

Devons-nous pondérer par le vote à l'élection précédente?

Présenté au Colloque francophone sur les sondages,
Tenu à Tanger, 22-25 mars, 2010.

Par Claire Durand,
Isabelle Valois et Mélanie Deslauriers,
département de sociologie,
Université de Montréal

Un peu d'histoire

La pondération par le vote déclaré à l'élection précédente

- ◆ Est apparue en 1948 (Hellevik, 2009) et est couramment utilisée surtout en Europe, dans les pays/firmes qui utilisent la méthode des quotas.
- ◆ Visait au départ à compenser pour la sous-représentation des partis d'extrême-gauche (Hellevik, 2009)
- ◆ A visé plus récemment à compenser pour la sous-représentation des partis d'extrême-droite.
- ◆ A été recommandée (Market Research Society, GB 1992) pour atténuer les effets des variations dans la composition socio-politique des échantillons.

Deux questions...

- ◆ 1. Le vote déclaré à une élection précédente est-il fiable?
 - ◆ Quels sont les déterminants de la fiabilité du rappel?
- ◆ 2. Quelles sont les conséquences d'utiliser une telle mesure sur la validité de l'estimation du vote par l'intention de vote?

1. Le vote déclaré à l'élection précédente est-il fiable?

- ◆ Selon les auteurs qui ont étudié cette question, la distribution du rappel de vote est souvent biaisée:
 - ◆ Elle tendrait à sous-estimer les partis de l'extrême-droite.
 - ◆ Elle tendrait à sous-estimer les petits partis.
 - ◆ Elle tendrait à être biaisée en faveur du gagnant ou en faveur du parti pour lequel le répondant a désormais l'intention de voter.
 - ◆ Elle varierait avec le temps écoulé depuis la dernière élection.

Les déterminants de la non-fiabilité du rappel de vote

- ◆ La mémoire : involontairement, les répondants se trompent dans leur réponse au rappel de vote.
- ◆ La dissimulation : les répondants dissimulent l'information soit en ne répondant pas à la question sur l'intention de vote, soit en donnant une information fausse : “A people of liars”, explication répandue.
- ◆ L'échantillon est biaisé de sorte que l'on rejoint différemment les répondants selon leur vote, soit dû à la procédure d'échantillonnage, soit dû à la non collaboration de certains répondants.

Aux causes des conséquences

- ◆ Les problèmes de mémoire entraînent que la distribution du rappel de vote variera avec le temps.
- ◆ La dissimulation entraîne la mauvaise représentation de certains groupes, sans égard au temps.
- ◆ Le biais échantillonnal entraîne aussi la mauvaise représentation de certains groupes, sans égard au temps.

L'utilisation du rappel de vote pour redresser l'échantillon

Des conséquences différenciées?

- ◆ Dépendent des raisons de la non-fiabilité:
 - ◆ En l'absence d'évolution des intentions de vote, si la mauvaise distribution est due à la non réponse aléatoire, une possible amélioration de l'estimation.
 - ◆ Lorsqu'il y a évolution des intentions de vote et sous-représentation non aléatoire de certains groupes, l'utilisation des informations sur le rappel peut avoir des conséquences néfastes (Waldahl et Aardal, 1982):
 - ◆ Sous-estimation du vote dont l'intention est en hausse.
 - ◆ Sur-estimation du vote dont l'intention est en baisse.
 - ◆ C'est ce qui s'est passé en France en 2002 (sous-estimation de Le Pen), en 2007 (surestimation de Le Pen) et... aux régionales de 2010.

Objectif de la recherche

- ◆ Revisiter la question de la pondération par le rappel de vote, question qui n'a pas été étudiée depuis très longtemps (sauf Hellevik, 2009).
 - ◆ Le rappel de vote est-il biaisé? Toujours?
 - ◆ Les prédicteurs de mauvais rappel sont-ils encore les mêmes?
 - ◆ Comment prévoir si et dans quels contextes l'utilisation de cette pondération entraînera des biais? Comment les compenser?

