

CALAGE DES SONDAGES ÉLECTORAUX EN CAS DE SOUVENIR OU D'INTENTION DE VOTE DISSIMULÉS OU ANACHRONIQUES

Pierre Marie Windal

Conseil en marketing et expert auprès de la Commission des sondages¹

¹ 25 rue Narcisse Aussenard, 95450 Sagy, France
pierre@windal-conseil.com

Résumé. Pour améliorer la précision des estimations des intentions de vote recueillies par sondage, la plupart des instituts calent l'échantillon sur une ou plusieurs élections de référence en fondant cette pratique sur l'hypothèse implicite que le souvenir de vote est sincère. Cette note a pour objet d'analyser les biais que produit ce calage lorsque cette hypothèse n'est plus totalement respectée, volontairement, en cas de dissimulation, ou involontairement, en cas d'oubli ou d'erreur. Pour chacun des quatre cas traités – dissimulation du souvenir de vote chez un autre candidat, dissimulation du souvenir de vote chez les non réponses, dissimulation du souvenir et de l'intention de vote, souvenir de vote anachronique – le biais est explicité et un nouvel estimateur est proposé. Cette mesure du biais est ensuite illustrée par trois exemples tirés des élections présidentielles et régionales en France.

Mots-clés. Calage, souvenir de vote, dissimulation, biais, lissage, redressement.

1 Introduction

Cette note a pour objet d'analyser les biais que produisent un souvenir de vote ou une intention de vote dissimulés, ou un souvenir de vote anachronique sur les intentions de vote issues de sondages électoraux. On se souviendra, par exemple, du 1^{er} tour de l'élection présidentielle de 2002, en France, où, pour des raisons expliquées en particulier par Bachelet (2007), aucun institut de sondage n'est parvenu à prédire la présence de J.M. Le Pen au 2nd tour de cette élection. On pensera aussi, en Grande-Bretagne, au *Shy Tory Factor* (syndrome des conservateurs timides – E. Noelle-Neumann 1993), l'une des raisons avancées par les analystes politiques pour expliquer la réélection récente et imprévue de David Cameron (SIG^{lab} 2015).

- La dissimulation du souvenir de vote consiste à dire que l'on a voté pour le candidat X alors que l'on a, en réalité, voté pour le candidat Y. Elle affecte la matrice de transfert du souvenir vers les intentions de vote et produit une matrice biaisée. Certains électeurs du FN ont, par exemple, surtout dans le passé, été soupçonnés de dissimuler leur vote.
- La dissimulation conjointe du souvenir et de l'intention de vote procède du même mécanisme et produit elle aussi une matrice de transfert biaisée, mais en sens inverse du cas précédent.
- Le souvenir de vote anachronique, volontaire ou non, consiste à faire comme si l'on avait voté pour X alors que l'on a, en réalité, voté pour Y. Elle n'affecte pas la matrice de transfert – ces électeurs de X se comportent désormais comme des électeurs de Y – mais elle conduit à utiliser une population de référence erronée. À La Réunion, par exemple, à l'occasion des régionales 2015, certains électeurs de P. Vergès et de M. Vergoz ont manifestement pris le parti de D. Robert, le gagnant de 2010 et fait comme s'ils avaient voté pour lui.

¹ Les opinions et propositions de cette communication sont de la seule responsabilité de l'auteur et n'engagent en rien la Commission des sondages. Les chiffres cités sont extraits des médias, des sites des instituts et des communications personnelles de l'auteur.

Nous étudions ci-dessous les conséquences que ces phénomènes – dissimulation ou anachronisme – entraînent sur les estimations des intentions, suggérons des pistes pour minimiser le biais et concluons par une invitation à constituer une base de « normes » comme l’une des pistes qui permettrait de distinguer une simple erreur d’échantillonnage d’un phénomène de dissimulation ou de rétroaction du souvenir.

