

# **Should we use recall of previous vote(s) to weight electoral polls?**

Presented at the Wapor/Aapor joint Annual Conference

Chicago, May 11-13, 2010

By Claire Durand, Isabelle Valois and Mélanie  
Deslauriers,  
Université de Montréal

# Presentation

- ◆ A bit of History
- ◆ Literature : Three questions
- ◆ Methodology
- ◆ Is declared past vote reliable?
- ◆ What are the causes of bad recall?
- ◆ What are the consequences of adjustment using voter recall?
- ◆ Conclusion

# A bit of history

## Adjusting according to declared past vote

- ◆ Started in 1948 (Hellevik, 2009) and is still used in Europe, mostly where quota sampling is used.
- ◆ Was first developed to compensate for the under estimation of the vote for Communist parties after the Second World War (Hellevik, 2009)
  - ◆ And has been recently used to correct the under estimation of the extreme-right vote.
  - ◆ Was recommended in 1992 (Market Research Society, GB 1992) in order to compensate for socio-political biases in the samples

# Three questions...

- ◆ 1. Is declared past vote reliable?
- ◆ 2. What are the causes of unreliable recall?
- ◆ 3. What is the impact of adjusting by declared past vote on the accuracy of the estimation of the vote?

# 1. Is declared past vote reliable?

## Literature

- ◆ Literature of the question states that recall of past voting behavior is often biased:
  - ◆ It tends to underestimate the support for the extreme right or left-wing parties.
  - ◆ It tends to underestimate the support for small parties.
  - ◆ It tends to be biased towards the winner or towards the party the respondent intends to vote for.
  - ◆ The quality of recall vary with the time interval since the preceding election(s).

## 2. What are the causes of a bad recall?

Literature points to ...

- ◆ The sample is unbiased but there are
  - ◆ Memory problems: respondents make involuntary mistakes when they recall for whom they voted in the last election.
  - ◆ Dissimulation : respondents either do not answer the question or voluntarily give an inaccurate answer.
- ◆ The sample is biased: not all members of the population have the same chance to be part of the sample. This is mostly due to the sampling procedure or to the propensity of some respondents to refuse to participate in surveys.

### 3. Specific causes lead to specific consequences

In the distribution of voter recall

- ◆ Memory problems may lead to an increase in inaccuracy of the distribution of voter recall with time.
- ◆ Dissimulation and sample bias may both lead to the under-representation of some political parties whatever the time interval between elections.

### 3. Specific causes lead to specific consequences

In the estimation of voting intention

- ◆ When voting intentions are stable, if the bad distribution of voter recall depends on random non-response, using voter recall to adjust may improve the estimates.
- ◆ When voting intentions are not stable and when specific groups are not randomly underrepresented, use of recall data to adjust may deteriorate the quality of estimates (Waldahl et Aardal, 1982):
  - ◆ Underestimation of vote for a party when intention has been rising and overestimation of vote when intention has been decreasing.
  - ◆ Happened in France, Presidential 2002, Presidential 2007 and Regional election 2010 at the Regional



# Aim of research

- ◆ Revisit this question (that was not been studied for a while except for Hellevik, 2009).
  - ◆ Is the recall of past voting behavior biased? In which circumstances?
  - ◆ Are the predictors of bad recall consistently the same?
  - ◆ Can we determine in which contexts the use of recall data may create bias and how we can minimize it?

# Methodology

3 cases : case no 1

## ◆ Quebec: Elections of March 2007 and December 2008

### ◆ Context :

- ◆ Poor prediction of the results for the March 2007 election.
- ◆ Strong progression of voting intention for a right-wing populist party (ADQ): on the rise before the 2007 election and declining afterwards.

### ◆ Information:

- ◆ Measure of recall for 2003 and 2007
- ◆ Measure of bad recall : Declared vote for the March 2007 election in April 2007 and November 2008.
- ◆ Concurrent evolution of recall and voting intentions (2007-2010).
- ◆ Impact of adjustment according to declared past vote.

