Should Recall of Previous Votes Be Used to Adjust Estimates of Voting Intention?

Special issue

Claire Durand, Mélanie Deslauriers, Isabelle Valois, Department of sociology, Université de Montréal

20.02.2015

How to cite this article: Durand, C., Deslauriers, M . & Valois I. (2015). Should Recall of Previous Votes Be Used to Adjust Estimates of Voting Intention? *Survey Insights: Methods from the Field, Weighting: Practical Issues and 'How to' Approach.* Retrieved from http://surveyinsights.org/?p=3543

Abstract

Correction of estimates of voting intention using voter recall of previous votes is frequent in electoral polls, particularly in Europe. However, research on the impact of its use is scarce. The results presented in this paper confirm that voting is not a salient, memorable behaviour for all voters. People who always vote the same way and identify with a political party are likely to accurately remember their vote, but in most developed countries, change of allegiance is now common. A substantial portion of the electorate changes its mind between or during campaigns, and switchers seem to have more difficulty remembering how they voted. Recall error is not random. Voters' misremembering a previous vote to reconcile it with how they currently wish to vote (reconciliation) and, above all, the difficulty in reaching voters for far-right or populist parties/candidates, and in convincing them to reveal their true vote or voting intention are the main explanations for error. Memory failure also plays a role, but the overall impact appears to be weak. Finally, reconciliation and social desirability also play roles when it comes to correcting estimates using recall of past voting, but the overall impact is weak. The results presented here show that, at best, the practice does not have a significant, substantial, impact on estimates. We suggest that researchers and pollsters would do better to focus their energies on tackling the problem itself instead of working on a posteriori correction.

Keywords

adjustment, electoral polls, vote estimates, voter recall, weighting

Acknowledgement

The authors wish to thank the Canadian Election Study and the Panel Electoral Français as well as the Centre de recherche sur l'opinion publique (CROP) who provided the data on which this paper is based. They are also most thankful to the reviewers whose suggestions greatly help improve this paper. Finally, this paper benefited from the financial support of SSHRC grant no. 410-2009-1582.

After collecting survey data using random samples, it is usually necessary to weight the data and/or make adjustments using external data. Weighting compensates for the sampling process. When the post-weighting figures obtained in the sample differ significantly from census figures on major variables such as age, sex, or region of residence, the sample may be adjusted to better reflect the demographic composition of the population. Weighting and adjustments are based on fairly reliable information such as sampling ratio, response rate, number of people in the household, and census figures. When quota samples are used, the quality criterion is whether

quotas are respected. Therefore, no information is usually gathered about sampling ratio, response rates or number of people in the household. The selection process is based on the distribution of characteristics in census data. In addition, an indicator of social class is often used. In France, this is usually the profession of the head of household; elsewhere, it may be a rough evaluation ("low," "medium," "high") made by the interviewer. With this method, samples automatically reflect the census for the socio-demographic variables used to determine the quotas. However, samples often suffer from political bias, i.e., they do not reflect the political composition of the electorate, possibly due to self-selection. The distribution of declared previous votes, for example, frequently does not match election results. To compensate for this bias, some pollsters make adjustments based on respondents' recall of their vote in previous elections when estimating voting intention.

According to Taylor (1995) and Hellevik (2009), this practice is widespread among private pollsters conducting pre-election polls, at least in Europe. It is commonly used in Germany, the United Kingdom, Denmark, Portugal (Taylor, 1995), Norway (Hellevik, 2009), the Netherlands (Foekema, 2009), and France (Taylor, 1995; author reference, 2004, 2008; Bachelet, 2007). It was originally used to compensate for a systematic underestimation of the communist vote in the first polls conducted after World War II in Europe (Hellevik, 2009); it was suggested as a way to reduce sample variance and/or correct for political bias following the failure of the polls in Great Britain in 1992 (Jowell, 1993); it is now justified by the fact that samples tend to be biased against the far-right-wing vote (Bachelet, 2007; Hellevik, 2009).

Given its widespread use, there are surprisingly few studies that examine the practice as a whole, its relevance, and its consequences. Is voter recall reliable, and if so for how long? What is the impact of using recall to adjust the estimates of voting intention? This paper aims to bridge some of the knowledge gap regarding this practice by examining the impact of using declared past vote to adjust estimates of voting intention. To this end, we use data gathered in different electoral contexts. However, one must first ask whether voter recall is valid and reliable in the first place.

Is voter recall valid and reliable?

