

# A Look at the 2012 Presidential Election

Eric Hare & Andee Kaplan

Dec. 12, 2012

## Abstract

Motivated by the 2010 Citizens United ruling and the subsequent birth of “Super PACs”, this paper uses independent expenditures data from the Federal Elections Commission, in conjunction with presidential polling data to analyze the 2012 presidential campaign. Using R, and the packages `XML`, `ggplot2`, `plyr`, `lubridate`, `knitr`, `forecast`, and `zoo`, we scrape data from these sources and analyze them in order to highlight interesting trends in campaign spending. Furthermore, we correlate these trends in spending over time to the changes in the polls. Ultimately, we did not find a clear and direct relationship between increases in spending and changes in public support. However, our analysis does provide evidence for some commonly held views of the candidate’s spending habits and geographical areas of strength and weakness.

# 1 Introduction

In 2010, the United States Supreme Court released its decision on *Citizens United v. Federal Election Commission*. The decision found that it is a violation of the constitutional right to free speech for limits to be placed on money spent by corporations and labor unions to support or oppose political candidates. For the first time since the passage of the Tillman Act in 1907, corporations and unions could spend unlimited amounts of their own money on the presidential election. Such power led to the coining of the term Super PAC, as well as a cloud of uncertainty as to what effect such money could have on the election outcome.

In this report, we analyze the independent expenditures data set from the Federal Election Commission in an attempt to determine the effect of this spending on the 2012 election. We subset these expenditures to analyze only those which were filed after the Republican primary, so that we can focus on those which benefit President Barack Obama and Governor Mitt Romney. We also retrieve data from the NationalPolls.com database of polling results to quantify the effect of this spending. The results of our analysis suggest that the spending may have had a measurable impact on public opinion, but external factors, such as the presidential debates, had much stronger impacts.

## 2 Data

We are primarily working with two data sources to complete our analysis: independent expenditures (PAC) data and polling data. The PAC spending data comes from the Federal Election Commission and the polling data comes from <http://nationalpolls.com/>. We are primarily working with the fields `spe_nam`, `exp_amo`, `exp_dat`, and `pur` from the spending data. These provide us with information on the organization filing the expense, the amount, the date, and the purpose of the expense. We also make extensive use of our created field ‘`beneful_can`’, which describes the candidate benefiting from the expense. A full description of the fields in our dataset are provided in the following table.

### 2.1 Variable Definitions

The variables present in the PAC spending data are defined as follows:

Tag	Field Name	Data Type	Description	Explanation
can_id	Candidate ID	Character	unique ID of candidate for or against whom the expenditure was made	First character indicates office sought - H=House, S=Senate, P=Presidential. Columns 3-4 are the state abbreviation for Congressional candidates. NOTE - this information is provided by filers and may be missing - in these cases of- fice, state, district and candidate name should appear.
spe_id	Spender ID	Character	Unique ID of committee, individual or group making expenditure	Unique FEC ID assigned to the entity submitting reports of independent ex- penditures
spe_nam	Spender Name	Character	Name of committee, individual or group making expenditure	
ele_typ	Election Type	Character	code for specific election for which ex- penditure was made	First character indicates election - P=Primary, G=General, S=Special. Next four characters indicate election year.
can_off_sta	Candidate State	Character	Postal state abbreviation for the candi- date	
can_off_dis	Candidate District	Number	District number for the candidate	District location if spending for/against House candidate.
can_off	Office	Character	Office Sought by Candidate	(H=House, S=Senate, P=President)
can_par_aff	Party	Character	Party abbreviation for candidate	Dem=Democrat Rep=Republican
exp_amo	Expenditure Amount	Currency	Dollar amount of specific expenditure	Min: \$0.00 Max: \$17661251.08 Mean: \$11306.42 Median: \$26.61
exp_dat	Expenditure date	Date	Date of specific Expenditure MM/D- D/YYYY	