1 Le biais de dissimulation du souvenir de vote

Deux cas de dissimulation du souvenir sont étudiés, selon que cette dissimulation se niche chez les Y (les autres candidats) ou chez les non-réponses. Dans ce dernier cas, une fraction des électeurs de X déclarent à tort ne plus se souvenir de leur vote pour ne pas avoir à le communiquer. Dans les deux cas, on suppose que l’intention, elle, n’est pas affectée par cette dissimulation du souvenir. À titre d’exemple, il est aujourd’hui, en France, plus socialement acceptable d’avouer un vote frontiste que cela ne l’était en 2002. Si l’on hésite encore à dire que l’on a dans le passé voté pour ce parti (souvenir de vote), on hésite beaucoup moins à dire que l’on votera aujourd’hui pour lui (intention de vote).

1.1 Dissimulation chez les autres candidats

Pour simplifier l’analyse, considérons le candidat X (par exemple, le FN) et une catégorie Y regroupant les autres candidats. Les données du problème sont les suivantes :

Souvenir de vote			
	Candidat X	Autres candidats : Y	Commentaires
Poids dans la population	N_1	$N_2 = 1 - N_1$	
Poids dans l’échantillon	n_1	$n_2 = 1 - n_1$	$n_1 < N_1$
Taux de dissimulation	d_1	$-d_1$	$d_1 \geq 0$

Souvenir de vote vers intention de vote					
Vers ↓ / De →		Réel		Observé	
		X	Y	X	Y
X		v_{11}	v_{21}	v_{11}	$v_{21} + b$
Y		v_{12}	v_{22}	v_{12}	$v_{22} - b$

On suppose que $d_1\%$ des électeurs ont voté X et déclaré Y comme souvenir. On suppose également qu’ils se comportent comme les électeurs de X qui n’ont pas dissimulé leur vote. Ils sont donc v_{11} à voter pour X et non v_{21} comme les véritables électeurs de Y.

La quantité « b » dénote le biais que produit cette dissimulation sur la matrice de transfert. Ce que l’on observe, ce n’est donc pas v_{21} , mais :

$$v_{21}^b = v_{21} \left(1 - \frac{d_1}{n_2}\right) + v_{11} \frac{d_1}{n_2}, \text{ soit un biais } b = \frac{d_1}{n_2} (v_{11} - v_{21})$$

Soit $\text{Int}(X|R)$ l’intention de vote à l’égard de X si l’on redresse² l’échantillon sur le souvenir de vote, $\text{Int}(X|NR)$ l’intention si l’on ne redresse pas et $\text{Int}(X|vrai)$ l’intention si l’on redressait cet échantillon avec connaissance du taux de dissimulation d_1 .

² Dans cette note, le seul redressement auquel l’on fait référence est celui basé sur le souvenir du vote. Si le terme « brut » est employé, il englobe le redressement sociodémographique.

$$\text{Int}(X|\text{vrai}) = v_{11}N_1 + v_{21}N_2$$

$$\text{Int}(X|R) = v_{11}N_1 + (v_{21} + b)N_2 \quad \text{soit un biais égal à } bN_2$$

$$\text{Int}(X|NR) = v_{11}n_1 + (v_{21} + b)n_2 \quad \text{soit un biais égal à } v_{11}(n_1 - N_1) + v_{21}(n_2 - N_2) + bn_2$$

Que l'on redresse ou non, l'estimation de l'intention sera biaisée.

- Si l'on redresse, le biais vient de ce que l'on utilise un taux de transfert erroné : $(v_{21}+b)$ au lieu de v_{21} .
- Si l'on ne redresse pas, le biais est double : biais d'échantillonnage (on utilise n_1 au lieu de N_1 et n_2 au lieu de N_2) ; biais de dissimulation (bn_2). La bonne nouvelle, c'est que ces deux biais sont de signes contraires. Il n'est pas certain que l'intention X non redressée soit moins précise que l'intention redressée.