The reliability of recall data in general, and of voter recall specifically, was questioned early on in research on survey methodology (Dex, 1995). The validity of respondents' declarations regarding their participation in the vote can be assessed in some countries—Sweden (Anderson & Granberg, 1997) and the U.S. (Traugott & Katosh, 1979; Katosh & Traugott, 1981)—using voter records. However, the validity of the declared vote cannot be verified using external information. Therefore, researchers assess the reliability of voter recall by comparing two different individual declarations, and they assess its validity by comparing the distribution of the declared vote with election results.

A limited number of researchers, mostly interested in voter stability, have examined the reliability of declared past vote. Studies were conducted in England (Benewick, Birch, Blumber & Ewbank, 1969; Himmelweit, Biberian & Stockdale, 1978), the U.S. (Weir, 1975; Granberg & Holmberg, 1986), the Netherlands (Van der Eyk & Niemoller, 1979), Sweden (Granberg & Holmberg, 1986), and Norway (Waldahl & Aardal, 1982, 2000; Hellevik, 2009). With the notable exception of Norway, most studies were conducted in the 1960s and 70s.

This research indicates that inconsistency in recall of previous votes, i.e., when the recalls of two previous votes do not match, varies between 7 and 39 percent but is usually in the 10–25-percent range (Benewick et al., 1969; Weir, 1975; Himmelweit et al., 1978; Van Der Eijk & Niemoller, 1979; Waldahl & Aardal, 1982; Anderson & Granberg, 1997). More recently, Wahldal and Aardal (2000) found levels close to 30 percent for the Norwegian elections of 1993. Hence, voting does not appear to be a salient behaviour that remains vivid in people's minds for an extended period. According to Waldahl and Aardal (2000), the level of misreporting is likely to rise with recent trends in many countries toward higher numbers of parties and candidates and increased voter volatility.

A number of factors have been hypothesized as possible determinants of recall accuracy, whether socio-demographic (primarily age and education level), political (stability of voting behaviour, party identity), or contextual (passage of time, change in voting intention). According to Himmelweit et al. (1978) and Markus (1986), age could affect the type of error that respondents make, and according to both Van Der Eijk and Niemoller (1979) and Waldahl and Aardal (1982), the various effects of age may cancel each other out. The general conclusion is that age is indirectly related to accuracy of recall mostly in that age groups differ with regard to political behaviours and attitudes, i.e., stability of preferences, identification with a political party, and involvement in the electoral process. Education level, thought to be an indicator of political sophistication, is hypothesized as being negatively correlated with misreporting. However, empirical findings generally do not

support this hypothesis (Weir, 1975; Himmelweit et al., 1978; Waldahl & Aardal, 1982; MacDermid, 1989).

Recall reliability is generally hypothesized to be positively associated with the salience and stability of the behaviour or attitude under scrutiny and negatively associated with time and with changing attitudes (Dex, 1995; Waldahl & Aardal, 2000). Research on voter recall confirms that people who consistently vote for the same party tend to better recall their past vote (Himmelweit et al., 1978; Van der Eijk & Niemoller, 1979; Waldahl & Aardal, 1982, 2000). However, the relationship between interest in the electoral campaign and accuracy of recall appears to be weak (Weir, 1975), while political interest in general is either weakly related to recall accuracy (Van Der Eijk & Niemoller, 1979) or not related at all (Waldahl & Aardal, 1982). By contrast, identification with a political party is consistently and strongly related to recall accuracy (Weir, 1975; Van Der Eijk & Niemoller, 1979; Waldahl & Aardal, 1982; Granberg & Holmberg, 1986).

The direction of incorrect recall varies: Respondents may err in the direction of their current opinions (Himmelweit et al., 1978; Markus, 1986, Joslyn, 2003) or in the direction of the winner or most popular opinion. Recall that drifts toward the winner is associated with weak party identification (Weir, 1975). Waldahl and Aardal (1982) find an interaction effect such that among unstable voters, when party identification is weak, there is a drift toward the winner, and when it is strong, the drift is toward the current allegiance. However, authors seem to agree that there is a greater tendency to declare a past vote in line with the intended vote or current opinion than there is to drift toward the winner or most popular opinion (Himmelweit et al., 1978; Waldahl & Aardal, 1982, 2000; Smith, 1984; Granberg & Holmberg, 1986; Markus, 1986; MacDermid, 1989; Dex, 1995; Joslyn, 2003). Misreporting may also be related to the type of party for which the vote is cast. Votes for smaller parties tend to be more frequently misreported than votes for major parties, due both to party identification and to voter stability (Benewick et al., 1969; Himmelweit et al., 1978). Votes for parties of the extremes—recently the far-right—are also thought to be under-declared (Bachelet, 2007; Hellevik, 2009).