agg_amo	Aggregate amount	Currency	Total amount expended during the calendar year, per election, per office sought	Min: \$-39124.40 Max: \$161321977.00 Mean: \$514486.68 Median: \$202014.82
sup_opp	Support or Oppose	Character	S=Support, O=Oppose	Describes whether the expenditure was made to support or oppose the candidate.
pur	Purpose of expenditure	Character	description of the expenditure, e.g. television or radio ad	
pay	name of payee	Character	Name of the person or vendor or other entity receiving this payment	
file_num	Filing number	Number	Unique identifier for a submission (which may report several disbursements)	
amn_ind	Amendment Indicator	Character	New report or amendment to a report	
tra_id	Transaction ID	Character	Unique identifier for the transaction (unique within the specific filing)	
ima_num	Image number	Number	Image location for page on which transaction appears	
rec_dt	Filing receipt date	Date	Date on which transaction was submitted to FEC MM/DD/YYYY	
prev_file_num	Previous filing number	Number	Reference to a filing being amended	For electronic filings the previous filing number references the filing being amended. For new filings and paper filings this field will be null

The variables present in the polling data are defined as follows:

Tag	Field Name	Data Type	Description
Pollster	Polling Company	Character	Company that conducted the poll.
State/US	State	Character	State poll was conducted of. If national poll, then value is “National”.
Date	Poll Date	Date	Range of dates that the poll was being conducted.
Obama	Support for Mr. Obama	Number	Integer rounded percent of support in the poll.
Romney	Support for Mr. Romney	Number	Integer rounded percent of support in the poll.

## 2.2 Cleaning

The majority of the data cleaning had to be performed on the PAC spending data. We made sure all the columns were formatted correctly in regards to data type (dates are date, numbers are numeric, etc.) before cleaning and reformatting the data more extensively.

One challenge we faced was that the purpose of independent expenditures column is a free text field on the FEC reporting form. For example, the Conservative Super PAC Americans for Prosperity tended to use verbose descriptions of the spending purpose. An entry from mid-August read “oppose advertising-tv production (voted for).” This description lists whether it was an ad in support or opposition, that it was a television ad, and the name of the ad. In contrast, the anti-abortion Women Speak Out PAC preferred short descriptions, such as “ads” for their early-October expense. The result was that when trying to explore what PACs spent the majority of their money on, we were unable to group expenditures together. To solve this issue we searched for matches to patterns in the purpose field. We chose these patterns by looking at the expenditure purposes and manually finding common threads among the purposes. From these patterns we were able to create buckets that each expense fell in to, as well as high level buckets that more generally classified the expenditures. The buckets we chose are:

**Ads** Advertisement spending, including television, radio, and online

**Direct Contact** Direct voter contact, such as canvassing

**Salary** Payments made to staff of the organization

**Swag** Clothing, signs, and other promotional material

**Transport** Transportation costs, such as taxis or van rentals

**Other** All expenses that do not fit into the above categories

An additional complication we faced was with the support/oppose column. This column in conjunction with the candidate column are used to indicate which candidate benefits from the expenditure. An example would be if the support/oppose value is “oppose” and the candidate name is “Romney”, then Mr. Obama benefits because the money is being spent to “oppose Romney”. Likewise, if the support/oppose column equals “support” and the candidate name is “Romney”, then Mr. Romney benefits. We solved this by adding a new field which simply stores the candidate that benefits from the particular expense. That is, an opposition advertisement expense entered with a candidate name of Obama will benefit Romney.

The polling data did not require much data cleanup. We split the date range and formatted the end date as a date for our use. We also removed some differences in state naming by the different pollsters. One limitation we faced is that a fairly significant number of polls were excluded from the nationalpolls.com database. Some new national polling outfits, such as the RAND Corporation, were excluded, as were a large number of online polls, including Google Consumer Surveys. We were unable to find a methodology for the inclusion/exclusion of certain polls, but we suspect that polls without a track record in previous presidential elections were likely to be excluded. As it turns out however, online surveys performed more strongly in the 2012 election than did automated phone surveys or live interviewer surveys according to columnist Nate Silver.<sup>1</sup> Unfortunately, we could not find a better source of polling data and therefore chose to perform our analysis without these online surveys.

