# Working towards a predictive platform for elections

**PsephoAnalytics**

psephoanalytics.blogspot.ca/
@PsephoAnalytics
PsephoAnalytics@gmail.com

*October 13, 2015*

# Understanding the "Political Brain"

> "If [the Enlightenment idea of reason] were right, politics would be universally rational. If the people are made aware of the facts and figures, they should naturally reason to the right conclusion. Voters should vote their interests; they should calculate which policies and programs are in their best interests, and vote for the candidates who advocate those policies and programs. But voters don't behave this way. They vote against their obvious self-interest; they allow bias, prejudice, and emotion to guide their decisions; they argue madly about values, priorities, and goals. Or they quietly reach conclusions independent of their interest without consciously knowing why. Enlightenment reason does not account for real political behavior because [it] is false."
>
> *George Lakoff, "The Political Mind", p. 8*

> "*The political brain is an emotional brain.* It is not a dispassionate calculating machine, objectively searching for the right facts, figures, and policies to make a reasoned decision."
>
> *Drew Westen, "The Political Brain", p. xv*

Political psychologists like George Lakoff and Drew Westen have long held that over-simplified "rational" models of voters do not help accurately predict their actual behavior. This often includes a "left-right" scale that political pundits love. Researchers note that this is an inaccurate metaphor: there is not a clearly unified 'mainstream' worldview. Terms like "conservative", "liberal", or even "environmentalist" cannot fairly explain the complexity of voters' (and candidates') beliefs[1].

[1] That said, neurologically, if a voter is more "conservative" it is because there are more of those receptors at the voter's synapses. As the voter continues thinking about issues where they are already "conservative", it becomes more likely that "conservative" will begin to unconsciously bind to other issues as well (called neural binding). Changing someone's mind really is about changing their *brain*. This is why candidates should focus on repeating their (moral) worldview (not specific policy positions) as often as possible. (It is also why candidates should not engage in their opposition's frames, even to argue against them.)

What most behavioural researchers have found is the decision-making (e.g., voting) often boils down to emotional, unconscious factors. For example, research[2] has shown that snap judgments (i.e., simply seeing the face of a candidate) allow voters to assess competence. Even additional information that would normally be gathered over a campaign only dilutes the impact of an initial impression, but does not eliminate it.

Further, much research shows that an abundance of options with trade-offs starts to make all options look unappealing (think of comparing detailed policy platforms). In such cases, people tend to seize at rationalization to help them decide. Given that evolution has "trained" our minds to distinguish simply between good and bad, a democratic process where relatively little objective room separates policy platforms makes distinguishing better from slightly worse difficult.[3] In such cases, falling back on simpler ways to decide seems understandable.

So, in attempting to build up our voting agents, we will need to at least:

- include multiple issue perspectives, not just a simple evaluation of "left-right";
- include data for non-policy factors that could determine voting; and
- not prescribe values to our agents beyond what we can empirically derive.

It may become clear that we do not have access to data that would give us a complete picture of how our agents should vote, so we are simply aiming to explain as much as we can (including the use of proxies).

## So what data can we use?

Given that we are unable to peek into voters' minds (and remember: we are trying to avoid using polls[4]), we need data for (or proxies for) factors that might influence someone's vote. In all cases below, we gathered (or created) and joined data for the 2006, 2008, and 2011 Canadian federal elections (as well as the 2015 election, which will be used for predictions).

---

[2] See Hallinan, p. 69

[3] See the *Wisdom of Crowds* (pp. 128, 142-3)

[4] This is true for a number of reasons: first, we want to be able to simulate elections, and therefore would not always have access to polls; second, we are trying to do something fundamentally different by observing behaviour instead of asking people questions, which often leads to lying (e.g., social desirability biases: see the "Bradley effect"); third, while polls in aggregate are generally good at predicting outcomes, individual polls are highly volatile.

## Follow the Leader

The political scientist John Zaller has shown that voters' perception of a party strongly follow that of the party leaders and other "elites". (This has since been *partly* refuted; see "Should the mass public follow elite opinion? It depends…".) Nonetheless, the theory goes that most voters do not have the time to think through all issues, and will therefore take on the beliefs of "elites" that they trust. (We will not discuss here the relative benefits and costs, either socially and individually, of such behavior.)