# Methodology

3 cases : Case no 2

- ◆ French presidential elections- April 2002 and April 2007
- ◆ Context:
  - ◆ Underestimation of support for Le Pen, a right-wing populist candidate, in 2002, and overestimation of the same vote in 2007.
  - ◆ Support for Le Pen increased before the 2002 election and decreased afterwards.
- ◆ Information:
  - ◆ Recall of past vote collected after the first round of the presidential election by different pollsters (Durand, et al. 2004 & Durand, 2008).
  - ◆ Recall of vote cast in the first round of the Presidential elections of 2002 and 2007 (French electoral panels, PEF 2002, 2007)
  - ◆ Impact of adjustment by declared past vote on the quality of the estimation of the vote (French electoral panel for 2007).

# Methodology

Three cases : Case no 3

- ◆ Canada: 2004, 2006 and 2008 elections (Canadian Election Study) leading to minority governments
- ◆ Context:
  - ◆ Multiple elections with minority governments (June 2004, January 2006, October 2008).
  - ◆ Information:
    - ◆ Recall of past votes for 2000, 2004, 2006 and 2008.
    - ◆ Measure of bad recall : Recall questions asked twice for 2004 and 2006.
    - ◆ Impact of adjustment by declared past vote on the quality of estimates of the vote in 2004, 2006 and 2008.

# Recall of past vote

## Quebec 2003 and 2007

	Quebec 2003			Quebec 2007			
	Vote	Declared vote in 2007 (CROP>	Declared vot> in CES04	Vote	Declared vote After 2007 elect>	Declared vot> Crop - nov 0>	Declared vote Subsample 2008
PLQ	46,0%	43,1%	44,7%	33,1%	29,8%	33,8%	36,3%
PQ	33,2%	41,9%	44,1%	28,4%	29,6%	34,9%	34,8%
ADQ	18,2%	10,6%	9%	30,8%	32,1%	23,1%	21,6%
Others	2,6%	4,4%	2,2%	7,7%	8,5%	8,2%	7,3%
Total	100%	100%	100%	100%	100%		
Void vote	1,25%	0%		1%	0%	0	0,1%
Didn't vote>	29,58%	17%	12,6%	28,77%	18,8%	15,3%	13,4%

- Declared vote is reliable only right after the election in 2007.
- The deterioration of recall is substantial for ADQ (right wing populist party).

# Recall of past vote

## French electoral Panel (PEF) France 2002 & 2007

France 2002				France 2007 (PEF)		
	Vote	Declared vote in 2002 (pef)	Declared vo> in 2007 (pef>		Vote	Declared vote after the 1st roun>
Daniel GLUCKSTEIN	0,7%	0,4%	0,4%	Gérard Schivardi	0,3%	0,3%
Arlette LAGUILLER	5,7%	6,4%	3,9%	Arlette Laguiller	1,3%	2,0%
Olivier BESANCENOT	4,3%	5,2%	6,8%	Olivier Besancen>	3,4%	5,7%
Robert HUE	3,4%	3,9%	3,3%	Marie-George Buffe>	1,9%	3,6%
Lionel JOSPIN	16,2%	19,7%	29,8%	José Bové	1,1%	1,7%
Christiane TAUBIRA	2,3%	2,2%	1,8%	Ségolène Royal	25,9%	31,4%
Jean-Pierre CHEVEN>	5,3%	5,9%	2,7%	Dominique Voynet	1,6%	1,2%
Noël MAMERE	5,3%	7,1%	4,4%	Frédéric Nihous	1,1%	0,8%
Corinne LEPAGE	1,9%	1,4%	0,9%	François Bayrou	18,6%	21,5%
François BAYROU	6,8%	7,9%	4,3%	Nicolas Sarkozy	31,2%	23,5%
Jacques CHIRAC	19,9%	21,5%	27,4%	Philippe de Villier>	2,2%	1,8%
Alain MADELIN	3,9%	2,8%	1,1%	Jean-Marie Le Pen	10,4%	6,4%
Christine BOUTIN	1,2%	1%	0,1%			
Jean SAINT JOSSE	4,2%	3%	0,9%			
Bruno MEGRET	2,3%	1,3%	0,3%			
Jean-Marie LE PEN	16,9%	10,3%	11,9%			
Total	100%	100%	100%			
Void vote	2,4%	1,7%	3,8%		1,2%	
Didn't vote	28,4%	14,1%	9,6%		16,2%	

- For 2002 as well as 2007, distribution of declared vote for LePen is mediocre, even immediately after the election.