### 3 Findings

Our key findings are summarized in the following sections. We begin by analyzing temporal and categorical trends in spending for Super PACs benefiting each of the candidates. We then take a closer look at the organizations which are spending the most on each side. Next, we analyze changes in the polling data with an emphasis on swing states. We conclude by attempting to correlate these changes to spending changes over time.

#### 3.1 Type of Spending by Week

Figure 1 displays the spending in each of the six categories by independent organizations benefiting Mr. Obama and Mr. Romney for each week since April 25, 2012. The spending amounts are split by whether the goal was to support or oppose the candidate. There are three markers which signify major events in the campaign: the selection of Representative Paul Ryan as Mr. Romney’s vice presidential nominee, the Republican National Convention, the Democratic National Convention, the 47% video, and the debates. Interesting trends can be observed in advertisement spending, and direct contact. Ads in support of both candidates show increases since the end of July. However, ads in opposition of the candidates were airing well before. You can, however, see an increase in negative ads benefiting Mr. Romney. A noticeable increase in spending on direct contact for Obama is also evident beginning in August, while the organizations benefiting Mr. Romney have maintained consistent spending in negative direct contact beginning in mid-May.

---

<sup>1</sup>Which Polls Fared Best (and Worst) in the 2012 Presidential Race

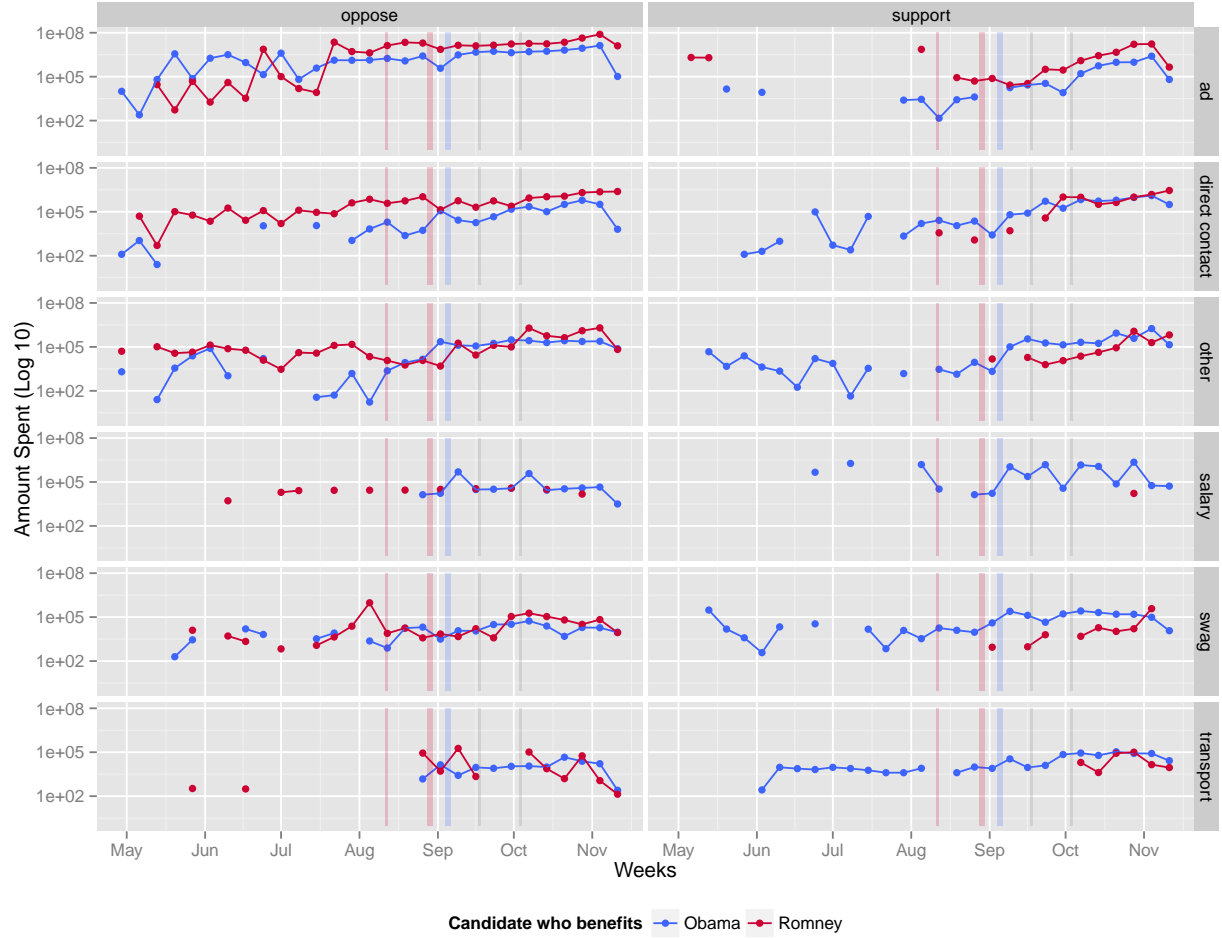


Figure 1: Spending by Super PACs in support or opposition of candidates

### 3.2 Spending Categories

These plots shows the comparison of amount spent for each candidate by categories of independent expenditure, since April 25, 2012. The categories are mined from the self-reported text description of the expense. The left plot shows the spending to benefit Mr. Obama against the spending to benefit Mr. Romney by category, with values to the right of the reflection line showing PACs supporting Mr. Romney are outspending PACs supporting Mr. Obama on these categories, and vice versa for those to the left. The right plot shows bar chart comparisons of the spending in main categories, on a log scale. Organizations supporting Romney are greatly outspent those supporting Obama, and particularly in direct contact and advertisements. On the other hand, organizations supporting Obama only outspent those supporting Romney in salary.

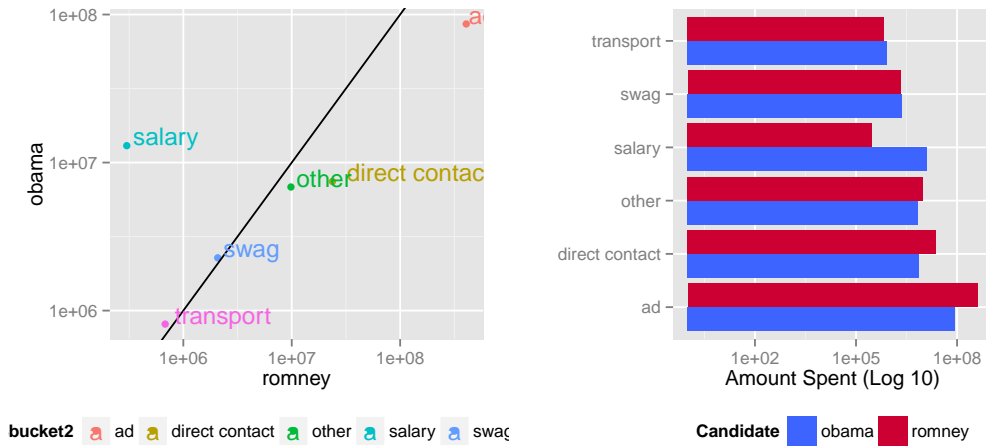


Figure 2: Types of spending by category for Romney and Obama Super PACs

### 3.3 Spending by Independent Organizations

This chart displays the total spending by the top independent organizations split by candidate. The cumulative amounts spent is displayed vertically, by the benefiting candidate. The organizations supporting Romney have spent significantly more than those for Obama. In fact, two Romney Super PACs (Restore our Future, Inc and Americans for Prosperity) have spent more than all the organizations supporting Obama combined.

