

2M11-2668.1

Université de Montréal

L'effet des facteurs méthodologiques sur la qualité
de l'estimation des intentions de vote

par

Sébastien Vachon

Département de sociologie, Université de Montréal

Faculté des arts et sciences

Mémoire présenté à la Faculté des études supérieures
en vue de l'obtention du grade de
Maître ès sciences (M.Sc.)
en sociologie

janvier, 1998

© Sébastien Vachon, 1998



HM

15

U574

1999

v.003



Library of the Faculty of Arts and Social Sciences
University of Toronto

1999

Department of Economics, University of Toronto

Faculty of Arts and Social Sciences

Library of the Faculty of Arts and Social Sciences
University of Toronto
127 St. George Street, Toronto, Ontario M5S 1A5
Canada

*

1999

1999

Université de Montréal
Faculté des études supérieures

Ce mémoire intitulé:

L'effet des facteurs méthodologiques sur la qualité
de l'estimation des intentions de vote

présenté par:

Sébastien Vachon

a été évalué par un jury composé des personnes suivantes:

Paul Bernard - président-rapporteur
Claire Durand - directrice de recherche
André Blais - codirecteur
Richard Nadeau - membre du jury

Mémoire accepté le: 14 janvier 1999

Sommaire

Les sondages constituent un mode d'enquête largement utilisé par les pouvoirs publics et para-publics, les médias, les banques, les entreprises privées, etc. Parmi tous les sondages produits, les plus visibles sont, sans nul doute, les sondages pré-électoraux. Dans cette perspective, cette recherche met en lumière l'effet de certaines pratiques méthodologiques, reliées à la réalisation des sondages, sur l'estimation des intentions de vote des résidents du Québec pendant la campagne fédérale de 1997. Après examen de la documentation sur le sujet, il apparaît que les pratiques méthodologiques moins strictes, c'est-à-dire qui risquent de faire augmenter la non-réponse, sont associées à une moins bonne estimation des intentions de vote.

Pour les fins de l'étude, 10 sondages et 7 strates de sondage, c'est-à-dire des portions de sondages pan-canadiens, ont été analysés. Afin de déterminer quelles pratiques méthodologiques sont reliées à une moins bonne estimation des intentions de vote, il est nécessaire d'avoir une base de comparaison, un modèle de l'évolution de l'intention de vote tout au long de la campagne. Pour ce faire, l'analyse de série chronologique ARIMA a été employée. Afin de mesurer les pratiques, de classer les sondages en fonction de la plus ou moins grande maximisation de l'échantillon, c'est-à-dire des efforts mis pour obtenir une entrevue auprès des personnes moins coopératives et plus difficiles à joindre, des analyses en classes hiérarchiques ont été réalisées. Enfin, l'impact de la maximisation de l'échantillon sur l'estimation des intentions de vote est examiné au moyen d'analyses de variance et de covariance.

Il apparaît qu'une faible maximisation de l'échantillon entraîne une moins bonne estimation

des intentions de vote pour le Parti libéral du Canada et pour le Bloc québécois et une sous-estimation des intentions de vote pour le Parti conservateur. Ces résultats s'expliquent, du moins en partie, par le peu d'effort qui a été mis pour réduire la non-réponse.

Les différentes pratiques méthodologiques ne s'équivalent pas, certaines permettent d'obtenir un échantillon plus représentatif de la population à l'étude et par le fait même, des résultats plus fiables. Au Canada, comme dans les autres pays industrialisés, on constate une absence de consensus sur tous les aspects de la méthodologie de sondage (Taylor, 1994, 1997; Durand, Vachon et Tanguay, 1998; Voss, Gelman et King, 1995), malgré que la recherche en méthodologie de sondage soit relativement bien développée. De nombreuses autres recherches devront être menées afin de mieux comprendre les conséquences liées à l'utilisation de certaines pratiques méthodologiques, notamment l'utilisation de quotas.