Finally, as with all retrospective data, the passage of time is hypothesized to have an impact on erroneous reporting (Himmelweit et al., 1978; Waldahl & Aardal, 1982, 2000; Dex, 1995; Joslyn, 2003). However, the effect of time may be due to changing views and consistency rather than to memory failure.

In summary, inaccurate recall may be attributed to three phenomena, which may occur concurrently: memory failure, the tendency of voters to misremember a previous vote to reconcile it with how they currently wish to vote (reconciliation), and the social desirability of the declared vote (Waldahl & Aardal, 2000). These factors may be more or less prominent in different elections and, therefore, have different consequences. *Memory failure* causes the distribution of recall data to deteriorate over time. *Reconciliation* favours parties whose support is increasing, and therefore recall of vote for these parties rises with changing voting intentions. *Social desirability* affects the distribution of recall data, notwithstanding the passage of time, and is generally detrimental to parties of the extremes and to populist parties that tend to get bad press.

Poor distribution of recall data may also be due to bias in coverage or sampling. There is no easy way to partial out whether this is the case, but we may postulate that if recall does not deteriorate over time, discrepancies are more likely due to sampling bias or to social desirability. What are the consequences of using recall of previous vote to weight samples and produce estimates of voting intentions? According to Waldahl and Aardal (1982),

"…*recall data overrate the political stability among voters*, particularly in periods with great movements in the electorate […]. The underestimation of changes of party […] signifies that the progress of parties with increasing voters' support is underrated in political barometers, and that the falling off for parties on the decline is underrated." (p. 107-108).

An example of the possible consequences predicted by these authors can be seen in the French presidential elections of 2002, in which the vote for the far-right candidate Jean-Marie Le Pen was underestimated (author reference), and of 2007, in which the same vote was overestimated (author reference). However, the above quote [PC1] presumes that inconsistent voter recall is mostly explained by reconciliation, which is not necessarily the case. In addition, if poor recall is more common among certain specific groups, the impact may vary according to the context of each election.

In which contexts is recall of previous votes reliable and valid? In which contexts is the use of recall data to correct weighting a sound practice? These are the questions that this article seeks to address.

Research Hypotheses

Based on the reviewed literature, our research hypotheses are as follows:

1) ***Regarding reliability of recall of past votes,*** we hypothesize that a) since authors state that the level of inconsistent recall will increase over time, and since our data is recent, the level of inconsistent recall in our data should be higher than what was seen in the reviewed literature; b) Inconsistent recall is mostly related to respondents' socio-political characteristics, i.e., interest in the campaign, stability of preference, and identification with a political party.

2) ***Regarding the validity of recall of past votes,*** we hypothesize that the distribution of recall data is more likely to be biased with the passage of time (memory failure hypothesis), with change in support for the different parties (reconciliation hypothesis) and when the party for which recall is reported is a populist or far-right-wing party or candidate (social desirability hypothesis). The vote for small parties or candidates will also tend to be misreported due to the fact that it is less stable than the vote for the other parties.

3) ***Regarding the impact of adjustment using recall of past votes,*** we hypothesize that the impact will be influenced by the same factors that influence recall validity, i.e., memory failure, reconciliation and social desirability.

Methodology

Data

To test our hypotheses, we had to find data where the questions asked included both voting intention and recall of a previous vote. In addition, to assess recall reliability we needed panel data where the vote recall question was asked at two points in time. Finding such data is not an easy task, since adjustment using recall of past votes is a practice used primarily by private pollsters, and data from these sources are rarely shared with researchers. We also sought some diversity in the data in order to have sufficient variance to test the different hypotheses. Our data came from three sources:

a) Various surveys conducted by the Canadian Election Studies (CES) team since 2000. Canada has had a number of close elections in recent years due to successive minority governments. Elections were held in November 2000, June 2004, January 2006, October 2008 and May 2011. The CES uses a panel design whereby respondents are contacted by telephone before and after each election. The team kept track of some of the 2004 CES survey respondents in subsequent elections, which allows for two different recalls of the 2004 vote by the same respondents. Most of the fieldwork for the CES polls is conducted by an academic pollster.

b) Surveys conducted for the French Presidential elections in 2002 and 2007. The *Panel electoral français* (PEF) is the French academic national election study. For the two-round presidential elections, surveys were conducted before the first round and between the two rounds except for 2002.[1] A private pollster conducted the surveys face-to-face using methods that are usual for pre-electoral polls in France, i.e., probability sampling with quotas (PSQ).

c) Surveys conducted by CROP, a private Quebec pollster with whom we conducted an academic research project. Quebec also had a number of elections over a short period of time due to the election of two minority governments. Elections were held in March 2007, December 2008, September 2012 and April 2014. Two instances of recall of the 2007 vote by the same respondents are available.