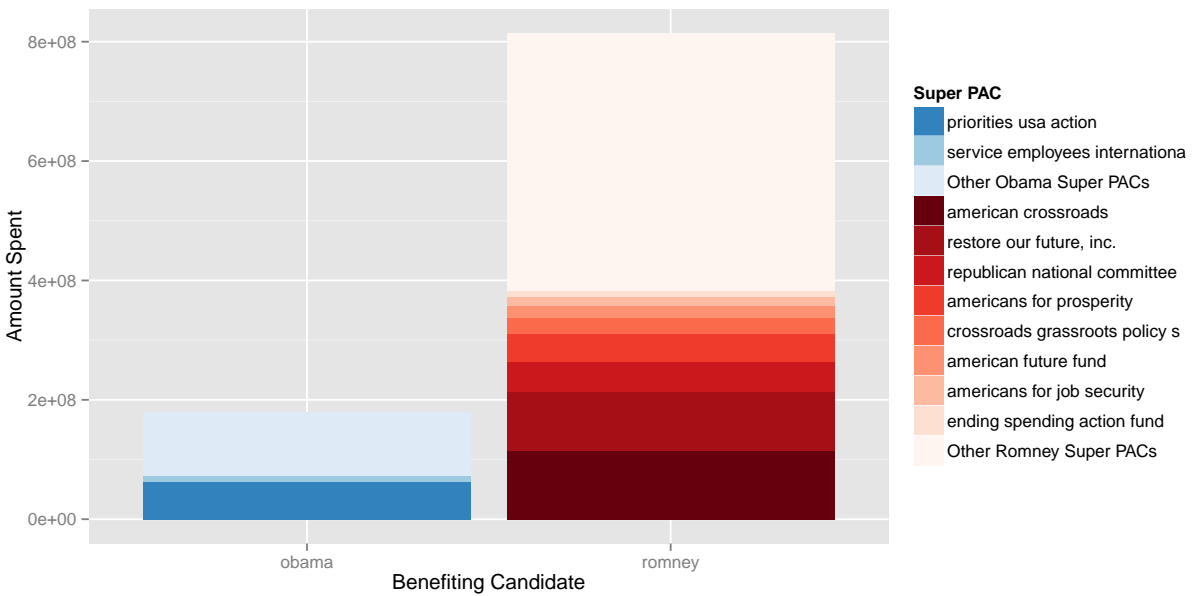


Figure 3: Spending by Super PAC, stacked by candidate



### 3.4 Swing State Trends

This plot displays the change in polling support for both Romney and Obama over time, in 12 swing states. The data includes all polls in the National Polls database from April 25th to now. Vertical lines indicate difference between the two candidates, and color shows who is ahead. There are five markers which signify major events in the campaign: the selection of Paul Ryan as Mitt Romney’s vice presidential nominee, the Republican National Convention, the Democratic National Convention, the 47% video and the first presidential debate. The plots suggest that the reaction to these major campaign events was not consistent state to state. Some states, such as Wisconsin, have produced drastic changes in polling support over time (“bounces”), while others, such as North Carolina, have maintained a consistent margin of support for the two candidates.