TABLE DES MATIÈRES

Page d'identification du jury	iv
Sommaire	v
Liste des tableaux	x
Listes des figures	xii
Remerciements	xiii
INTRODUCTION	1
1. PROBLÉMATIQUE	3
1.1 Pratiques méthodologiques	4
1.1.1 Devis de sondage	4
1.1.1.1 Mode de génération de l'échantillon	4
1.1.1.2 Mode de sélection dans le ménage	7
1.1.1.3 Utilisation de quotas	10
1.1.2 Conduite des opérations	12
1.1.2.1 Prise de rendez-vous	13
1.1.2.2 Récupération des refus	13
1.1.2.3 Nombre de tentatives de contact	14
1.1.2.4 Durée de la période de collecte	15
1.1.2.5 Substitution de numéros de téléphone	16
1.1.3 Redressement	17
1.2 Représentativité de l'échantillon	20
1.2.1 Représentativité des personnes moins coopératives	21
1.2.2 Représentativité des personnes plus difficiles à joindre	23
1.3 Qualité des estimés	26
2. Hypothèse	32

3. Méthodologie	34
3.1 Données	34
3.2 Mesures	35
3.2.1 Mesures de l'intention de vote	35
3.2.2 Mesures de la mauvaise estimation des intentions de vote	37
3.3 Classification des pratiques	38
3.4 Analyses des relations	39
4. ARTICLE: «Pratiques méthodologiques et estimation des intentions de vote»	42
4.1 Contexte	42
4.2 Méthodologie	50
4.2.1 Données	50
4.2.2 Modèle d'analyse	50
4.3 Les résultats	53
4.3.1 Mesures	53
4.3.1.1 Intentions de vote	53
4.3.1.2 Caractérisation des pratiques	57
4.3.2 Effets de la maximisation de l'échantillon sur l'estimation des intentions de vote	59
4.3.2.1 Erreurs relatives	59
4.3.2.1 Erreurs absolues	61
4.4 Conclusion	63
Références (pour l'article)	64
5 DISCUSSION	68

6 CONCLUSION	71
RÉFÉRENCES	73
ANNEXE 1	77
ANNEXE 2	79
ANNEXE 3	88

LISTE DES TABLEAUX

I	Redressement des résultats pour les hommes, avec données imputées et réelles	18
II	Profil du répondant coopératif et du répondant moins coopératif	22
III	Profil du répondant demeurant dans un ménage où il ya eu initialement refus (Chevalier et Durand, 1997)	23
IV	Profil du répondant plus difficile à joindre	25
V	Nombre d'appels faits pour compléter l'entrevue et intentions de vote (élection américaine de 1988)	29
VI	Nombre d'appels faits pour compléter l'entrevue et intentions de vote (élection américaine de 1984)	29
VII	Intentions de vote des Québécois à l'élection fédérale de juin 1997 en fonction de leur choix lors du référendum de 1995 sur la souveraineté	30
VIII	Analyse de série chronologique ARIMA pour le Bloc québécois	54
IX	Analyse de série chronologique ARIMA pour le Parti conservateur	56
X	Analyse de série chronologique ARIMA pour le Parti libéral du Canada	57

XI	Analyse de variance et de covariance mesurant l'effet de la maximisation de l'échantillon sur l'estimation des intentions de vote - erreurs relatives	60
XII	Analyse de variance et de covariance mesurant l'effet de la maximisation de l'échantillon sur l'estimation des intentions de vote - erreurs absolues	62

LISTE DES FIGURES

1	Modèle pour le Bloc québécois, projection, sondages et strates, maximisation de l'échantillon	54
2	Modèle pour le Parti conservateur, projection, sondages et strates, maximisation de l'échantillon	55
3	Modèle pour le Parti libéral du Canada, projection, sondages et strates, maximisation de l'échantillon	57
4	Analyses en classes hiérarchiques - catégorisation des pratiques méthodologiques	58

REMERCIEMENTS

Je tiens à remercier Claire Durand, ma directrice, pour m'avoir si bien encadré dans ce processus de recherche. Ces conseils, tant sur le plan de la forme que du contenu, ont été des plus précieux pour la rédaction de mon mémoire. Sa grande disponibilité, son enthousiasme et sa rigueur sont quelques unes des qualités que j'ai particulièrement appréciées. De plus, elle m'a donné la chance de participer à différents projets qui m'ont permis d'acquérir de précieuses connaissances et de développer une bonne expérience de recherche. Je tiens également à remercier mon codirecteur André Blais. Ces conseils m'ont été des plus utiles.