On notera que le redressement surestime cette intention (biais positif) alors que l'absence de redressement la sous-estime (biais négatif si $(N_1 - n_1) > d_1$, vrai par hypothèse). En l'absence d'information sur le degré de dissimulation (d_1), on peut toujours identifier le degré qui équilibre les deux biais (en valeur absolue, puisqu'ils sont de sens contraire), c.-à-d. qui produit la même estimation de l'intention X que l'on redresse ou non. Ce taux de dissimulation neutre, dénoté d_1^{minmax} est la solution de l'équation :

$$(bN_2)^2 = [v_{11}(n_1 - N_1) + v_{21}(n_2 - N_2) + bn_2]^2, \text{ d'où il ressort que } d_1^{\text{minmax}} = \frac{(N_1 - n_1)n_2}{(N_2 + n_2)}$$

Avec ce degré de dissimulation, les deux biais sont égaux à :

$$\text{Biais}^{\text{minmax}} = \frac{N_2(N_1 - n_1)(v_{11} - v_{21})}{N_2 + n_2}$$

Il se trouve que ce biais neutre est le minimum du biais maximal. En d'autres termes, quel que soit le degré de dissimulation d_1 , le biais induit par le redressement corrigé du biais minmax (ou le non-redressement, puisque les estimations seront équivalentes) sera inférieur à $\text{Biais}^{\text{minmax}}$.

L'estimateur proposé est donc égal à :

$$\text{Int}(X|\text{biais minmax}) = \text{Int}(X|R) - \text{Biais}^{\text{minmax}}, \text{ en estimant le taux de transfert inconnu } v_{21} \text{ par la}$$

quantité : $\frac{v_{21}N_2 - v_{11}d_1^{\text{minmax}}}{N_2 - d_1^{\text{minmax}}}$

1.2 Dissimulation chez les non-réponses

Si une fraction des électeurs de X se dissimule dans les non-réponses, la matrice de transfert souvenir/intention $\{v_{ij}\}$ n'est pas affectée. Tout se passe comme si le poids de X dans l'échantillon n'était pas de n_1 , mais de $\frac{n_1 + d_1}{1 + d_1}$ et celui de Y , pas de n_2 mais de $\frac{n_2}{1 + d_1}$. Cette dissimulation conduit à utiliser un poids trop fort pour les électeurs de X alors que précédemment on utilisait le bon poids, mais un v_{21} trop élevé.

Le biais sera supérieur à celui de la dissimulation de X chez Y , car elle pèse sur le poids lui-même, pas seulement sur l'une des composantes de la matrice de transfert. À titre de comparaison, voici les formules équivalentes à celles du cas précédent.

$$\text{Int}(X|R) = \frac{v_{11}N_1(n_1+d_1)}{n_1(1+d_1)} + \frac{v_{21}N_2}{1+d_1}$$

$$\text{Int}(X|NR) = \frac{v_{11}(n_1+d_1)}{(1+d_1)} + \frac{v_{21}n_2}{1+d_1}$$

$$\text{Biais}(X|R) = \frac{d_1}{1+d_1} \left[\frac{v_{11}N_1n_2}{n_1} - v_{21}N_2 \right]$$

$$\text{Biais}(X|NR) = \frac{(v_{11}-v_{21})(d_1N_2+n_1-N_1)}{1+d_1}$$

$$d_1^{\text{minmax}} = \frac{n_1(N_1-n_1)(v_{11}-v_{21})}{v_{11}(N_1n_2+n_1N_2)-2n_1N_2v_{21}}$$

$$\text{Biais}^{\text{minmax}} = \frac{(N_1-n_1)(v_{11}-v_{21})(v_{11}N_1n_2 - v_{21}n_1N_2)}{v_{11}n_2(N_1+n_1)-v_{21}n_1(N_2+n_2)}$$

2 Le biais de dissimulation du souvenir et de l'intention de vote

2.1 Dissimulation chez les autres candidats

Contrairement aux deux cas précédents, l'électeur sondé dissimule *et* son souvenir *et* son intention de vote, ce que résume le schéma ci-dessous.