Further, there is evidence that "allegiance to party – a largely emotional allegiance – remains the central determinant of voting behavior"[5] where it exists. When it does exist, even *seeing* their own party's candidates and opposition party's candidates activate different parts of a partisan's brain.

We therefore assign leader "likeability" scores to the various party leaders of the three major (national) parties, using polls that ask questions about leadership characteristics and formulaically compare them to party-level polls around the same time. This provides an estimate of how much influence a party leader had on their party's showing in the polls, and *should* account for much of the within-party variation that we see from year to year. (We also use party identifiers, to identify a true "base".)

## What's in a name?

Many voters focus more locally too. Canadian research (using the 2000 federal election) found that "44 per cent of Canadian voters formed a preference for a local candidate and that this preference had an effect on vote choice independent of how people felt about the parties and the leaders. The findings suggest that the local candidate was a decisive consideration for 5 per cent of Canadian voters".[6] Factors that can easily be attributed to local candidates will be useful; we currently use two:

Literature (e.g., Gelman and King's paper, "Estimating Incumbency Advantage without Bias") has suggested that being an incumbent is worth a few percentage points in an election. Further, parties often try to enlist "star candidates" to draw out additional voters. Threehundredeight.com founder Éric Grenier has estimated that a star candidate multiplies the base proportion of votes by somewhere around 1.15 (some with more or less), or about 3-6 percentage points.

[5] Westen, p.7

[6] Blais et. al., 2003

Therefore, for all 366 Toronto-based candidates across the three elections, we identify two things: whether they are an incumbent, and whether they are a "star" candidate, by which we mean would they be generally known outside of their riding? This yields 64 candidate-year incumbents (i.e., an individual could be an incumbent in all three elections) and 29 candidate-year stars in Toronto races.

## Got issues?

The final component that we consider from a candidates' perspective is the parties' platforms.

Despite the above, there are certainly voters who are, at least somewhat, influenced by the policy positions of party platforms. For policy to influence voting behavior, voters must a) care about an issue, b) have an opinion on that issue, and c) be able to tell the difference between candidates' positions on that issue. This combination of conditions is often hard to meet.

Further, changing minds can be hard as "several experiments have found that people evaluate evidence that disconfirms their cherished beliefs much more critically than evidence that supports [them]."[7] And it seems that policies matter only as far as they influence voters' emotions. So, we would not expect this factor to be as influential as the others. It is more likely that we'll find that a couple key issues outweigh the others.

Nonetheless, we want to include policy in the voting decision. In order to define such positions, we start with a detailed analysis of party platforms for the three elections, generating position scores for 175 year-party-topic combinations over the three elections. Each topic is scored on a "progressive-conservative" scale of 0 to 100 and grouped into six topic types (economy, environment, foreign policy, general government, social, and health & education). Averaging individual topic scores across these types allows for comparisons across the elections.

## Demographic divides

Now that we've looked at aspects of candidates, parties, and their leaders, we can move to attributes of the voters themselves.

The "gender gap" in voting has been well studied, though the degree to which it exists and reasons why vary (even one of us studied this for our MA thesis). Nonetheless, it is typically found that, in modern economies, women tend to vote for more progressive parties than men. Similarly, young

[7] Westen, p. 100

voters are typically more progressive as well (though also less likely to vote; see our study on turnout).

Other demographic and socio-economic factors (e.g., race, religion, marital status, wealth/ income) have also been shown to influence voting outcomes.