J'aimerais exprimer ma reconnaissance au Fonds pour la formation de chercheurs et l'Aide à la recherche (FCAR) pour m'avoir accordé une bourse de maîtrise en recherche. Grâce à cette bourse, j'ai pu me consacrer à plein temps à la réalisation de mon mémoire. Finalement, je tiens à remercier toutes les firmes de sondage qui ont accepté de participer à cette recherche. Leur collaboration était essentielle à l'aboutissement de ce projet.

Les sondages constituent un mode d'enquête largement utilisé par les pouvoirs publics et parapublics, les médias, les banques, les entreprises privées, etc. Qu'ils soient largement diffusés par les médias ou qu'ils demeurent dans l'ombre, ils affectent la vie des citoyens. Parmi tous les sondages produits, les plus visibles sont, sans nul doute, les sondages pré-électoraux. «Les sondages se multiplient et envahissent les salles de rédaction des médias. Les journalistes ont d'ailleurs amplifié le phénomène, à tel point que la nécessité des sondages en démocratie est de plus en plus contestée» (Lachapelle, 1991: 3). Au Canada, comme dans la majorité des démocraties occidentales, ils occupent une place centrale dans le déroulement des campagnes électorales. «The modern campaign for public office is now built around polling» (Cantril, 1991: 9). De fait, on peut constater que leur nombre ne cesse d'augmenter (Lachapelle, 1991). Ainsi, «il est essentiel de s'assurer que les données diffusées sont fiables quel que soit l'impact des sondages. Le droit du public à une information juste, principe fondamental dans les démocraties pluralistes, demeure la pierre d'assise de toute démonstration visant à bien juger à la fois de la validité des sondages et de leur valeur scientifique» (Lachapelle, 1991: 16).

La présente recherche a permis la rédaction d'un article qui, au début septembre, a été soumis aux Cahiers canadiens de sociologie (lettre de soumission à l'Annexe 1). Celui-ci a été rédigé en collaboration avec mes deux codirecteurs Claire Durand et André Blais. Ainsi, c'est sous la forme d'un mémoire par article que cette recherche est présentée.

Dans la première section, la problématique trace un portrait des différentes opérations nécessaires à la réalisation d'un sondage. Une description détaillée des différentes pratiques méthodologiques permet de mettre en lumière ce processus. Dans la deuxième, le modèle d'analyse est décrit de façon exhaustive: types d'analyses employées, variables sélectionnées, élaboration des

indicateurs, etc. Dans la troisième, l'article intitulé «Pratiques méthodologiques et estimation des intentions de vote» est présenté. Dans la quatrième section, les résultats et certains aspects de la méthodologie d'analyse sont discutés. Finalement, un résumé des principaux résultats est présenté et quelques suggestions de recherches sont proposées.

1 Problématique

Les sondages d'opinion occupent une place des plus enviables en tant que mode d'enquête. C'est ainsi que, aujourd'hui plus que jamais, la qualité et la fiabilité de la méthodologie de sondage prennent toute leur importance. De fait, dans la mesure où les sondages peuvent affecter la perception que les électeurs se font de la force ou de la faiblesse des différents partis et dans la mesure où cette perception peut affecter le vote (Johnson, Blais, Brady et Crête, 1992), il devient impératif d'estimer la fiabilité de ces sondages.