# Recall of past vote

## Private pollsters - France 2002 (Durand et al., 2004)

Election	Reported vote					
	First round presidential 2002	Sofres April 21 (election day)	Sofres May 31 & June 1	Ipsos April 30 & May 2	Ipsos June 5	CSA June 5-6
	%	%	%	%	%	%
<b>Extreme-left</b>	10.5	8.7	9.5	9.5	13.7	7.4
<b>Traditional Left</b>						
Hue	3.4	4.7	3.7	3.1	3.8	3.4
Chevènement	5.3	6.9	4.8	5.0	4.9	4.2
Taubira	2.3	3.5	2.4	1.8	2.3	1.6
Jospin	16.2	18.4	24.4	20.0	22.3	22.5
Mamère	5.3	7.1	6.8	8.0	6.5	7.6
<b>Traditional Right</b>						
Bayrou	6.8	10.7	7.8	7.5	6.8	8.7
Boutin	1.2	0.8	1.0	1.0	0.8	1.4
Chirac	19.9	19.4	22.7	23.0	24.7	27.5
Lepage	1.9	1.8	1.3	2.1	1.8	0.9
Madelin	3.9	3.5	2.5	3.7	2.5	2.2
Saint-Josse	4.2	5.4	2.4	4.0	2.5	3.8
<b>Extreme-right</b>	19.2	9.2	10.8	11.1	7.8	8.0
<i>Proportion who report their vote</i>				71.2	78.8	

The vote for Le Pen is under-represented by all the pollsters (by approximately half) even on election day (Sofres)

# Recall of past vote

## Private pollsters - France 2007 (Durand, 2008)

**Table 4**  
**Reported Vote in the First-Round Election in Polls Conducted**  
**at the Beginning and at the End of the Between-Round Period**

	Reported vote April 22-25			Reported vote May 2-4					
	vote	C.I.	BVA	IPSOS	IFOP	BVA	IPSOS	IFOP	CSA
<b>Sarkozy</b>	<b>31.2</b>	28.3, 34.1	31.1	33.9	31.1	29.8	29.4	32.0	31.1
<b>Royal</b>	<b>25.9</b>	23.2, 28.6	31.5*	25.4	28.2	29.5*	28.9*	27.6	25.5
<b>Bayrou</b>	<b>18.6</b>	16.2, 21.0	20.1	19.9	20.4	21.0	22.3*	22.3*	23.3*
<b>Le Pen</b>	<b>10.4</b>	8.5, 12.3	3.3*	7.1*	5.9*	4.5*	5.1*	4.2*	5.6*
<b>Total right</b>	<b>63.6</b>	60.6, 66.6	57.5*	63.3	60.2*	58.1*	60.0*	61.2	63.3

\*: estimate outside the confidence interval at 95%

Note: Ipsos grouped Le Pen with another extreme right-wing candidate. We subtracted the share of this candidate from the estimate as if the estimate were perfect. CSA's estimates included non voters; we recalculated the proportions excluding them.

- The declared vote for LePen is highly under-represented between the two rounds of the election and after the second round.



# Recall of past vote

## Canadian Election Study - Canada 2000 and 2004

	Canada 2000		Canada 2004			
	Vote	Declared vote in 2004	Vote	Declared vote in 2004	Declared vote in 2006 (pre, rdd>	Declared vote in 2006 (post)
PLC	40,8%	52,3%	36,7%	35%	47,5%	44,9%
PC (PPC+all>	37,7%	24,5%	29,6%	31,7%	25,3%	26,9%
NPD	8,5%	10,4%	15,7%	17%	12,3%	13,2%
Bloc QC	10,7%	11,5%	12,4%	11,5%	11,2%	11,4%
Others	2,2%	1,3%	5,6%	5,2%	3,7%	3,7%
Total	100%	100%	100%	100%	100%	100,1%
Didn't vote	39,1%	21,6%	39,1%	15%	17,6%	12,3%

- Declared vote is accurate right after the election.
- However, with time, the distribution of recall deteriorates especially among the conservative - PC (underestimation) and the liberal- PLC voters (overestimation).