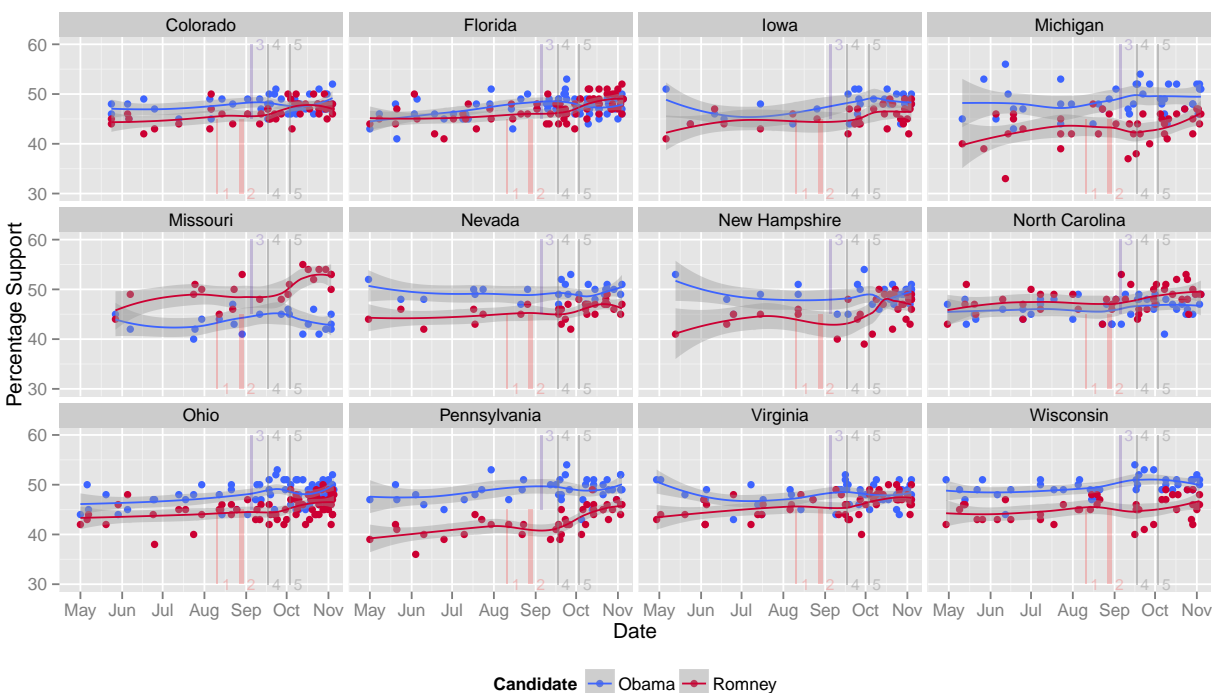


Figure 4: Polling Averages for Obama and Romney by swing state

### 3.5 Spending by Week

Looking at weekly spending by PACs benefiting Romney, we can see a sharp increase in spending that occurred during the week of July 18th. We look at the national and swing-state polls to determine if there is any noticeable effect due to these changes in spending.