D'un point de vue théorique, pour s'appuyer sur la théorie des probabilités et pouvoir ainsi mesurer l'erreur d'échantillonnage, c'est-à-dire l'importance des biais occasionnés par la généralisation des résultats à la population étudiée, il est nécessaire que la génération de l'échantillon se fasse au hasard et que chaque membre de la population ait une chance connue d'être sélectionné. Si, en théorie, la meilleure façon de procéder à cette sélection semble relativement claire, dans la pratique, les choses ne sont pas si simples. En effet, il n'y a pas une façon standardisée et codifiée, mais une pléiade de pratiques méthodologiques différentes. Les sondages, bien que scientifiques, sont aussi des productions sociales. «Les techniques utilisées pour produire et pour construire les données sont, comme les données elles-mêmes, des construits sociaux: les techniques statistiques se développent et se transforment continuellement sous l'influence de conditions historiques particulières» (Irvine, Miles et Evans, 1979 dans Blondiaux, 1991: 756). De fait, au Canada comme à l'étranger, les techniques statistiques et les pratiques méthodologiques inhérentes à la réalisation des sondages sont loin d'être standardisées dans l'ensemble des firmes (Taylor, 1994, 1997; Voss, Gelman et King, 1995; Durand, Vachon et Tanguay, 1998). Taylor (1997), dans un article intitulé «The very

different methods used to conduct telephone surveys of the public», fait la revue des méthodes employées par 83 firmes de sondage répartis dans 17 pays. Sa conclusion première est l'absence de consensus sur tous les aspects de la méthodologie de sondage.

Dans cette perspective, cette recherche vise à mettre en lumière l'effet des pratiques méthodologiques, reliées à la réalisation des sondages, sur l'estimation des intentions de vote. Plus précisément, cette recherche évalue les pratiques associées aux sondages mesurant les intentions de vote des résidents du Québec ayant droit de vote à l'élection fédérale de 1997.

1.1 Pratiques méthodologiques

La réalisation d'un sondage demande un nombre important d'opérations. Une description détaillée des différentes pratiques méthodologiques permettra de mettre en lumière ce processus. Dans un premier temps, il sera question du devis de sondage, dans un deuxième, des pratiques reliées à la conduite des opérations et dans un troisième, du redressement.

1.1.1 Devis de sondage

1.1.1.1 Mode de génération de l'échantillon

La génération de l'échantillon constitue la toute première étape de la réalisation d'un sondage téléphonique. À ce moment, on génère, ou on fait générer par un sous-traitant, un échantillon de numéros de téléphone. Théoriquement, pour obtenir un échantillon aléatoire, il est nécessaire d'avoir une base de sondage qui comprend tous les membres de la population à l'étude et chacun de ceux-ci doit avoir une chance connue d'être sélectionné. En pratique, les choses sont un peu plus

compliquées. De fait, peu importe le mode de génération utilisé, seules les personnes demeurant dans un ménage privé ayant le téléphone ont une chance d'être sélectionnées dans un sondage téléphonique, excluant ainsi les personnes en foyer d'accueil, hospitalisées, sans domicile fixe, etc. Il n'y a jamais adéquation parfaite entre la population à l'étude et la base de sondage employée.

L'annuaire téléphonique, fourni par les compagnies de téléphone, a été grandement utilisé au début des sondages téléphoniques comme base de sondage. Il permet le recours à des méthodes simples de sélection. Par exemple, pour un annuaire donné, un premier numéro est choisi au hasard et ensuite, suivant un certain pas de sélection (à toutes les trois, quatre colonnes ou pages), les autres numéros sont sélectionnés de façon à constituer un échantillon aléatoire systématique. L'utilisation de cette base a comme avantage de maximiser le nombre de ménages valides dans l'échantillon de départ. Le contact peut être également plus facile et plus personnalisé parce que l'on possède le nom d'une des personnes du ménage. L'utilisation de l'annuaire téléphonique sur papier comme base de sondage est de moins en moins fréquente. Il est maintenant possible d'obtenir la liste des numéros de téléphone inscrits dans les annuaires sur cd-rom. Ainsi, des logiciels permettent maintenant de sélectionner au hasard, de façon aléatoire simple, les numéros de téléphone nécessaires à l'élaboration de l'échantillon désiré. L'annuaire téléphonique, qu'il soit sur papier ou sur cd-rom, n'inclut pas les numéros de téléphone non inscrits (numéros confidentiels, récents, déménagements, etc). Dans les milieux urbains, c'est environ 25 p. cent des ménages privés qui sont dans cette situation (Taylor, 1997; Tremblay, 1981).