# Recall of past vote

## Canadian election study - Canada 2006 and 2008

	Canada 2006			Canada 2008		
	Vote	Declared vote in 2006	Declared vot> in 2008 (pre>	Vote in 2008	Declared vot> in 2008	
PLC	30,2%	28,1%	30,5%	26,3%	24,8%	
PC (PPC+allian>	36,3%	38%	37,8%	37,7%	37,5%	
NPD	17,5%	17,9%	14,1%	18,2%	18,1%	
Bloc QC	10,5%	10,2%	12,3%	10%	12,1%	
Others	5,5%	5,7%	5,3%	7,8%	7,5%	
Total	100%	100%	100%	100%	100%	
Didn't vote	35,3%	10,3%	17,1%	41,2%	14,7%	

- Distribution of declared vote is accurate, even two years after (vote of 2006 declared in 2008) for all the parties.

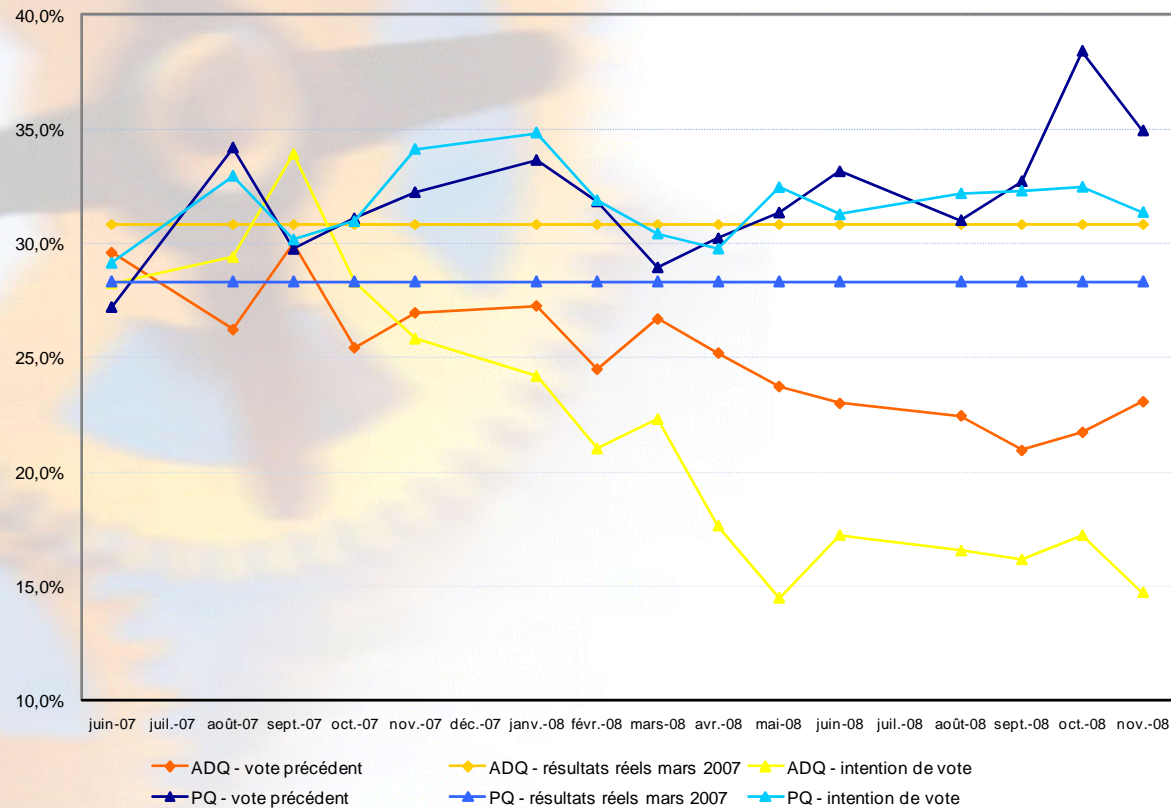
# 1. Is recall of past voting behavior reliable?