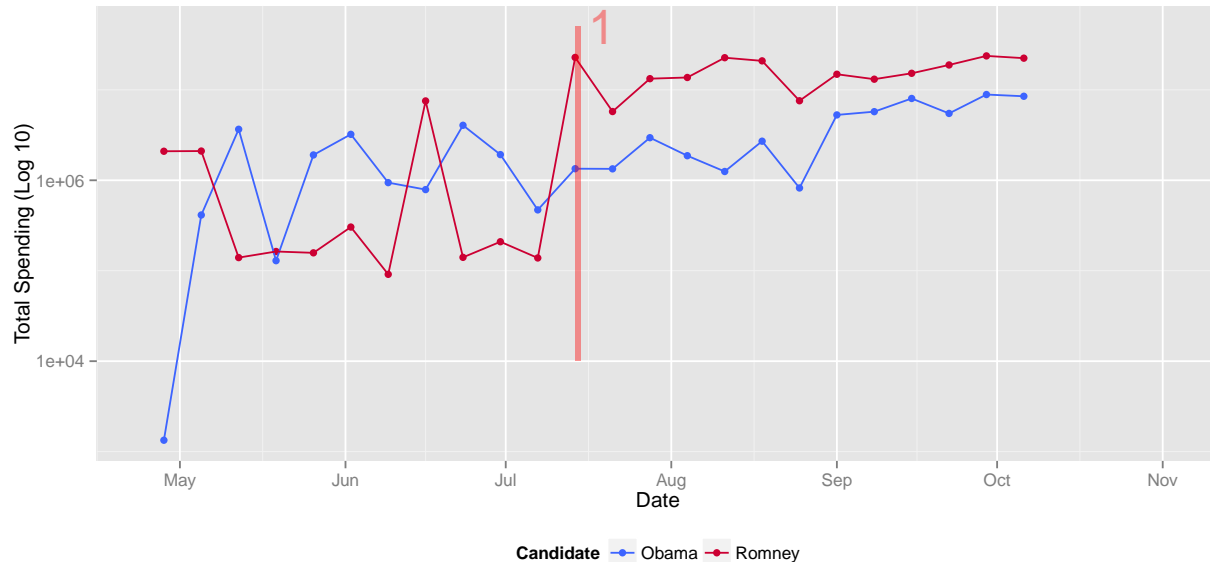


Figure 5: Weekly Spending by Super PACs supporting Obama and Romney

### 3.6 Effect of Spending on Polls

This plot shows the polling margin (Obama - Romney) over time, colored by swing states versus the national polls. It can be seen that Mr. Obama consistently maintained an advantage in swing states relative to his national numbers. A marker is placed on July 18th, corresponding to the sharp increase in spending by Romney-supporting PACs. It does not seem that this spending increase had a measurable effect on the overall trend in the polls at this time.

Note that this plot utilizes exponential smoothing to show the trends in polls over time rather than a LOESS smoother, as we have used in all other plots. When using the LOESS smoother, the future polling numbers were heavily affecting the smoothed trend line, which was undesirable. This is because the LOESS smoother takes points past and future with equal weight into account when showing the trend. However, the exponential smoothing method assigns exponentially decreasing weights over time utilizing past points. We implemented the exponential smoothing method using the `forecast` and `zoo` packages.

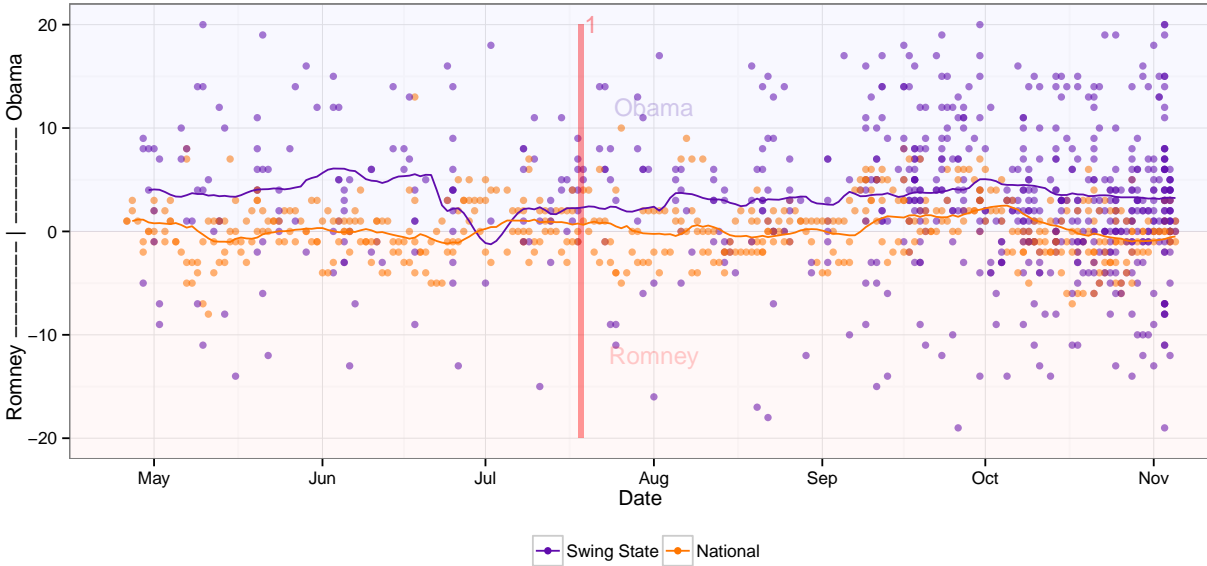


Figure 6: Polling Margin (Obama - Romney) over time, swing states versus national polls

This plot compares the difference in spending by week by the PACs to the difference in polling between Mitt Romney and Obama with a one week lag. The points are colored by week and certain weeks are labeled to indicated important events. The goal is to see if there is a relationship between PAC spending and poll results. In both Obama-supporting and Romney-supporting PACs we can see a weak positive relationship between spending and polling.