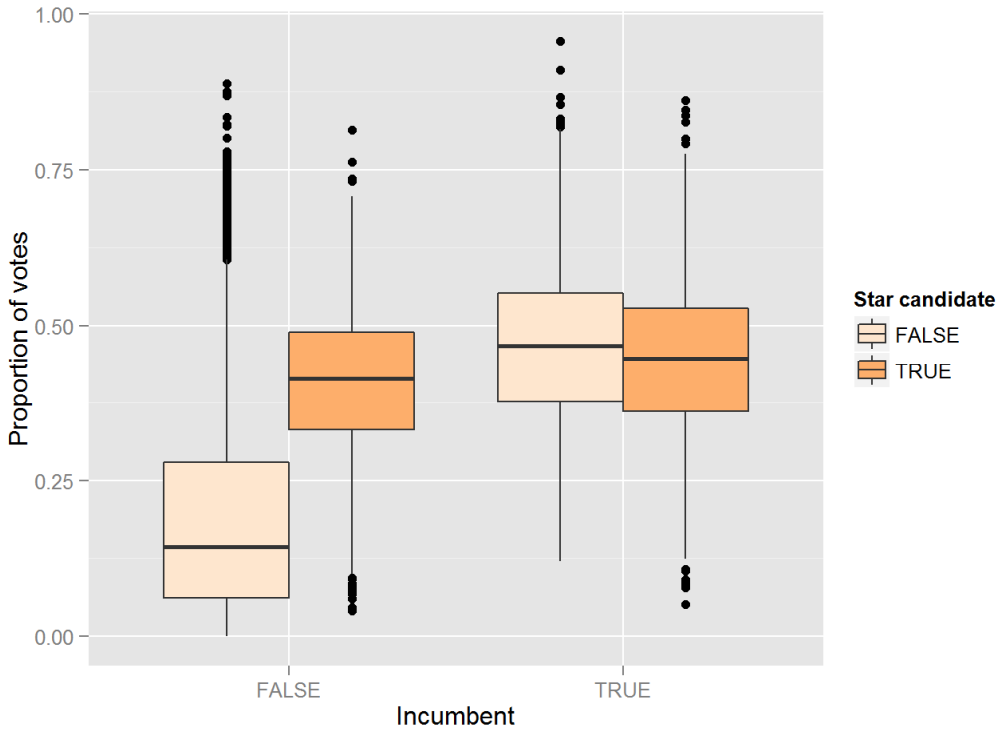
In this case, we use data from the 2011 Census, specifically age, gender, and family income.

# Empirical results: factors that influence voting

## Starting at a high-level

Regressing leader likeability, incumbent, star candidate, and party data against the proportion of votes received across ridings yields some interesting results. First all these variables are statistically significant (as is the interaction of star candidate and incumbency). This isn't a surprise, given the literature around what it is that drives voters' decisions. (Note that we haven't yet included demographics or party platforms.)

Let's break down the results. We start with a simple plot showing the proportion of votes received for the four types of candidate.



A mixed-effects model of these data show that being either a star candidate or an incumbent can boost a candidates share of the vote by 21%. But, as is clear in the figure above, being both an incumbent *and* a star candidate does not give a candidate an incremental increase. In other words, the effects are not additive, as evidenced by the statistically significant interaction between the two effects.

Furthermore, the two effects are equivalent to belonging to a party (21%). So, being a member of a major party and either incumbent or a star candidate (i.e., with name recognition) offers the best chance of winning an election. Perhaps not a surprising finding overall, but the relative equality of the three effects, plus negative interaction of incumbency and star candidate, are useful nuances. We do need to be careful with this interpretation though. For the period we are considering in Toronto ridings, the Conservatives haven't had any incumbents and the NDP haven't had any stars who weren't also incumbents.

Likeability matters too. A linear, mixed effects model finds that leader likeability is associated with a 0.3% change in the proportion of votes received by a candidate. So, a leader that essentially polls the same as their party yields their Toronto-based candidates about 14 points. (As an example of what this means, Stephane Dion lost the average Liberal candidate in Toronto about 9 points relative to Paul Martin. Alternatively, in 2011, Jack Layton added about 16 points more to NDP candidates in Toronto than Michael Ignatieff did for equivalent Liberal candidates.) What we're trying to get here is some sense of what effect a leader has, as an individual.

Finally, party base matters: for example, being an average Liberal candidate in Toronto adds about 17 points over the equivalent NDP candidate. (We expect some of this will be explained with demographics and party platforms.)

## Drilling down on demographics

Given that we are developing agents, stopping at the high-level (despite some useful results overall) is obviously insufficient. We have to determine what effect, if any, demographics may have on voter preferences.