## Synthesis

- ◆ Where a right-wing/ populist party/candidate is present (Quebec, France), this vote is generally under-represented in voter recall.
  - ◆ However, for Quebec, it is not under-represented right after the election,
  - ◆ While in France, it is always under-represented, even on election day.
- ◆ For Canada,
  - ◆ The vote for the Conservative party in 2004 was under-represented in declared vote in 2006
  - ◆ but the same vote in 2006 was not underrepresented in declared vote after the elections of 2006 and 2008.

# Causes of bad recall

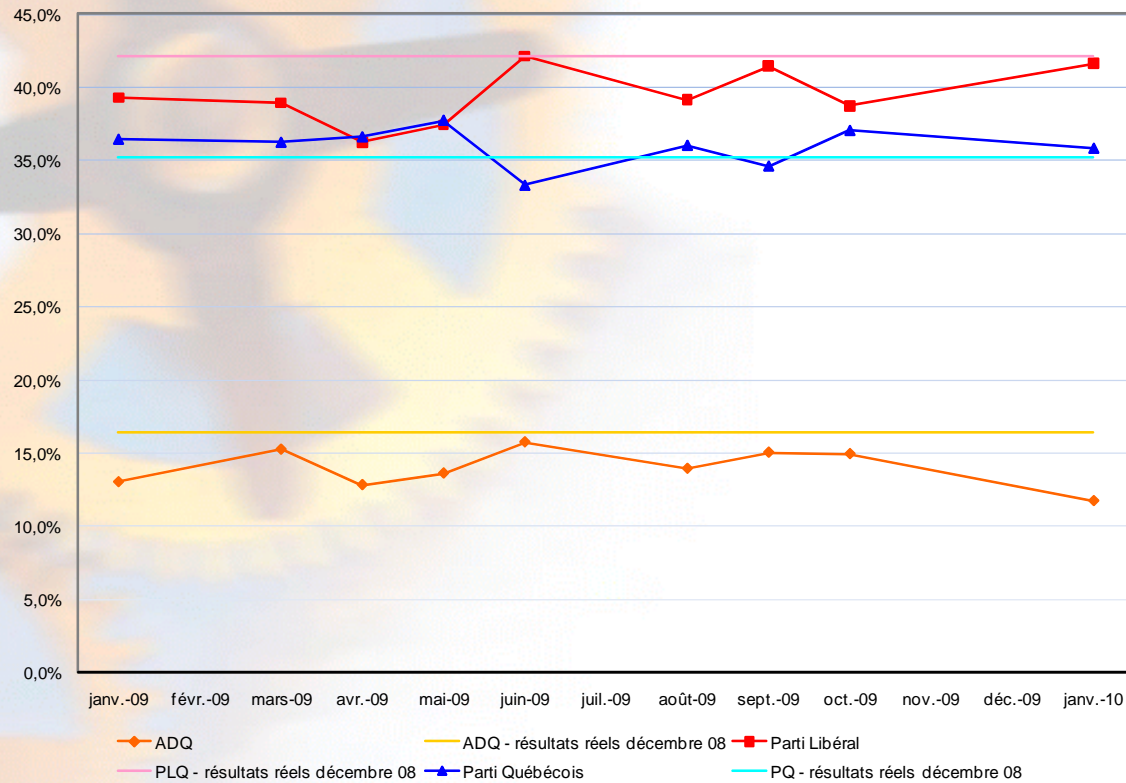
Quebec 2007-2008



- Recall follows voting intentions: decrease of declared vote for ADQ follows the decrease in voting intentions.
- Both are stable for the PQ.

# Causes of bad recall

## Quebec 2008-2010



- Recall follows voting intentions: no change in voting intentions, no change in recall of past vote.

## 2. What are the causes of bad recall?

Aggregate level : Quebec between March 2007 and November 2008

- ◆ Logistic regressions (18 surveys between June 2007 and November 2008)
- ◆ After controlling for socio-demographic factors related to voting intention, recall does vary significantly
  - ◆ As a function of time : recall of past vote for ADQ decreases with time.
  - ◆ As a function of voting intentions : recall of past vote for small parties decreases as voting intention for PQ (nationalist party) increases.