Souvenir de vote			
	Candidat X	Autres candidats : Y	Commentaires
Poids dans la population	N_1	$N_2 = 1 - N_1$	
Poids dans l'échantillon	n_1	$n_2 = 1 - n_1$	$n_1 < N_1$
Taux de dissimulation	d_1	$-d_1$	$d_1 \geq 0$

Souvenir de vote vers intention de vote					
		Observé		Réel	
Vers ↓ / De →		X	Y	X	Y
	X	v_{11}	v_{21}	v_{11}	$v_{21} - b$
	Y	v_{12}	v_{22}	v_{12}	$v_{22} + b$

On suppose que $d_1\%$ des électeurs ont voté pour le candidat X et déclaré Y comme souvenir et intention de vote. Cette dissimulation de l'intention se traduit par un transfert biaisé du souvenir Y vers l'intention X : on observe v_{21} alors que le transfert réel est égal à $v_{21} + b$.

Comme précédemment le biais b est égal à $\frac{d_1}{n_2}(v_{11} - v_{21})$, mais il agit cette fois en sens inverse. Les estimations des intentions sont données par :

$$\text{Int}(X|\text{vrai}) = v_{11}N_1 + v_{21}N_2$$

$$\text{Int}(X|R) = v_{11}N_1 + (v_{21}-b)N_2 \quad \text{soit un biais égal à } -bN_2$$

$$\text{Int}(X|NR) = v_{11}n_1 + (v_{21}+b)n_2 \quad \text{soit un biais égal à } v_{11}(n_1-N_1) + v_{21}(n_2-N_2) -bn_2$$

Sachant que $b > 0$, les intentions redressées à l'égard de X sous-estimeront le vote réel, et les intentions non redressées plus encore puisque des deux composantes (population et taux de transferts erronés) vont dans le même sens.

2.2 Dissimulation chez les non-réponses

Par hypothèse, ni la matrice de transfert (les v_{ij}) ni les poids des électeurs de X et de Y (les n_i) ne sont affectés, puisque les dissimulateurs se nichent dans les non-réponses. Tout se passe comme si le poids des électeurs de X dans la population n'était pas de N_1 , mais de $\frac{N_1+d_1}{1+d_1}$ et celui des électeurs de Y, de $\frac{N_2}{1+d_1}$.

		Souvenir de vote			
		Réel		Observé	
		X	Y	X	Y
Population		$\frac{N_1+d_1}{1+d_1}$	$\frac{N_2}{1+d_1}$	N_1	N_2
Échantillon		n_1	n_2	n_1	n_2

Les estimations des intentions sont données par :

$$\text{Int}(X|\text{vrai}) = v_{11} \frac{N_1+d_1}{1+d_1} + v_{21} \frac{N_2}{1+d_1}$$

$$\text{Int}(X|R) = v_{11}N_1 + v_{21}N_2$$

Le biais induit par le redressement s'élève à $\frac{d_1 N_2}{1+d_1} (v_{21} - v_{11})$.

3 Biais de souvenir de vote et lissage

Soit l'estimateur lissé $\text{Int}(X|\text{lissé}) = \alpha \text{Int}(X|R) + (1-\alpha)\text{Int}(X|NR)$ où $0 \leq \alpha \leq 1$ est un coefficient de lissage à déterminer. Lorsque les estimateurs redressés et non redressés encadrent le vote réel, existe-t-il un α tel que les deux estimateurs $\text{Int}(X|\text{lissé})$ et $\text{Int}(X|\text{biais minmax})$ soient égaux ?

La réponse est « oui », est la valeur de cet α est égale à $\frac{1}{2}$. Ainsi, l'estimateur le moins biaisé (biais maximal minimum) en l'absence d'information sur d_1 est tout simplement la moyenne des intentions redressées et non redressées.

4 Le biais de souvenir de vote anachronique

Les données du problème sont les suivantes :

		Souvenir de vote		
		Candidat X	Autres candidats : Y	Commentaires
Poids dans la population		N_1	$N_2 = 1 - N_1$	
Poids dans l'échantillon		n_1	$n_2 = 1 - n_1$	$n_1 > N_1$
Taux de dissimulation		d_1	$-d_1$	$d_1 \geq 0$