Méthodologie

Trois cas : cas no 1

- ◆ Québec: élections de mars 2007 et décembre 2008
 - ◆ Mauvaise prédiction de l'élection de mars 2007.
 - ◆ Présence d'un parti populiste de droite.
 - ◆ Forte évolution des intentions de vote pour le parti populiste de droite (ADQ) : en hausse avant l'élection de 2007, en baisse après l'élection.
- ◆ Informations:
 - ◆ Double rappel de vote 2007 et 2008 (mesure de mauvais rappel).
 - ◆ Évolution du rappel de vote (2007-2010) et des intentions de vote.
 - ◆ Impact de la pondération par le vote à l'élection précédente

Méthodologie

Trois cas : Cas no 2

- ◆ Élections présidentielles françaises d'avril 2002 et d'avril 2007
 - ◆ Présence d'un parti d'extrême droite.
 - ◆ Sous-estimation de l'intention de vote pour l'extrême-droite en 2002 et sur-estimation du même vote en 2007 (Durand et coll. 2004, 2007, 2008).
 - ◆ Evolution des intentions de vote pour l'extrême-droite en hausse avant 2002, en baisse après 2002.
 - ◆ Analyses présentées dans POQ pour 2002 et IJPOR pour 2007.
 - ◆ Rappels de vote après le premier tour de la présidentielle, diverses firmes de sondage.
 - ◆ Panel électoral français de 2002 :
 - ◆ Rappel de vote après le premier tour.
 - ◆ Panel électoral français de 2007 :
 - ◆ Rappel de vote pour les deux tours.
 - ◆ Impact de la pondération par le rappel de vote sur la qualité de l'estimation du vote.

Méthodologie

Trois cas : le troisième

◆ Canada: élections de juin 2004, janvier 2006 et septembre 2008

◆ Informations:

- ◆ Estimation de la qualité du rappel de vote 2000, 2004, 2006, 2008.
- ◆ Double rappel de vote pour 2004 et 2006 (mesure de mauvais rappel).
- ◆ Impact de la pondération par le vote à l'élection précédente pour 2004, 2006 et 2008.

1. Quelle est la fiabilité du rappel de vote pour ces élections

- ◆ Pour le Québec comme pour la France, avec présence d'un parti populiste de droite ou d'extrême droite, le vote pour ce parti est sous-représenté, que le parti soit en ascension ou en descente
 - ◆ MAIS pour le Québec, il n'est pas sous-représenté au lendemain de l'élection.
 - ◆ Alors qu'en France, il est toujours sous-représenté.
- ◆ Pour le Canada, pour 2004, sous-représentation du Parti conservateur en 2006 mais bonne fiabilité pour l'élection de 2006 (déclaration en 2006 et en 2008).

Rappel de vote

Québec 2003 et 2007

	Québec 2003			Quebec 2007			
	Vote	Vote déclaré en 2007 (CROP>	Vote déclar> ds CES04	Vote	Vote déclaré après élec. 200>	Vote déclar> Crop - nov 0>	Vote déclaré ss-echant 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%
Autres	2,6%	4,4%	2,2%	7,7%	8,5%	8,2%	7,3%
Total	100%	100%	100%	100%	100%		
Vote blanc	1,25%	0%		1%	0%	0	0,1%
Non vote (>	29,58%	17%	12,6%	28,77%	18,8%	15,3%	13,4%

Le vote déclaré n'est bon que juste après l'élection (2007). Ensuite le rappel de vote se détériore, particulièrement surtout pour l'ADQ (droite populiste).