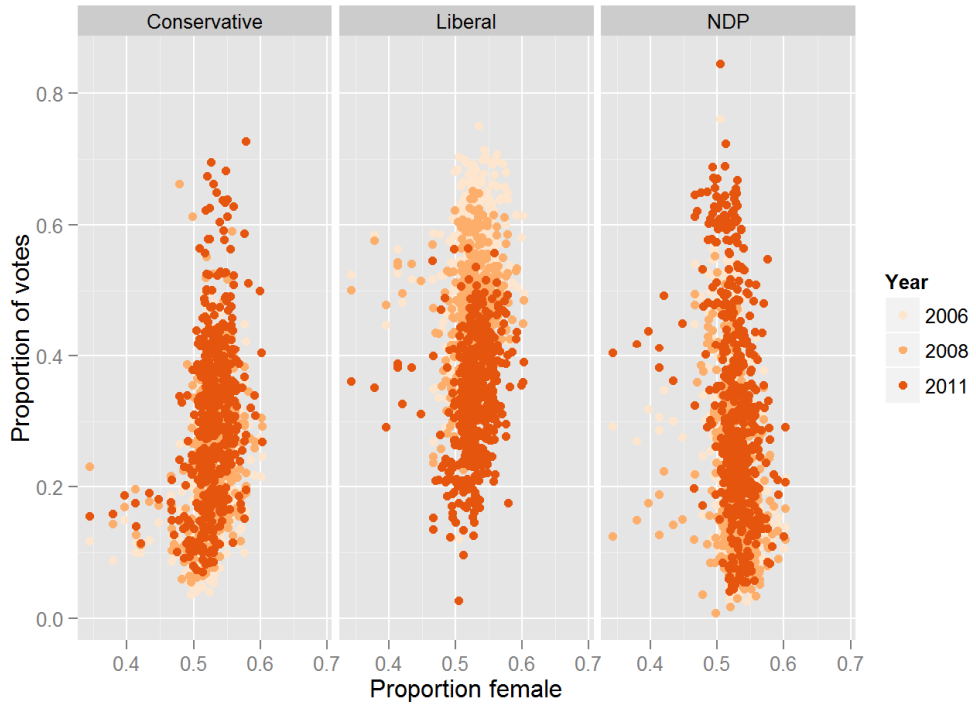
The linear, mixed effects model shows that the relationships between age, gender, family income, and the proportion of votes vary widely across the parties (as expected). In general:

- Age tends to increase support for Conservatives and Liberals (though the slope is higher for Conservatives; 0.01 vs 0.007), while decreasing support for NDP (-0.012).
- Family income tends to increase support for Conservatives (0.005/$10,000) while decreasing for the other two major parties by roughly the same magnitude.

- Gender is a surprise. We are detecting a strong signal for increased support by women for Conservatives, moderate support for Liberals, and strong negative support for NDP. This is not consistent with the notion that women favour more progressive parties and requires further analysis.[8]

One important point to make is that we are not actually tracking the demographics of voters themselves. Rather, we are using census data for the "neighbourhoods" in which the voters cast their ballots. Consequently, there is a chance that neighbourhoods with a higher proportion of women are voting for Conservative candidates, even if the individual women are more likely to vote for a different party. There is no evidence for this, however, and there is no particular reason to expect that such an effect would only influence Conservative support.

The figure below takes a closer look at the relationship between gender and the proportion of votes received by each party. The year of the election is indicated by the colour.




One evident pattern is that the points are oriented vertically. A strong gender effect would appear as a horizontal, or at least angled, line. For the most part, each party's support is clustered around

[8] It could simply be that the other variables have accounted for this. For example, some research (including the aforementioned MA thesis of ours) contemplates whether economic uncertainty for women (particularly those with children) was what drove their progressiveness - and we have accounted for income.

a relatively narrow band just above 50% female. This may be a limitation of these data. On average, Toronto is about 52% female (according to the census data) and this proportion is roughly the same across census tracts. There may simply not be enough variation in gender at the census tract level to reveal an important relationship with voting.

However, if we look a little closer at the bottom left of the panel for Conservatives, we see a cluster of census tracts with a relatively high proportion of males that has low support for Conservatives (below about 0.2 for proportion of votes). Liberals appear to have support from some of these high-male proportion census tracts between about 0.3 and 0.6 for proportion of votes and the NDP from about 0.1 to 0.5. The gender effect may, in fact, be due to specific, predominantly male census tracts that tend to *not* favour the Conservatives. This is subtly different from a female preference for the Conservatives.