On note cette fois-ci que $n_1 > N_1$ et que la négation du vote n'altère pas la matrice de transfert du

souvenir vers l'intention. On suppose que $d_1\%$ des électeurs ont voté X et déclaré Y comme souvenir. On suppose également qu'ils se comportent désormais comme les électeurs de Y. Ils sont donc v_{21} à voter pour X et non v_{11} comme les électeurs fidèles. Les estimateurs deviennent :

$$\text{Int}(X|\text{vrai}) = v_{11}(N_1 + d_1) + v_{21}(N_2 - d_1)$$

$$\text{Int}(X|R) = v_{11}N_1 + v_{21}N_2 \text{ soit un biais égal à } d_1(v_{21} - v_{11})$$

$$\text{Int}(X|NR) = v_{11}n_1 + v_{21}n_2 \text{ soit un biais égal à } (n_1 - N_1 - d_1)(v_{11} - v_{21})$$

Tout se passe comme si le poids de X dans l'élection de référence n'était pas de N_1 , le résultat de l'élection, mais N_1 augmenté des apostats d_1 . Que l'on redresse ou non, l'estimation de l'intention à l'égard de X sera biaisée faute d'utiliser la bonne (mais inconnue) population de référence.

- Si l'on redresse, le biais vient de ce que l'on utilise N_1 au lieu de $(N_1 + d_1)$ et N_2 au lieu de $(N_2 - d_1)$.
- Si l'on ne redresse pas, le biais vient de ce que l'on utilise n_1 au lieu de $(N_1 + d_1)$ et n_2 au lieu de $(N_2 - d_1)$.

On notera que le redressement sous-estime cette intention (biais négatif) alors que l'absence de redressement la surestime (biais positif). Comme pour la dissimulation du souvenir de vote, l'estimateur lissé :

$$\alpha \text{Int}(X|R) + (1 - \alpha) \text{Int}(X|NR)$$

avec un coefficient α de $1/2$ correspond à un d_1^{minmax} de $\frac{n_1 - N_1}{2}$ et un biais $^{\text{minmax}}$ de $\frac{(v_{21} - v_{11})(n_1 - N_1)}{2}$

5 Exemples d'application

5.1 Élection présidentielle de 2002 en France

Le premier tour de cette élection est un cas exemplaire de dissimulation du souvenir et de l'intention de vote. Nous reprenons de Bachelet (2007) les données de base servant à illustrer le calcul a posteriori du biais induit par cette dissimulation.

Rappelons que les instituts pris au piège de cette dissimulation avaient prédit en moyenne un écart de 4 points en faveur de Lionel Jospin pour un écart réel de 0.7 point en faveur de J.M. Le Pen. C'est moins l'inversion du score en elle-même qui a interpellé les commentateurs, le faible écart entre les deux candidats s'inscrivant dans la marge d'erreur associée à des échantillons utiles de l'ordre de 700 personnes, que le fait que tous les instituts se sont trompés, un évènement aléatoire hautement improbable³.

Avec l'assurance que donne la connaissance du passé, on sait ce qu'il aurait fallu faire en dehors du simple fait de publier les résultats du sondage plutôt que les résultats filtrés à l'aune de ce que les instituts pensaient qu'ils devaient être.

³ Ce phénomène n'est pas propre aux sondeurs français. Lors des élections américaines de mi-mandat de 2014, Nate Silver (2014) avait montré que dans la dernière phase de la campagne, les instituts dont les chiffres étaient jusqu'ici divergents avaient tendance à donner des résultats de plus en plus proches. En Grande-Bretagne, au lendemain du scrutin, Damian Lyons Lowe, directeur de l'institut Survation, a révélé que le dernier sondage qu'il a réalisé donnant les conservateurs à 37 % et les travaillistes à 30 %, l'étude n'a pas été publiée, car l'institut en a jugé les résultats trop éloignés de ceux des autres sondages publiés (SIG^{lab}, 2015)

	Le Pen	Autres candidats
Élection de 1995		
Vote réel	15,3 %	84,7 %
Souvenir de vote	7,5 %	92,5 %
Élection de 2002		
Vote réel	16,9 %	83,1 %
Intention de vote non redressée	9,0 %	91,0 %
Intention de vote redressée	14,0 %	86,0 %
Écart intention/vote	-2,9 %	