Rappel de vote

France 2002 et 2007

	France 2002			France 2007 (PEF)		
	Vote	Vote déclaré en 2002 (pef)	Vote déclaré en 2007(pef)		Vote	Vote déclaré après 1er tour
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 Besancenot	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 CHEVENE	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%			
Vote blanc	2,4%	1,7%	3,8%		1,2%	
Non vote	28,4%	14,1%	9,6%		16,2%	

Pour 2002 et pour 2007, la reconstitution du vote Le Pen est mauvaise dès le lendemain de l'élection.

Rappel de vote

Canada 2000 et 2004

	Canada 2000		Canada 2004			
	Vote	Vote déclaré en 2004	Vote	Vote déclaré en 2004	Vote déclaré > 2006 (pré rdd >	Vote déclaré en 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	10,7%	11,5%	12,4%	11,5%	11,2%	11,4%
Autres	2,2%	1,3%	5,6%	5,2%	3,7%	3,7%
Total	100%	100%	100%	100%	100%	100,1%
Non vote	39,1%	21,6%	39,1%	15%	17,6%	12,3%

Le vote déclaré n'est bon que juste après l'élection. Ensuite le rappel de vote se détériore, particulièrement pour les Conservateurs (sous-estimation) mais aussi pour les Libéraux (surestimation)

Rappel de vote

Canada 2006 et 2008

	Canada 2006		Canada 2008		Canada 2008	
	Vote	Vote déclaré en 2006	Vote déclaré en 2008 (pré>	Vote en 2008	Vote déclaré en 2008	Vote déclaré en 2008
PLC	30,2%	28,1%	30,5%	26,3%	24,8%	24,8%
PC (PPC+allian>	36,3%	38%	37,8%	37,7%	37,5%	37,5%
NPD	17,5%	17,9%	14,1%	18,2%	18,1%	18,1%
Bloc	10,5%	10,2%	12,3%	10%	12,1%	12,1%
Autres	5,5%	5,7%	5,3%	7,8%	7,5%	7,5%
Total	100%	100%	100%	100%	100%	100%
Non vote	35,3%	10,3%	17,1%	41,2%	14,7%	14,7%

Le vote déclaré est bon y compris deux ans plus tard (vote de 2006 déclaré en 2008) pour tous les partis.

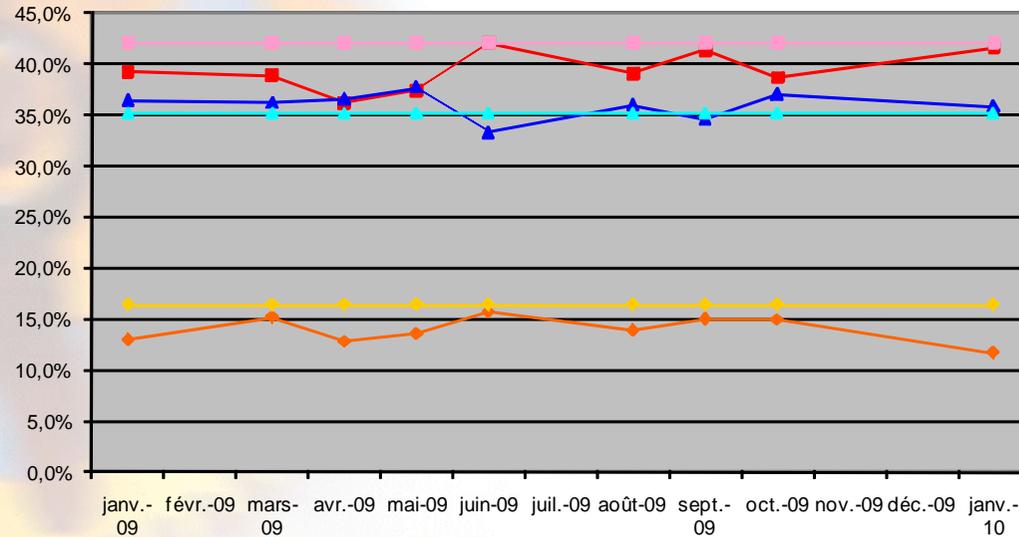
2. Le rappel de vote évolue-t-il avec le temps? Selon les partis? 2007-2008

1000%

Le rappel suit les intentions de vote : baisse du vote déclaré de l'ADQ (droite populiste) suit la baisse de l'intention de vote. Stabilité pour le PQ.