Pour pallier à ce problème, il est possible d'ajouter un nombre aux numéros de téléphone

sélectionnés dans l'annuaire. On ajoute, par exemple, 3 à tous les numéros de sorte que le 354-3590 (numéro sélectionné dans l'annuaire) devient le 354-3593. Cette méthode, communément appelée *cd-rom+*, permet d'accéder à davantage de numéros non publiés, tout en ayant moins de numéros non valides que lorsque l'on procède par génération aléatoire de numéros de téléphone (GANT) (Landon et Banks, 1977). Cependant, pour que cette méthode soit non biaisée, il faudrait que les numéros non listés se distribuent aléatoirement dans les différents échanges (Cantril, 1991; Lepkowski, 1988 dans Lavrakas, 1993). Comparée à la GANT, «[...] cette méthode ne donne pas une chance égale à tous les numéros attribués d'être choisis. De plus, de nouveaux échanges mis en service après la publication de l'annuaire ne pourront pas être sélectionnés»¹ (Landon et Banks, 1977: 297).

La génération aléatoire de numéros de téléphone (GANT) constitue un groupe de techniques d'échantillonnage probabiliste. Schématiquement, elle consiste à générer au hasard des numéros de téléphone en fonction des informations disponibles sur les échanges et les préfixes, les 3 ou 4 derniers chiffres des numéros de téléphone, parmi tous ceux susceptibles d'être attribués. Dans les régions où le pourcentage de numéros de téléphone non publiés est élevé, notamment dans les grands centres urbains, la génération aléatoire de numéros de téléphone est préférable. Étant donné que les numéros de téléphone sont générés au hasard, tous les ménages, que leur numéro soit publié ou non, font partie de la base de sondage (Dillman, 1979). Cependant, le taux de non-réponse est plus élevé pour les ménages ayant un numéro non publié (Dillman, 1979; Drew, Choudry et Hunter, 1988). Pour Taylor (1997), l'utilisation de la GANT, malgré les coûts supplémentaires qu'elle entraîne, est une nécessité:

¹ «[...] the design does not give all working numbers a chance to be included. For instance, new sets of numbers placed in service after directory publication will not be sampled.»

«un sondage récent réalisé par la firme Harris aux États-Unis montrait que 28 p. cent de toutes les entrevues complétées [...] provenaient de numéros de téléphone non listés dans l'annuaire et que ce segment de l'échantillon était significativement différent de l'échantillon des numéros publiés»² (Taylor, 1997: 424).

En conclusion, si on compare ces différents modes de génération de l'échantillon, il apparaît que l'utilisation de la GANT permet de générer un échantillon aléatoire non biaisé, c'est-à-dire où tous les ménages ayant le téléphone sont dans la base de sondage et ont une chance connue d'être sélectionnés.

1.1.1.2 Mode de sélection dans le ménage

Dans un sondage téléphonique visant l'ensemble de la population, un échantillon à deux degrés doit être utilisé: la sélection ou la génération de numéros de téléphone permet de sélectionner un ménage; la sélection dans le ménage permet de choisir un répondant. Traugott (1987) souligne toute l'importance qu'il faut accorder à cette étape du processus:

Les différences dans les modes de sélection du répondant dans le ménage peuvent affecter la qualité des données et ainsi, les conclusions qui sont tirées de plusieurs sondages. Ces effets sont dus à un certain nombre de facteurs, incluant la façon par laquelle les membres du ménage sont énumérés ou listés, la manière dont le répondant est sélectionné et la persistance avec laquelle les répondants sélectionnés sont rappelés dans le but d'obtenir une entrevue. Alors que les différences méthodologiques affectent les caractéristiques démographiques de l'échantillon, les conséquences les

² «A recent Harris survey in the US found 28% of all completed telephone interviews [...] were with unlisted phone numbers and this segment of the sample was significantly different from the samples of listed numbers.»

plus significatives vont se trouver dans les mesures d'attitudes et de comportements³ (Traugott, 1987: 48-49).