Transfert souvenir/intention de vote

	Le Pen	Autres candidats
Le Pen	63 %	5 %
Autres candidats	37 %	95 %

Sans connaître le taux de dissimulation (d_1), on peut se demander quel taux aurait produit une sous-estimation de l'écart constaté entre les intentions de vote et le vote réel (-2,9 %). Si Δ représente cet écart, les formules du §2.1 nous disent que d_1 doit être égal à $\frac{\Delta n_2}{N_2(v_{21}-v_{11})}$ soit 5,5 %. Ainsi, l'écart de 7,8 % observé entre le vote réel de 1995 à l'endroit de J.M. Le Pen (15,3 %) et le souvenir déclaré de ce vote (7,5 %) se répartit-il en 5,5 % pour la dissimulation et 2,3 % pour l'aléa. Ce dernier chiffre correspond en ordre de grandeur à ce que l'on observe aujourd'hui dans les sondages pour Marine Le Pen, la parole s'étant, en quelque sorte, libérée.

Alternativement, si les dissimulateurs se nichent dans les non-réponses, le taux de dissimulation d_1 qui produit le Δ recherché (cf. §2.2) est égal à $\frac{\Delta}{N_2(v_{21}-v_{11})-\Delta}$, soit un un taux de 6.3%. À titre comparatif, Bachelet (2007) avait estimé à 75% le pourcentage de dissimulation des électeurs de J.M. Le Pen en 2002, quantité proche de ce que l'on calcule ici : $6.3\% \div (15.3\% - 7.5\%) = 81\%$.

5.2 Le cas des régionales 2015 à la Réunion

À La Réunion, le fait que le seul institut à ne pas redresser politiquement son échantillon en 2010 ait fourni les prévisions les plus précises laissait craindre l'existence de souvenir de vote anachronique, une fraction de l'électorat du perdant de l'élection antérieure rejoignant en souvenir de vote le camp du vainqueur.

	D. Robert	Autres candidats
Régionales de 2010		
Vote réel	32,1 %	67,9 %
Souvenir de vote	46,5 %	53,5 %
Régionales de 2015		
Vote réel	40,4 %	59,6 %
Intention de vote non redressée	40,2 %	59,8 %
Intention de vote redressée	35,1 %	64,9 %
Écart intention/vote	-5,3 %	

À l'occasion d'un sondage on constate, par exemple, que 46,5 % des sondés déclarent avoir voté pour D. Roberts aux municipales de 2010 pour un vote réel de 32,1 %. Appliquant les formules du §4, il s'avère qu'il faut un taux de souvenir anachronique de 12 % pour expliquer l'écart constaté :

$d_1 = \frac{\Delta}{v_{21} - v_{11}}$ avec $\Delta = -5,3 \%$, $v_{11} = 57\%$ et $v_{21} = 12\%$, en se basant sur les transferts suivants :

Transfert souvenir/intention de vote		
	D. Robert	Autres candidats
D. Robert	57 %	12 %
Autres candidats	43 %	88 %

La presque totalité de l'écart constaté entre le vote réel et le souvenir de vote ($46,5\% - 32,1\% = 14,4\%$) correspond à du souvenir anachronique avec, pour conséquence immédiate, que les intentions non redressées sont bien meilleures que leurs homologues redressées, en l'occurrence pratiquement égales au vote réel, conformément à la théorie :

Biais de non-redressement = $(n_1 - N_1 - d_1)(v_{11} - v_{21})$, ce qui donne $(46,5\% - 32,1\% - 12\%)(57\% - 12\%) = 1,1\%$, soit un biais quasi nul.

5.3 Estimer le potentiel électoral de Marine Le Pen

Un exemple de dissimulation potentielle du souvenir de vote nous est fourni par un sondage d'octobre 2015, consacré au potentiel électoral de Marine Le Pen. Ce type de sondage se prête à la dissimulation, car il se limite au seul FN et a été réalisé par téléphone. Dans ce sondage, on constate un écart de près de 6.6 points entre le souvenir de vote en faveur de M. Le Pen à la présidentielle de 2012 et le vote réel, écart qui semble élevé par rapport à ceux habituellement constatés – de l'ordre de 2 points.