2. Le rappel de vote évolue-t-il avec le temps? Selon les partis? 2008-2010

Vote précédent déclaré pour l'élection prov. de décembre 2008 pondéré par recensement
excluant les NFP (non partis)



	janv.-09	mars-09	avr.-09	mai-09	juin-09	août-09	sept.-09	oct.-09	janv.-10
ADQ	13,0%	15,2%	12,8%	13,6%	15,7%	13,9%	15,0%	14,9%	11,7%
ADQ - résultats réels décembre 08	16,4%	16,4%	16,4%	16,4%	16,4%	16,4%	16,4%	16,4%	16,4%
Parti Libéral	39,3%	38,9%	36,2%	37,4%	42,1%	39,1%	41,4%	38,7%	41,6%
PLQ - résultats réels décembre 08	42,1%	42,1%	42,1%	42,1%	42,1%	42,1%	42,1%	42,1%	42,1%
Parti Québécois	36,4%	36,2%	36,6%	37,7%	33,3%	36,0%	34,6%	37,0%	35,8%
PQ - résultats réels décembre 08	35,2%	35,2%	35,2%	35,2%	35,2%	35,2%	35,2%	35,2%	35,2%

Le rappel suit les intentions de vote : aucune évolution de l'intention de vote, aucune évolution du vote déclaré.

Plus précisément,

Québec entre 2007 et 2008

- ◆ Régressions logistiques (18 sondages entre juin 2007 et novembre 2008)
- ◆ Toutes choses égales par ailleurs, soit après contrôle pour les déterminants socio-démographiques de l'intention de vote, le rappel varie significativement
 - ◆ Avec le temps (rappel de vote pour l'ADQ diminue).
 - ◆ En fonction de la variation dans les intentions de vote (rappel de vote pour les petits partis diminue en fonction de la hausse des intentions de vote pour le PQ).

France 2002 et 2007

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
	%	%	%	%	%	%
Extreme-left	10.5	8.7	9.5	9.5	13.7	7.4
Traditional Left						
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
Traditional Right						
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
Extreme-right	19.2	9.2	10.8	11.1	7.8	8.0
<i>Proportion who report their vote</i>				71.2	78.8	

Toutes les firmes sous-estiment le vote Le Pen d'environ la moitié, même le jour de l'élection (Sofres)

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
Sarkozy	31.2	28.3, 34.1	31.1	33.9	31.1	29.8	29.4	32.0	31.1
Royal	25.9	23.2, 28.6	31.5*	25.4	28.2	29.5*	28.9*	27.6	25.5
Bayrou	18.6	16.2, 21.0	20.1	19.9	20.4	21.0	22.3*	22.3*	23.3*
Le Pen	10.4	8.5, 12.3	3.3*	7.1*	5.9*	4.5*	5.1*	4.2*	5.6*
Total right	63.6	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.

Le vote déclaré pour LePen est fortement sous-représenté entre les deux tours et après le second tour.

La qualité du rappel varie-t-elle en fonction de caractéristiques individuelles?

Prédicteurs de la qualité du rappel individuel

- ◆ Mauvais rappel Québec 2007-2008, CES 2004-2006 et 2006-2008 ont les mêmes prédicteurs:
 - ◆ (age et scolarité : effet indirect via intérêt pour la campagne) (-) sauf pour CES 2006-2008
 - ◆ Intérêt pour la campagne de départ (-)
 - ◆ Identification à un parti politique et force de cette identification (-) (CES seulement)
 - ◆ Stabilité des opinions politiques (-) (absent CES 2006-2008)
 - ◆ Intention de vote:
 - ◆ Pour ADQ (droite populiste) pour Québec 2007 (vs autres partis)
 - ◆ Pour petits partis vs partis traditionnels (CES)