## 2. What are the causes of bad recall?

At the individual level

- ◆ Bad recall for Quebec 2007-2008 (25%), Canada (CES) 2004-2006 (21%) & 2006-2008 (29%) have the same characteristics:
  - ◆ age & education : indirect effect mediated by interest in the campaign for CES 2006-2008
  - ◆ Interest for the initial campaign (-)
  - ◆ Identification with a political party and strenght of identification (-) (CES only)
  - ◆ Stability of political opinions (-) (not available for CES 2006-2008)
  - ◆ Voting intentions:
    - ◆ For ADQ (right wing populist party) for Quebec 2007 (vs other party)
    - ◆ For small parties vs traditional parties (CES)

### 3. What is the impact of adjusting using recall of past vote?

#### Quebec 2007 & 2008

- ◆ Adjusting by declared vote at the preceding election...
  - ◆ In 2007, while support for ADQ is rising:
    - ◆ Overestimation of ADQ vote (28.6% vs à 25.9%). Final vote: 30.8%.
    - ◆ Underestimation of PQ vote (24.3% vs 28.3%). Final vote: 28.3%.
    - ◆ Adjusting gives a better estimate of the ADQ vote but a worse one for the PQ vote.
  - ◆ In 2008, while support for ADQ is decreasing:
    - ◆ Overestimation of ADQ vote (17.3% vs 14.7%). Final vote: 16.4%.
    - ◆ Underestimation of PQ vote (27.2% vs 31.4%). Final vote : 35.2%.
    - ◆ Deterioration of the estimate for the PQ only.



### 3. What is the impact of adjusting using recall of past vote?

France 2002 & 2007

- ◆ Past analyses (Durand et al., 2005; Durand, 2007, 2008) have shown that the French pollsters have underestimated the Le Pen vote in 2002 and over-estimated it in 2007.
- ◆ However, estimation of the vote for the 2nd round was good using adjustment by declared vote for the first round.
- ◆ The PEF in 2007 shows that adjusting using declared past vote in the first round in 2002...
  - ◆ Gives 1.7 points more to Le Pen (9.6% vs 7.9%).
  - ◆ Gives a better estimate of Le Pen's vote (at 10.4%).
  - ◆ In this case, voting intention was collected using a secret ballot;
  - ◆ This tends to validate the hypothesis that the under-representation of LePen's vote may be due to dissimulation.

### 3. What is the impact of adjusting using recall of past vote?

Canada 2004, 2006, 2008

- ◆ In 2004: adjusting using declared vote in 2000 leads to
  - ◆ Substantial overestimation of the Conservative vote, a party whose support was decreasing between the 2 elections (39.9% vs 34.2%) : Vote: 29.6%.
  - ◆ Underestimation of the vote for the Liberal Party of Canada (also decreasing but less substantially)(28.9% vs 33.4%). Vote : 36.8%.
- ◆ In 2006 & 2008, adjustment does not change anything.

# Conclusion

- ◆ Different processes?
  - ◆ France : quotas, bad distribution even on election day
  - ◆ Canada/Quebec : random sampling, deterioration w. Time
- ◆ Voting recall is less reliable...
  - ◆ When a right wing populist party is present.
  - ◆ With time, when there is movement in support for the different parties.
- ◆ Recall is less reliable among individuals who are less stable politically, prone to vote for a populist party.
- ◆ Adjusting using recall of past vote has varying consequences depending of the causes of bad recall: it may
  - ◆ Improve estimates (France 2002, Quebec 2007),
  - ◆ Deteriorate estimates (Quebec 2008, Canada 2004)
  - ◆ Have no effect (Canada 2006, 2008).

# Conclusion

- ◆ Measuring recall allows to assess whether
  - ◆ There is a likely bias in the sample or
  - ◆ There is a “spiral of silence” type of process happening.
- ◆ We should be very cautious when...
  - ◆ There is evolution in support for a right wing populist party.
  - ◆ This is more so since the clientele of the populist parties has similar characteristics to those who tend not to accurately recall their previous voting behavior.