Dans son article, l'auteur fait la recension d'un certain nombre de méthodes. La grille de Kish, utilisée, au moins à l'occasion, par 47 p. cent des firmes québécoises et par aucune firme canadienne (Durand, Vachon et Tanguay, 1998), est la plus exigeante en temps et ferait augmenter les risques de refus de la personne contactée selon certains auteurs (Cantril, 1991; Groves et Lyberg, 1988; Lavrakas, 1993), car elle exige l'énumération complète de tous les membres du ménage par leur âge, demande pouvant être perçue comme intrusive. Cependant, elle a l'avantage d'être strictement probabiliste, c'est-à-dire que chaque membre du ménage a une chance connue d'être sélectionné. Un peu moins exigeante que la première, la grille de Trodahl-Carter exige de l'interviewer qu'il demande le nombre de personnes dans le ménage et le nombre d'hommes (ou de femmes). Grâce à ces deux questions, une matrice permet de choisir au hasard la personne qui doit être interviewée en fonction du sexe et de l'âge..

La sélection en fonction de la date d'anniversaire la plus rapprochée, à venir ou passée, est employée par les deux tiers des firmes canadiennes⁴ (Durand, Vachon et Tanguay, 1998). Elle est

³ «Variations in techniques for selecting respondents within sampled households can affect the quality of data and therefore the substantive conclusions which are drawn from many surveys. These effects derive from a number of factors, including the way in which individuals residing in the household are enumerated or listed, if at all, how a designated respondent is subsampled, and the persistence with which designated respondents are called back in order to obtain an interview. While the methodological differences affect the demographic characteristics of the sample, the most significant consequences will be found in attitudinal or behavioral measurements which may be taken from the respondents.»

⁴ La sélection en fonction de la date d'anniversaire est utilisée, au moins à l'occasion, par 60 p. cent des firmes québécoises et par 70 p. cent des firmes du Canada anglais (Durand et al., 1998).

rapide, peu intrusive et permet d'obtenir un bon taux de réponse (Salmon et Nichols, 1983; Lavrakas, 1993). Cependant, cette méthode augmente les risques d'auto-sélection, ce qui a pour effet de contribuer à la sous-représentation des personnes moins coopératives ou plus difficiles à joindre. De fait, «dans 20-25 p. cent des ménages, il y [a] une autre personne éligible dont la date d'anniversaire [est] plus récente que celle du répondant»⁵ (Lavrakas, Merkle et Bauman, 1993: 112). Autrement dit, une fois sur quatre, la mauvaise personne est sélectionnée.

La méthode «homme le plus jeune, femme la plus vieille» est encore plus simple que la précédente: l'interviewer demande, en alternance, à parler à l'homme le plus jeune dans le ménage, à la femme la plus âgée. Cette méthode n'est pas vraiment aléatoire: les membres du ménage n'ont pas une chance égale d'être sélectionnés. Cependant, du point de vue empirique, c'est-à-dire du respect des paramètres démographiques, cette méthode permet de contrôler la sous-représentation des jeunes hommes dans l'échantillon, surtout quand un faible nombre de tentatives sont faites pour joindre les répondants. Elle permet, d'une certaine façon, de faire du redressement a priori.

Plus une méthode est exigeante, plus le nombre de refus sera élevé (Salmon et Nichols, 1983). Cependant, les méthodes plus exigeantes sont strictement probabilistes, c'est-à-dire que chaque membre du ménage a une chance connue d'être sélectionné et on peut donc s'attendre à ce que les paramètres démographiques de l'échantillon soient plus conformes à ceux de la population à l'étude. Inversement, les méthodes plus rapides et moins intrusives ne sont pas probabilistes.