Canadian and Quebec elections are parliamentary and follow the British model. French elections are two-round presidential elections where the candidates who finish first and second in the first round are the only ones that advance to the second round. In all, we can compare 40 series of measures of declared past vote to the official results of twelve different elections: five Canadian, four Quebec, and three French. Detailed information on the different samples is available from the authors.

Measures

Recall reliability is assessed by *inconsistent recall.* It is measured by the discrepancy between two reports of the same voting behaviour, including not having voted. It is a dichotomous variable that takes the value of 1 when a respondent states a different past voting behaviour to two different requests. Cases where respondents either

refuse to report their voting behaviour or state that they do not remember it are categorized as missing cases. Inconsistent recall is available for the Canadian election of 2004 and the Quebec election of 2007. For these two measures, the second instance of recall is taken between 18 months and two years after the first. Both follow the election of a minority government and are characterized by a political landscape that is moving beyond the typical two-party choice usually seen in “first past the post” types of electoral systems. In terms of predictors, *age* is in three categories, under 35, 35 to 54, and 55 and over. *Education* is measured in years of schooling. Socio-political characteristics are *interest in the campaign* for which recall is asked, *stability of voting intentions*—measured by the fact that a respondent had similar voting intentions (including intention not to vote) in two different elections[2]—*voting intention in the campaign for which recall is asked*, *voting intention for the next election*, *presence of party identification,* and *strength of party identification* (CES only).

There are 40 series of *declared past vote*, with 160 entries overall, since each series comprises the information for the three main parties or candidates and for the smaller parties grouped together. There are two exceptions: Canada 2004, where there were four main parties, and the second round of the French presidential election of 2007, where there were only two candidates. Each main party/candidate is characterized as either far-right-wing or populist, conservative, center, or left-wing; all the other parties are grouped together in a “small party” category. We characterize as far-right-wing or populist any parties or candidates whose main platforms are identity politics and the struggle against “big government.” These parties are clearly not identical, but they appeal to the same type of voter, and their support tends to vary substantially between elections. Such a party or candidate is present in all the French and Quebec elections; for Canada, it is only present in the 2000 election.

The number of series of past vote recalls varies with each election, from one (Canada 2011) to 17 (Quebec 2007). The mode is two recalls per election. Retaining all the data would have been redundant because for any given election, one entry is the exact combination of the other entries. Therefore, the information for the left-wing parties or candidates has been withdrawn, since this type of party is the only one to appear in all elections. This leaves 119 entries for analysis, since there is no recall data available for the second round of the 2007 French presidential election.

There are two measures of *bias in voter recall*, i.e., relative difference and absolute difference between the recall for each party or candidate and the vote. *Relative difference* indicates the direction of bias. A negative result means that the vote for a party or candidate is under-recalled. *Absolute difference* indicates the level of bias. The measure of *elapsed time* is the number of months between the vote and its recall. This is used to test the memory failure hypothesis. *Vote change* is computed as the difference between the vote for which recall is asked and the vote in the subsequent election. A positive measure occurs when the vote share for a given party has increased. It is used to test the reconciliation hypothesis and therefore only for the measure of relative error. Since the change may not yet have occurred at the moment when recall is reported, it is not an ideal measure, but it is the best one available and it provides an idea of whether support for a given party tends to be unstable. In order to assess the social desirability hypothesis, i.e., whether the type of party or candidate makes a difference, the fact that *the party or candidate is far-right or populist* is used. In addition, the fact that the party is *a small party* is entered, as this is hypothesized to deteriorate recall.

Estimations of voting intention for the next election using recall of past voting are available for eight elections: four Canadian, two Quebec and the first and second rounds of the 2007 French election. In Canada/Quebec, data are originally weighted according to the most recent census information for cells formed by age group, gender and region, together with mother tongue in the case of Quebec. In France, quotas are implemented based on age group, gender, region/size of community, and profession of the head of household. In order to assess the impact of using recall of the previous vote to adjust estimates, the original weights are corrected so that recall of past vote reflects the original vote perfectly. To compute the dependent variable, we first calculate the difference between each estimate of voting intention—census weighted and corrected according to recall—and the vote. The dependent variable is computed as the *difference between these two estimates*. A positive difference occurs when the correction based on recall improves the estimates while a negative difference means that the correction worsens the estimate. The independent variables are *vote* (as a control), *elapsed time between elections*, *difference between the vote and the vote in the preceding election*, and *type of party* (far-right wing or populist, small party) for which voting intention is estimated.

Analysis

To assess which factors are related to inconsistent recall, logistic regressions are performed for each instance of double recall. We present the results of the bivariate regressions for each variable and those of the parsimonious

final models. This allows us to examine the net contribution of each variable as well as their remaining combined effect.