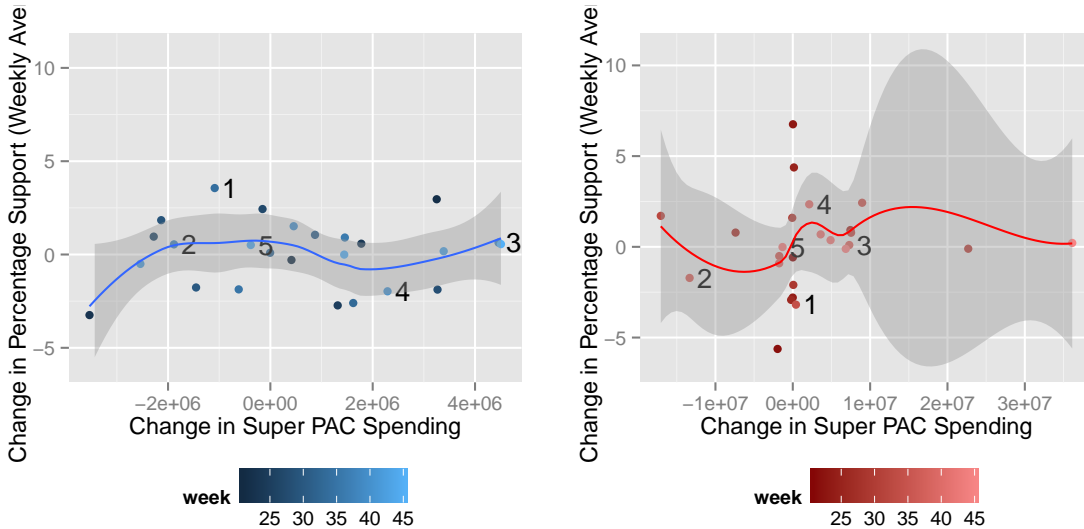


Figure 7: Change in polling over change in spending by candidate

## 4 Conclusions/Future Work

Ultimately, our findings suggest that while election spending itself was a very significant part of the 2012 elections, it is not clear whether the existence of Super PACs had a measurable impact on the outcome. Instead, the events that garnered large amounts of media coverage, such as the 47% video and the first presidential debate, seemed to have the most noticeable effect on the polling. Furthermore, our analysis did not take into account the spending done by the candidates themselves. It may be that such spending muted the effect of Super PAC spending.

Our analysis also makes clear that Mr. Obama's polling held up more strongly in the swing states than in the national polling throughout the campaign. This advantage was evident in the election results, as well. According to the current certified vote tally as of December 4th<sup>2</sup>, Mr. Obama won the overall popular vote by 3.6%. But he won Colorado, the state that put him over the 270 electoral votes needed for victory, by 5.4%. This suggests that Mr. Obama could have lost the popular vote by nearly 2% and still have been elected president of the United States.

An extension of this project should take into account spending done by the candidates themselves, in addition to independent expenditures. This would allow for a more informed look at the relative spending difference, and how it may or may not correlate with changes in the polling averages. Another area of exploration would be the purpose field of the independent expenditures data set. Our analysis attempted to categorize spending entries into broad categories, but many of the entries in the database have more specific information. For instance, an ad buy may list the name of the ad, as well as the state the ad is running in. It may also be possible to incorporate another data source in order to link the expenditure entries with ad buys in particular states.

---

<sup>2</sup>Wasserman's 2012 National Popular Vote Tracker