).

Chapter 7

Conclusions

In conclusion, after performing a comprehensive data analysis, the results of this paper imply that Secretary Clinton's visits had a statistically significant effect on her share of the vote on the county level. The models found that President Trump and Governor Johnson's visits did not have a statistically significant effect on their share of the vote on the county level. Additionally, the study found that there does not exist statistically significant support for the prediction that visits later in the campaign have more of an effect, however.

This paper represents the first research on the effect of visits in presidential campaigns completed using final vote totals and such a specific level of aggregation; it helps shed light onto the effects of a complex, and extremely resource intensive aspect of each presidential election cycle. Further, it is the first work to investigate this phenomenon as it relates to the 2016 election.

Ultimately, it was determined that campaign visits can have an effect for some candidates, but not for others.

Perhaps the biggest implication, though, is that if Hillary visited more counties in October similar to the others she had visited, she may have won the election. This is especially salient given how thin the margins in key states were and gives credence to the argument of critics who declared that a lack of coverage of some swing states may have cost her the election.

While this research represents a step forward for literature on campaign visits, there exist identifiable recommendations for future researchers. It would be particularly useful to perform a more advanced statistical analysis based on the proximity of counties that did not receive visits with counties that did. In such a case, it would be possible to solve the aforementioned potential

issue of biasing the effects of visits downward. Additionally, such a study would allow for the examination of how far-reaching of an effect campaign visits have on the surrounding area. The findings in that case would be particularly helpful for campaign strategy.

Additionally, as my work highlights that some candidates' visits have an effect on their share of vote, while others do not, past studies have similarly found the same phenomena occurring. Future researchers may consider conducting a meta-analysis of various campaigns and look to identify which types of candidates tend to have impactful visits. It could be the case that candidates belonging to one party in particular have larger electoral effects of visits, or that the charisma of candidates plays a particularly important role.

In an ideal world, a researcher is able to perform a randomized experiment on the presidential campaign level, as has been performed on lower levels, such as gubernatorial elections. While unlikely to be accomplished, such a study would represent the first of its kind on the presidential scale.

Finally, it would be useful to perform a robust statistical analysis investigating how campaigns decide where to visit, as this would help solve the endogeneity problem that can make research in this area so difficult.

Bibliography

Beckett, Lois. “Trump Digital Director Says Facebook Helped Win the White House.” *The Guardian*, Guardian News and Media, 9 Oct. 2017

Bowman, Karlyn. “When People Make Up Their Minds about Their Presidential Vote.” American Enterprise Institute, American Enterprise Institute, 5 Aug. 1996,

Druckman, J. N. (2004), Priming the Vote: Campaign Effects in a U.S. Senate Election. Political Psychology, 25: 577–594.

Dumenco, Simon. *The States Where Clinton and Trump Are Advertising the Most Until Election Day*. Ad Age, 21 Oct. 2016

Herr J. Paul 2002. The impact of campaign appearances in the 1996 election. The Journal of Politics 64 (3): 904–13.

Hill Jeffrey S., Rodriquez Elaine E., Wooden Amada E. 2010. Stump speeches and road trips: The impact of state campaign appearances in presidential elections. PS: Political Science & Politics, 43 (2): 243–54.

Hillygus, D. S. 2004, Models of Voting in Presidential Elections: The 2000 U.S. Election. Presidential Studies Quarterly, 34: 691–693.

Holbrook Thomas. 1994. Campaigns, national conditions, and U.S. presidential elections. American Journal of Political Science 38 (4): 973–98.

Holbrook Thomas, McClurg Scott. 2005. The mobilization of core supporters: Campaigns, turnout, and electoral composition in United States presidential elections. American Journal of Political Science 49 (4): 689–703.

Isenstein, Libby, et al. “NJ Travel Tracker.” Nj-Travel-Tracker.herokuapp.com, 2016.

“Presidential Election Results: Donald J. Trump Wins.” The New York Times, The New York Times, 9 Aug. 2017

Kirkegaard, Emil O. W. Open Quantitative Sociology and Political Science. Open Quantitative Sociology and Political Science, 23 May 2016.

Lapowsky, Issie. “What Facebook Isn't Saying About Trump and Clinton's Campaign Ads.” Wired, Conde Nast, 28 Feb. 2018

Migliozzi, Blacki, et al. “Clinton and Trump Have Very Different Strategies for the Final Stretch.” Bloomberg.com, Bloomberg, 5 Oct. 2016,

Shaw Daron. R. 1999. A study of presidential campaign event effects from 1952 to 1992. The Journal of Politics 61 (2): 387–422.

Shaw Daron, Gimpel Jim. 2012. What if we randomize the governor’s schedule? Evidence on campaign appearance effects from a Texas field experiment. Political Communication 29 (2): 137–59.

Shepard, Steven, et al. “Did the Clinton Campaign Really Stop Polling at the End?” POLITICO, 29 Sept. 2017

Smith, Allan. “Hillary Clinton Disputes One of the Most Common Critiques of Her Campaign in Her New Book.” Business Insider, Business Insider, 13 Sept. 2017

Stanage, Niall. “Will Trump or Clinton Win? The 11 States Deciding the Race.” *TheHill*, 7 Sept. 2016

Wicks, Buffy. “Hillary Clinton: How She Is Going to Win Today.” Time, Time, 8 Nov. 2016

Wood, Thomas. “What The Heck Are We Doing in Ottumwa, Anyway? Presidential Candidate Visits and Their Political Consequence.” *The ANNALS of the American Academy of Political and Social Science*, vol. 667, no. 1, 2016, pp. 110–125.