Assimilons pour l'exemple la différence entre cette « norme » et l'écart constaté, soit 4.6, à de la dissimulation (X chez Y) et posons les transferts (plausibles) suivants entre le souvenir et l'intention de vote :

$v_{11} = .89$, $v_{21} = .12$, ce qui donne avec les formules du §1.1 :

$$v_{21}^b = .12(1 - .046 \div .883) + .89 \times .046 \div .883 = .16$$

Les intentions de vote à l'égard du FN sont :

$$\text{Int}(\text{FN}|\text{vrai}) = (.89)(.183) + (.12)(.817) = 26.1 \%$$

$$\text{Int}(\text{FN}|\text{R}) = (.89)(.183) + (.16)(.817) = 29.4 \%, \text{ soit un biais de } + 3.3 \text{ points.}$$

$$\text{Int}(\text{X}|\text{NR}) = (.89)(.117) + (.16)(.883) = 24.5 \%, \text{ soit un biais de } -1.6 \text{ point}$$

associées à un biais minmax de : $[(.817)(.183 - .117)(.89 - .12)] \div (.817 + .883) = + 2.44$ points, que l'on utilise pour corriger l'intention :

$$\text{Int}(\text{FN}|\text{R corrigé biais minmax}) = 29.4 \% - 2.44 \% = 27 \%$$

$$\text{Int}(\text{FN} | \text{lissé avec } \alpha = 1/2) = (29.4 \% + 24.5 \%) \div 2 = 27 \%$$

Dans le cas présent, si l'on accorde suffisamment de crédit à la « norme » de 2 %, il eût été préférable de ne pas redresser les résultats. Le biais négatif de 1.6 point est inférieur au biais minmax (+2.4 points), lequel est inférieur au biais induit par le redressement (+ 3.3 points).

Alternativement, l'institut aurait pu utiliser l'estimation du taux de dissimulation ($d_1 = 4.6$) pour éliminer le biais du taux de transfert – substituer v_{21} à v_{21}^b – ce qui reviendrait à estimer l'intention FN par $\text{Int}(\text{FN}|\text{vrai})$. On obtiendrait le même résultat avec l'estimateur lissé doté d'un α de 0.33.

6 Conclusion et voies de recherche

Il est difficile de prouver l'existence de dissimulation ou de rétroaction du souvenir de vote. Tout au plus peut-on, en cas d'écart jugé « anormal » entre le brut et le redressé, soupçonner que cet écart résulte de la somme de l'erreur d'échantillonnage et de la dissimulation ou de la rétroaction en gardant à l'esprit que d'autres causes peuvent produire les mêmes effets.

Le soupçon de dissimulation du vote doit être étayé par d'autres indices que la seule sous-estimation du souvenir de vote. Réintégrer le brut par lissage peut s'avérer très pénalisant en l'absence de dissimulation. Il faut que le contexte s'y prête, comme dans l'exemple cité en §5.3 (mode de recueil téléphonique, sujet sensible, sous-estimation beaucoup plus forte que d'habitude, etc.)

La matrice de transfert du souvenir vers l'intention de vote, impactée en cas de dissimulation, gagnerait à être rapprochée de matrices types, propres à chaque type d'élection, pour détecter d'éventuelles anomalies.

Ce n'est qu'en se dotant de « normes » issues d'élections antérieures que l'on peut espérer passer du simple diagnostic *ex post facto* à la prédiction.

Bibliographie

[1] Bachelet D. (2007), Proposition 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 du Marketing*, février 2007, n° 211 – 1/5.

[2] Noelle-Neumann E. (1993), *The Spiral of Silence : Public Opinion - Our Social Skin*, The University of Chicago Press.

[3] SIG^{lab} (2015), <http://www.siglab.fr/fr/pourquoi-les-sondages-nont-ils-pas-vu-venir-la-reelection-de-david-cameron>.

[4] Silver N. (2014), Here's Proof Some Pollsters Are Putting A Thumb On The Scale, <http://fivethirtyeight.com/features/heres-proof-some-pollsters-are-putting-a-thumb-on-the-scale/>