To assess the predictors of both the validity of recall data and the possible improvement provided by adjustment using recall of past voting, linear regressions are performed with standard errors corrected for clusters of election and recall spells. The predictors are entered hierarchically (control variable, time, vote change, type of party). The results of each regression allows us to assess the contribution of each variable. The process leads to a parsimonious final model.

Results

Reliability of recall data

The prevalence of inconsistent recall is 21% for the Canada 2004 election and 25% for the Quebec 2007 election. Our data are in the upper end of the range but do not stand out compared to recent figures reported in the literature (7 to 39%, with a usual range of 10 to 25%). The hypothesis of an increase in inconsistent recall over time due to higher volatility in the electorate, as suggested by Waldahl and Aardal (2000), is not confirmed, at least in our data, even though both elections are characterized by movement in support for the different parties.

Table 1. Predictors of inconsistency in recall

		Canada – 2004			**Quebec – 2007**		
		Bivariate		**Parcimonious**	**Bivariate**		**Parcimonious**
		exp(b)		**exp(b)**	**exp(b)**		**exp(b)**
Age (REF=18-34)			*				
35-54		0.568	*	0.499 *	1.094		0.809
55 +		0.517	**	0.601	0.641		0.413 *
	N	1304			542		
R^2 Nagelkerke		0.010			0.016		
Education		0.778	**		0.666	**	0.641 **
	N	1290			542		
R^2 Nagelkerke		0.007			0.033		

Interest in campaign	0.925	**				0.632	***	0.764	*	
N	1302					542				
R^2 Nagelkerke	0.011					0.047				
Voter stability (REF=unstable)										
Stable	0.173	***	.193	***		0.234	***	0.254	***	
N	1018					467				
R^2 Nagelkerke	0.145					0.128				
Voting intention A (REF=Liberals)		***		**	**Voting intention A (REF=LPQ)**		***			
CP	1.634	*	1.815	*	**ADQ**	2.661	***			
NDP	3.176	***	3.605	***						
BQ	0.946		1.260		**PQ**	1.034				
Spoiled/Others	4.250	***	1.777		**Spoiled/Others**	1.575				
DK/Will not tell	2.330	**			**DK/Will not tell**	2.053				
N	1297					542				
R^2 Nagelkerke	0.059					0.048				
Voting intention B (REF=Liberals)		**			**Voting intention B (REF=LPQ)**		**			
CP	0.867				**ADQ**	0.861				
NDP	1.354									

BQ	0.618		**PQ**	1.645	*	
Spoiled/Others	2.098	**	**Spoiled/Others**	1.736		
DK/Will not tell	1.755	†	**DK/Will not tell**	3.264	**	
N	1296			542		
R^2 Nagelkerke	0.026			0.040		
Party identification (REF=Liberals)		***				
CP	0.662	†				
NDP	1.448					
BQ	0.770					
Other Party	1.098					
None of these	2.437	***				
N	1111					
R^2 Nagelkerke	0.043					
Strength of identification	0.521	***	0.514	***		
N	1072					
R^2 Nagelkerke	0.088					
N total			**870**			**467**

R^2	Cox & Snell	0.143	0.129
	Nagelkerke	0.254	0.194

†:p<.10; *: p<.05; **: p<.01; ***: p<.001.

Understanding inconsistent recall may help comprehend why there is bias in the distribution of recall data. Hypothesis 1 posits that inconsistency in recall is primarily explained by socio-political variables. In both instances, either age alone or both age and education remain significant in the final parsimonious regression model. Older and more educated people are more likely to consistently recall their previous voting behaviour. Interest in the campaign is a significant predictor of consistency for both models. Voter stability appears to be the most substantial predictor in both models, explaining 14.5% of the variance for the Canada 2004 model and 12.8% of the variance for the Quebec 2007 model. For the Canada 2004 election, inconsistent recall is concurrent with having intended to vote for the two parties whose support was increasing between the two elections. Finally strength of party identification (available for Canada only) is a coherent and strong predictor of consistency in recall (8.8% of variance explained in bivariate regression). Bivariate regressions show that voting intention for the next election is significant in both regressions: Undecided voters or people who do not reveal their intentions are the most likely to be inconsistent when reporting past behaviour.

Finally, the total explained variances for the parsimonious models are 25.4% for Canada 2004 and 19.4% for Quebec 2007. Therefore, except for the fact that the impact of socio-demographic variables is not totally mediated by socio-political variables, our hypothesis is generally supported by the data. People who are interested in the campaign, more likely to identify with a political party and, more importantly, stable in their preferences are less likely to be inconsistent in their recall.