1. Le rappel de vote est-il fiable?

- ◆ Il est possible que des processus différents soient à l'oeuvre en France (sondages avec quotas) et au Canada (aléatoire).
- ◆ En France, mauvaise distribution même juste après l'élection.
- ◆ Au Canada, bonne distribution après l'élection mais parfois détérioration avec le temps.
- ◆ De plus, diverses caractéristiques socio-politiques -- intérêt pour la campagne, intention de vote, identification partisane -- sont liées à la non-fiabilité du rappel.

2. Quel est l'impact de la pondération utilisant le rappel de vote?

Québec 2007 et 2008

- ◆ La pondération par le vote déclaré à l'élection précédente...
 - ◆ En 2007, alors que l'ADQ est en hausse:
 - ◆ Surestime l'ADQ (28,6% p/r à 25,9%). Vote final: 30,8%.
 - ◆ Sousstime le PQ (24,3% p/r 28,3%). Vote final: 28,3%.
 - ◆ Améliore prédiction pour l'ADQ, détériore pour le PQ.
 - ◆ Fait passer l'ADQ devant le PQ, ce qui sera le résultat final.
 - ◆ En 2008, alors que l'ADQ est en baisse:
 - ◆ Surestime l'ADQ (17,3% p/r à 14,7%). Vote final: 16,4%.
 - ◆ Sousstime le PQ (27,2% p/r 31,4%). Vote final : 35,2%.
 - ◆ Détérioré la prédiction pour le PQ seulement.

Quel est l'impact de la pondération par le vote à l'élection précédente?

France 2002 et 2009

- ◆ Les analyses passées (Durand et coll., 2005, Durand, 2007, 2008) ont montré que les sondeurs français ont généralement sous-estimé Le Pen en 2002 et l'ont surestimé en 2007. Toutefois, la pondération par le vote déclaré au premier tour donne de bons résultats pour l'estimation du 2ème tour.
- ◆ Le PEF 2007 montre que la pondération par le vote déclaré au premier tour de 2002
 - ◆ Donne 1,7 points de pourcentage de plus à Le Pen (9,6% p/r à 7,9%).
 - ◆ Améliore la prédiction de Le Pen (à 10,4).
 - ◆ Dans ce cas, l'intention de vote était donnée dans une urne, ce qui tendrait à valider l'hypothèse d'une dissimulation du vote Le Pen dans les sondages habituels.

Quel est l'impact de la pondération par le vote à l'élection précédente?

Canada 2004, 2006, 2008

- ◆ En 2004: La pondération par le vote déclaré de 2000
 - ◆ Surestime fortement le parti Conservateur qui est en baisse (39,9% p/r à 34,2%) : Vote: 29,6%.
 - ◆ Sousstime le Parti Libéral du Canada (aussi en baisse mais moins fortement)(28,9% p/r à 33,4%).
Vote : 36,8%.
- ◆ En 2006 et en 2008, ne change à peu près rien.

Conclusions

- ◆ Le rappel de vote est moins fiable
 - ◆ Quand il y a présence d'un parti populiste de droite.
 - ◆ Avec le temps, quand il y a évolution des intentions de vote.
- ◆ Le rappel de vote est moins fiable chez les individus les moins stables politiquement, attirés par le vote populiste.
- ◆ L'impact de la pondération par le rappel de vote varie selon les causes de non-fiabilité: parfois amélioration (PEF2002, Québec 2007), parfois détérioration (Québec 2008, Canada 2004), parfois ne change rien au presque.

Conclusions

- ◆ Il faut être particulièrement prudent...
 - ◆ Lorsqu'un parti populiste de droite est en évolution,
 - ◆ D'autant plus que la clientèle de ces partis a des caractéristiques similaires à celles des personnes qui se souviennent "mal" de leurs votes précédents.