⁵ «[...] in 20-25% of the households there was another eligible whose birthday was more recent than that of the respondent.»

Ces méthodes de sélection du répondant sont moins intrusives que les méthodes probabilistes qui nécessitent l'utilisation de questions portant sur la composition du ménage, mais cet avantage a un coût: Sans l'utilisation des procédures probabilistes pour la sélection des répondants à l'intérieur des ménages, le sondeur n'a aucune façon de connaître, pour chaque répondant, la probabilité qu'il fasse partie de l'échantillon. Ainsi, la théorie des probabilités ne peut être utilisée pour vérifier si l'échantillon est représentatif. La variance de l'échantillon et l'erreur d'échantillonnage ne peuvent être calculés (Cantril, 1991: 99)⁶.

Les méthodes plus exigeantes ont également l'avantage de mieux contrôler le danger d'auto-sélection.

1.1.1.3 Utilisation de quotas

Une autre façon de sélectionner le répondant à l'intérieur du ménage consiste à mener des entrevues auprès de gens répondant à certains critères: sexe, âge, région, langue, etc. «Il n'y a pas alors de sélection dans le ménage jusqu'à ce que certains quotas soient remplis; la sélection se fait ensuite en recherchant des personnes dont les caractéristiques socio-démographiques permettent de remplir les quotas visés» (Durand, Vachon et Tanguay., 1998: 421). L'utilisation des quotas a, depuis les débuts des sondages, fait l'objet de nombreux débats (Blondiaux, 1991; Dreesbeke, Fichet et Tassi, 1987; March et Scarbrough, 1990; Stephenson, 1979).

Sur le plan théorique, l'utilisation des quotas ne se justifie pas. «L'échantillonnage probabiliste

⁶ «[...] these approaches to respondent selection are less invasive than probability methods that ask about household composition, but this advantage comes at a cost: Without the use of probability procedures in selecting respondents within households, the pollster has no way of knowing what chance those interviewed had of ending up in the sample. As a result, probability theory cannot be used to estimate the chance that the sample is unrepresentative. Therefore, the variance of the sample, and thus its "sampling error", cannot be computed.»

s'appuie sur la théorie mathématique de l'échantillonnage tandis que l'échantillonnage par quota ne le fait qu'approximativement. Ainsi, avec l'échantillonnage par quota nous perdons les bases théoriques pour tirer des conclusions à partir des résultats et nous ne savons jamais quand, ou à quelle fréquence, nos quotas vont nous induire en erreur»⁷ (Bradburn, 1992: 393). Selon Kish (1965; dans Blondiaux, 1991: 761), «l'échantillonnage par quota n'est pas une méthode scientifiquement définie. Au contraire, [chaque sondage par quota] ressemble à une production artistique, difficile à définir ou à décrire». Il est possible que les personnes plus difficiles à joindre et moins coopératives soient sous-représentées en utilisant cette pratique. Si dans chaque segment de la population (suivant la logique des quotas) seule une fraction de la population est sollicitée ou accepte de jouer le jeu, l'ensemble de la démarche est invalidée et la représentativité de l'échantillon sujette à caution» (Blondiaux, 1991: 765).

Sur le plan pratique, les avis sont plus partagés. Des auteurs (Taylor, 1997, Bradburn, 1992) affirment que l'utilisation des quotas permet, dans bien des cas, d'obtenir des résultats plus fiables.

Les firmes NOP et Harris, depuis 1969, utilisaient des échantillons probabilistes pour la plupart de leurs sondages d'opinion, incluant leurs sondages pré-électorales (jusqu'au milieu des années 70). Il y avait seulement deux problèmes importants. En moyenne, leurs résultats étaient légèrement moins précis dans la prédiction des résultats des élections que ceux obtenus grâce aux échantillons par quota. De plus, les coûts associés aux opérations de collecte étaient plus élevés. Malheureusement, mais sans grande surprise, NOP et

⁷ «[...] the probability sampling rests on the mathematical theory of sampling, while quota sampling does so only approximately. Thus with quota sampling we lose the theoretical basis for drawing inferences from the data, and we never know when (or how often) our quotas lead us astray».