Validity of recall data

The results presented in the upper portion of Table 2 pertain to relative error in recall and therefore to the direction of error. They show that on average, recall for non-left-wing parties—the only type of party retained for analysis—is one point lower than the vote (Model 0), which means that overall, the vote for left-wing parties or candidates tends to be over reported. However, this difference becomes non-significant when we control for the vote (Model 1). After controlling for the level of vote for each party, the passage of time by itself is not significantly related to the direction of recall error (Model 2). Reconciliation, however, explains part of the recall bias. Model 3 shows that for each point increase in the vote for a given party in the next election, recall of that vote is over-declared by 0.176 points. Finally, recall is affected by the type of party/candidate. Results presented in Model 4 show that the vote for far-right-wing or populist parties or candidates is on average 6.7 points less than for other parties. The fact that the party is small is not significantly related to relative error. After entering the variables for type of party, the coefficient for elapsed time remains small but becomes significant. Each month elapsed between the vote and its recall results in an average under-declaration of 0.03 points for non-left-wing parties. The total variance explained by the variables present in the model is 35.5%. In the final parsimonious model, elapsed time, vote change for the various parties between elections, and the fact that a party or candidate is far-right or populist are retained in the analysis. In this model, the intercept shows that there is a significant one point over-declaration of the vote for the different non-left-wing parties. This is counterbalanced by a decrease of 0.03 points in the declaration of the vote for a party with each month elapsed between the vote and its recall. The impact of the vote change for a given party becomes non-significant. It is retained because it improves the overall model. Given the fact that it was not an ideal measure of movement in support between elections, the fact that its contribution would nevertheless appear to be significant at one point is good news. Finally, far-right or populist parties tend to be under-declared by as much as 5.5 points. These variables explain 33.4% of the variance.

Table 2. Predictors of vote recall validity

Relative error

	Model 0	Model 1	Model 2	Model 3	Model 4	Model 5
Intercept	-1.000 ***	0.552	0.938	0.354	4.484 *	1.025 ***
Vote		-0.064 †	-0.063 †	-0.030	-0.105	
Elapsed time			-0.029	-0.029	-0.031 *	-0.031 *
Diff. vote following election				0.176 **	-0.035	0.099
Populist / Right wing					-6.696 ***	-5.481 ***
Small party/candidate					-2.027	
N	**119**	**119**	**119**	**116**	**116**	**116**
Adjusted R^2	**n/a**	**2.5% †**	**3.2%**	**11.9% ***	**35.5% *****	**33.4% *****

Absolute error

	Model 0	Model 1	Model 2	Model 3	Model 4
Intercept	3.506 ***	1.833 ***	0.391	-2.393	-0.267
Vote		0.069 *	0.067 *	0.114 *	0.063 *
Elapsed time			0.108 ***	0.108 ***	0.108 ***
Populist / Right wing				3.702 ***	2.856 ***
Small party/candidate				2.015	
N	**119**	**119**	**119**	**119**	**119**

	n/a	5.6%	23.2%	38.6%	35.8%
Adjusted R^2		*	***	***	***

†:p<.10; *: p<.05; **: p<.01; ***: p<.001.

We checked whether the results were robust by withdrawing recall of the French elections—where the electoral system is very different from the Canadian system—from the analysis. We also did this for the Canada 2000 election, where there were major changes in the configuration of political parties between the two elections. The results are basically the same.

Absolute error provides a measure of the total discrepancy between the vote and its recall. The analyses are presented in the lower portion of Table 2. Model 0 shows that absolute error is 3.5 points on average. When vote is added as a control (Model 1), the coefficient remains significant, and vote is also significant. The level of error is therefore proportional to the level of support. Model 2 shows that absolute error increases by .108 points for each month elapsed between the vote and its recall, which confirms the validity of the memory failure hypothesis. This model explains 23.2% of the variance. In addition, absolute error is 3.7 points higher for far-right-wing/populist parties or candidates. The coefficient is not significant for small parties or candidates. The complete model explains 38.6% of the variance. The final parsimonious model retains all of the variables except the measure of small parties/candidates and explains 35.8% of the variance. The results are robust. When the French elections and the Canada 2000 election are withdrawn from the analysis, the variables stay significant and the level of explained variance is similar.

In conclusion, hypothesis 2 is confirmed. Relative error in recall is related to reconciliation and to the type of party for which the vote was cast, either because of social desirability or because of coverage or sampling bias. It is only slightly related to memory failure. In addition, absolute error does increase substantially with the passage of time, and it also increases with far-right or populist parties and candidates. Finally, small parties are not generally more subject to bias in the declaration.