### Do issues actually matter?

Finally, we look at the influence that policy platforms have. Across all issues, it appears that only economic and environmental issues have meant enough to influence the Toronto electorate overall. This may be because other issues actually matter less, or (perhaps more likely) because media "elites" report that these are the most important issues.[9]

Based on a linear, mixed effects model, high platform scores for either the economy or the environment decrease the proportion of votes for the Conservatives. The economic effect is very small (-0.01) and the environmental one is as strongly negative (-0.037) as it is positive for the Liberals (0.34).

Increased scores in either the environmental or economic topics increase votes for the Liberals and the NDP. However, the Liberals benefit 3 times as much on economic topics, while the NDP benefit 1.3 times as much on environmental. This suggests an important split between Liberal and NDP supporters along the economic-environmental axis.

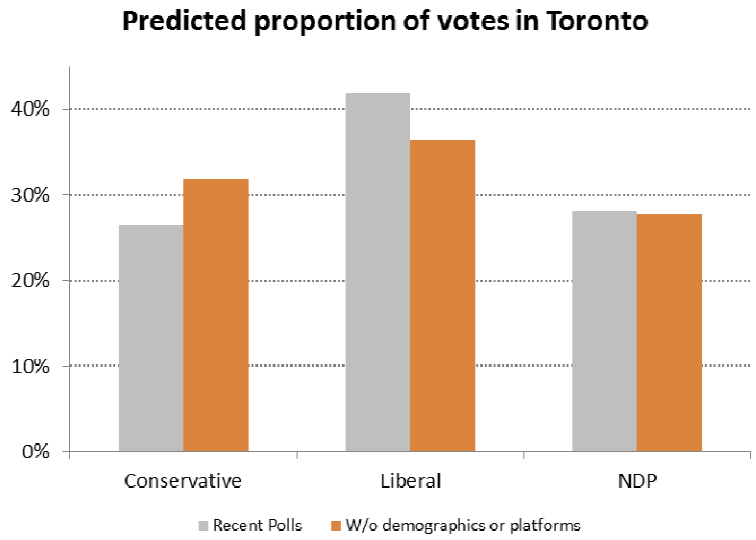
## Applying these results to the 2015 Canadian election

### Starting at the candidate level

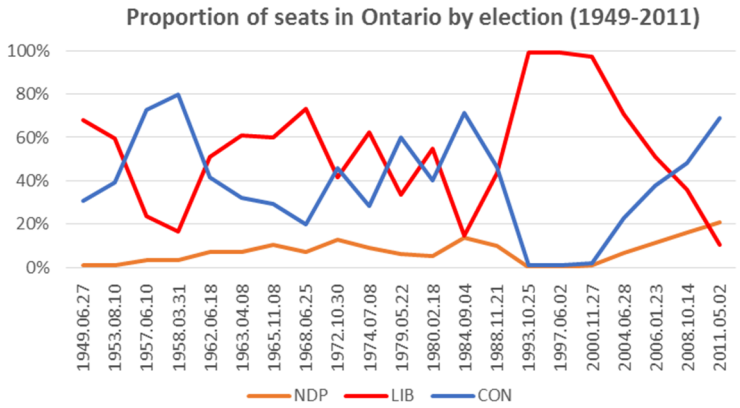
Starting with the high-level candidate (i.e., non-demographic) analysis yields some interesting results. To be clear, these are average results, so we can't yet use them effectively for predicting

[9] Our plan to build up an extensive agent based model includes media agents.

individual riding-level races (that will come later). But, if we apply them to all 2015 races in Toronto and aggregate across the city, we would predict voting proportions fairly similar to the results of a few recent polls (weighted by age and sample size) that showed Toronto-specific estimates:



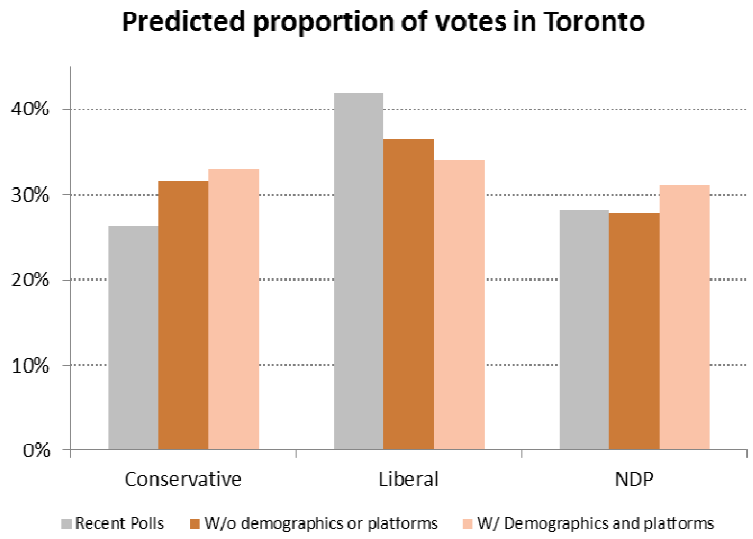

Given that we haven't used polls or included localized details or party platforms, these results are surprisingly good. The seeming shift from Liberal to Conservative is something that we'll need to look into further. It is likely highlighting an issue with our data: namely, that we only have three years of detailed federal elections data, and these elections have seen some of the best showings for the Conservatives (and their predecessors) in Ontario since the end of the second world war (the exceptions being in the late 1950s with Diefenbaker, 1979 with Joe Clark, and 1984 with Brian Mulroney), with some of the worst for the Liberals over the same time frame. That is, we are not picking up a (cyclical) reversion to the mean in our variables, but might investigate the cycle itself.




Next, we investigate what happens if we include policy and demographics.

## Adding in demographics & policy

Once we account for demographics and platforms (i.e., all of our variables), we get slightly less appealing results relative to the polls, but essential the same as before. While we still believe we are on the right track, there is still much work to go.




# Conclusion

We set out to understand (both theoretically and empirically) how to predict an election while significantly limiting the use of polls, and it appears that we are at least on the right track. Our Toronto-wide results are fairly in line with recent Toronto-specific polling results – though we'll see how right we all are come election day – which means that there may some inherent "truth" in the coefficients we have found.

Our next step is to try to use our beta distribution techniques to apply these results to individual voters (as well as political engagement scores) to run election simulations, though this will likely come after this election. There are also other factors that have been shown to influence voting decisions, including general evaluations of an incumbent's performance (perhaps this is part of

the cycle discussed above), candidate traits[10], personal "connectedness", or even the ballot ordering[11], which we have not (yet) included in our modeling.

## References


Blais, André, Elisabeth Gidengil, Agnieszka Dobrzynska, Neil Nevitte, and Richard Adeau. "Does the Local Candidate Matter? Candidate Effects in the Canadian Election of 2000." *CJP Canadian Journal of Political Science/Revue Canadienne De Science Politique* 36.03 (2003): n. pag. Web.

Gelman, Andrew, and Gary King. "Estimating Incumbency Advantage without Bias." *American Journal of Political Science* 34.4 (1990): 1142. Web.

Hallinan, Joseph T. *Errornomics: Why We Make Mistakes and What We Can Do to Avoid Them.* London: Ebury, 2009. Print.

Hochschild, Jennifer L. "Should The Mass Public Follow Elite Opinion? It Depends …." *Critical Review* 24.4 (2012): 527-43. Web.

Lakoff, George. *The Political Mind: A Cognitive Scientist's Guide to Your Brain and Its Politics*. New York: Penguin, 2009. Print.

Surowiecki, James. *The Wisdom of Crowds: With a New Afterword by the Author*. New York: Anchor, 2005. Print.

Westen, Drew. *The Political Brain: The Role of Emotion in Deciding the Fate of the Nation.* New York: PublicAffairs, 2007. Print.

Zaller, John. *The Nature and Origins of Mass Opinion*. Cambridge: Cambridge UP, 1992. Print.


[10] See http://www.icpsr.umich.edu/icpsrweb/instructors/setups2012/background/candidate-chars.jsp

[11] Jon Krosnick conducted a series of studies around the 2000 US presidential elections in states that rotate the order of names as they appear on ballots. He found significant benefits for whoever appeared first on the ballot – worth up to 2%.