The impact of using recall of previous votes to adjust estimates of voting intention

Table 3 shows the results of regressions with the difference between estimates as the dependent variable. Model 0 shows that the overall average difference between the two estimates is not different from zero, which means that using recall of past votes may improve some estimates while deteriorating others but that the overall impact is null. The vote for a given party does not have an impact on the difference (Model 1) nor does time elapsed between the preceding election and the one for which voting intentions are estimated (Model 2). However, Model 3 shows that if support for a given party has increased compared with the preceding election—the one for which recall is asked—the correction is more likely to improve the estimate of voting intentions. Overall, this model explains 14.2% of the variance. Finally, Model 4 shows that when estimating the vote for far-right or populist parties or candidates, using the correction worsens the estimate of voting intentions.

Table 3. Difference between estimates corrected using recall of past vote and census weighted estimates of vote intention

	Model 0	Model 1	Model 2	Model 3	Model 4	Model 5
Intercept	-0.554	0.722	-0.353	0.777	2.174	0.439
Vote		-0.048	-0.041	0.021	-0.026	
			0.026	-0.034	-0.019	

Elapsed time with preceding election						
Diff with vote in the preceding election				0.123 *	0.127 †	0.122 †
Populist/ far-right wing party/candidate					-2.969	-2.569 †
Small party/candidate					-0.958	
N	**22**	**22**	**22**	**21**	**21**	**21**
R^2		**2.76%**	**4%**	**14.28%**	**22.56%**	**0.2012**

† <.10 ; *: p<.05; **: p<.01; ***: p<.001.

Since there are only at most 22 entries, clustered within seven elections, with many independent variables and collinearity in the data, it is difficult to arrive at a sound parsimonious model. Model 5 explains 20.1% of the variance with two variables, i.e., the difference between the vote in the preceding election and the current vote, and the type of party for which the estimate is performed. However, the overall contribution of the variables is not significant (F=3.02, p=0.12).

These analyses lead us to conclude that there is little difference between corrected and uncorrected estimates of voting intention. There is scant evidence that either a large vote change between elections or the fact that the party/candidate whose support is estimated is far-right or populist make it more or less appropriate to correct using voter recall.

Discussion and conclusion

The results presented in this paper confirm that voting is not a salient, memorable behaviour among all voters. People who always vote the same way and who identify with a political party are likely to accurately remember their vote, but in most developed countries, change of allegiance is now common. A substantial portion of the electorate changes its mind between or during campaigns, and switchers seem to have more difficulty remembering how they voted. Recall error is not random. Voters' misremembering a previous vote to reconcile it with how they currently wish to vote (reconciliation) and, above all, the difficulty in reaching voters for far-right or populist parties/candidates, and in convincing them to reveal their true vote or voting intention are the main explanations for error. Memory failure also plays a role, but the overall impact on the direction of error is weak. Finally, reconciliation and the type of party for which the vote is cast also play roles when it comes to correcting estimates using recall of past voting, but the overall impact is weak.

The analyses presented in this paper confirm the hypotheses proposed by various authors. These authors usually had access to data from only one country and from one to three elections. The fact that we could validate these hypotheses using data from 12 elections, 40 series of recalls, in three constituencies, two countries and two electoral systems, lend confidence to their robustness.

There are some flaws in the research, however. Pollsters who adjust using recall of past voting usually collect their data using quota samples. Most of the data used in this research use RDD samples. In addition, the categorization of "right-wing populist" (RWP) might also be questioned, since we have placed Le Pen in France, the Canadian Alliance in Canada, and the Action Démocratique du Québec (ADQ) in Quebec in the same category. Le Pen is the most far-right wing candidate, followed by the Canadian Alliance. However, though ADQ policies focused on identity and fear of immigration, it could be categorized more as a populist than a far-right-wing party. We hypothesized that these parties or candidates may have the same effect because they appeal to the same type of voters. The fact that the results pertaining to RWP parties or candidates are robust tends to

confirm the validity of our hypothesis. Finally, the weighting correction we performed is a simplified version of what some pollsters do. These pollsters sometimes correct iteratively using numerous recalls of past votes. We could not use such sophisticated schemes when comparing a substantial number of elections held in diverse situations.

What, then, is the final conclusion? Should we or shouldn't we correct the weightings using recall of past votes? Since we know the conditions under which recall is likely to be biased, the bias might be first used to understand changes in support. For example, when voter recall overestimates support for a given party, it usually means that support for that party is increasing. And obviously, poor recall of voting for far-right and populist parties should be examined with caution. This is present for all three types of RWP parties or candidates, even though the elections themselves and the sampling methods used were different. Researchers and pollsters alike should strive to better understand this phenomena. Is it due to selection bias, to non-response bias or to social desirability? Is it always present at the same level in all elections? How might we correct for such bias or eventually eliminate it? The *Panel electoral français* of 2007, conducted face-to-face, used a ballot box to collect information on recall of past voting with some success. This procedure substantially reduced the bias in the recall of the Le Pen vote. Another avenue is to attribute the vote of non-disclosers non-proportionally to the different parties. This method is used in Quebec to correct for the known systematic bias in the estimation of voting intention with some success.

Is correction using recall of previous votes a good practice? The results presented here show that, at best, it does not have a significant, substantial, impact on the estimates. We suggest that the energies of researchers and pollsters would be better used to tackle the problem itself instead of working on *a posteriori* correction.

---

[1] The candidates for the second round of the 2002 election were Jacques Chirac (UMP) and the far-right-wing candidate Jean-Marie Le Pen. The election ended with 82 percent for Chirac. The results are too exceptional to be used in this analysis.

[2] This concept is usually measured in the literature as a similar voting behaviour in two elections. We feel that this is a problem, since the same variables are used to compute inconsistent recall and stability of the vote, creating a problem of endogeneity.

References

1. Anderson, H. E., & Granberg, D. (1997). On the Validity and Reliability of Self-Reported Vote: Validity without Reliability? *Quality & Quantity, vol. 31*(2), 127-140.
2. Bachelet, D. (2007). Propositions de nouvelles méthodes pour estimer les intentions de vote: Application au duel Le Pen – Jospin dans le premier tour de l'élection présidentielle de 2002. *Revue française de marketing* (Février 2007), 11-30.
3. Benewick, R. J., Birch, A.H., Blumer, J.G., Ewbank, A. (1969). The Floating Voter and the Liberal View of Representation. *Political Studies, 17*(2), 177-195.
4. Dex, S. (1995). The Reliability of Recall Data, a Literature Review. *BMS, 49*, 58-89.
5. Foekema, H. (2009). [Personal communication].
6. Granberg, D. & Holmberg, S. (1986). Prior Behavior, Recalled Behavior and the Prediction of Subsequent Voting Behavior in Sweden and the U.S. *Human Relations, 39*(2), 135-148.
7. Hellevik, O. (2009). *Political Weighting of Party Polls: Experiences from Norway*. Paper presented at the 62nd annual WAPOR conference, Lausanne, Switzerland.
8. Himmelweit, H. T., Biberian M. J., Stockdale, J. (1978). Memory of Past Vote: Implications of a Study of Bias in Recall. *British Journal of Political Science, 8*(3), 365-375.
9. Joslyn, M. R. (2003). The Determinants and Consequences of Recall Error about Gulf War Preferences. *American Journal of Political Science 47*(3), 440-452.
10. Jowell, R., Hedges, Barry, Lynn, Peter, Graham, Farrant, Heath, Anthony. (1993). The 1992 British Election: The Failure of the Polls. *Public Opinion Quarterly, 57*, 238-263.
11. Katosh, J. P. & Traugott, M. W. (1981). The Consequences of Validated and Self-Reported Voting Measures. *Public Opinion Quarterly, 45*(4), 519-535.
12. MacDermid, R. H. (1989). The Recall of Past Partisanship: Feeble Memories or Frail Concepts? *Canadian Journal of Political Science, 22*(2), 363-375.
13. Markus, G. B. (1986). Stability and Change in Political Attitudes: Observed, Recalled, and "Explained". *Political Behavior, 8*(1), 21-44.
14. Smith, T. W. (1984). Recalling Attitudes: An Analysis of Retrospective Questions on the 1982 GSS. *Public*

Opinion Quarterly, 48(3), 639-649.
15. Taylor, H. (1995). Horses for Courses: How Survey Firms in Different Countries Measure Public Opinion with Different Methods. *Journal of the Market Research Society, 37*(13), 211-219.
16. Traugott, M. W., & Katosh, J. P. (1979). Response Validity in Surveys of Voting Behavior. *Public Opinion Quarterly, 43*(3), 359-377.
17. Van der Eijk, C. & Niemoller, B. (1979). Recall Accuracy and its Determinants. *Acta Politica, 14*, 289-342.
18. Waldahl, R., & Aardal, B. O. (1982). Can We Trust Recall-Data? *Scandinavian Political Studies, 5*(2), 103-116.
19. Waldahl, R., & Aardal, B. O. (2000). The Accuracy of Recalled Previous Voting: Evidence from Norwegian Election Study Panels. *Scandinavian Political Studies, 23*(4), 373-389.
20. Weir, B. T. (1975). The Distortion of Voter Recall. *American Journal of Political Science, 19*(1), 53-62.