Harris ont joint les rangs et sont revenus à l'échantillonnage par quota⁸ (Taylor, 1994: 218).

D'autres auteurs (Cantril, 1991; Stephenson, 1979; March et Scarbrough, 1990) sont plus nuancés et notent certains biais sur le plan de la représentativité de l'échantillon. Ainsi, selon Cantril (1991), l'échantillonnage par quota strict, utilisé au moins à l'occasion par 82 p. cent des firmes du Canada anglais et par aucune firme québécoise⁹ (Durand, Vachon et Tanguay, 1998), entraîne une surreprésentation des personnes plus coopératives.

Bref, sur le plan théorique, l'utilisation des quotas ne se justifie pas. Cependant, dans la pratique, les avis sont plus partagés. Les auteurs ne s'entendent pas sur l'importance des biais causés par leur utilisation. Il n'en demeure pas moins que le principal problème, outre l'absence de base théorique, est l'absence de critère permettant de juger de la qualité de tels sondages.

1.1.2 Conduite des opérations

Dans la présente section, il sera question des différentes pratiques reliées à la conduite des opérations qui, selon la documentation consultée, peuvent avoir un effet sur la représentativité de l'échantillon et la qualité des estimés. Il sera question de la prise de rendez-vous, de la récupération des refus, du nombre de tentatives de contact, de la durée de la période de collecte, de la substitution

⁸ «Both Nop and Harris (from 1969) used probability samples for most of their opinion surveys, including their pre-election polls, until the mid-'70's. There were only two major problems. On average, their results were slightly less accurate in predicting elections than quota samples. And the fieldwork costs were much higher. Sadly, but unsurprisingly, NOP and Harris joined the pack and went back to quota sampling».

⁹ Ces informations sont basées sur l'auto-déclaration des firmes.

de numéros de téléphone et finalement, du redressement.

1.1.2.1 Prise de rendez-vous

Devant un refus potentiel ou en l'absence de la personne sélectionnée, la prise de rendez-vous permet de faire augmenter le taux de réponse et ce, surtout quand le questionnaire est long (Collins, Sykes, Wilson et Blackshaw, 1998). «Le plus souvent, la raison donnée pour refuser (57 p. cent) est que le moment où l'interview est sollicité ne convient pas ou dérange. Pour cette raison, il est souhaitable que le personnel interviewer puisse rappeler le répondant sélectionné à un moment plus approprié pour compléter l'entrevue¹⁰ (Cantril, 1991: 100)». Cette pratique permet de faire diminuer le nombre de refus circonstanciels: personnes occupées, fatiguées, etc. (Cantril, 1991).

1.1.2.2 Récupération des refus

La récupération des refus vise à faire diminuer les biais introduits par la sous-représentation des personnes moins coopératives. «Entre un cinquième et un tiers des ménages approchés par les interviewers [...] refusent de participer au sondage¹¹» (Cantril, 1991: 99). S'il y a récupération, les refus du répondant, c'est-à-dire de la personne sélectionnée et les refus du ménage, c'est-à-dire d'une tierce personne avant ou après la sélection, devraient se diviser en deux parts égales¹² (Chevalier et

¹⁰ «Most often the reason for refusals (57%) is the inconvenience or timing of the request for an interview. For this reason it is desirable that the interviewing staff have the flexibility to call a designated respondent back at a more convenient time to complete an interview».

¹¹ «From one fifth to one third of the households approached by interviewers [...] refuse to participate in a poll».

¹² S'il n'y a pas de récupération, les refus du ménage formeront la plus grande part des refus, car ils sont beaucoup plus fréquents. Cependant, si l'on fait de la récupération, une bonne

Durand, 1997; Collins, Sykes, Wilson et Blackshaw, 1988).

La récupération des refus fait augmenter significativement le taux de réponse, soit entre 7 et 12,3 p. cent (Chevalier et Durand, 1997 [11,2]; O'Neil, 1979 [12,3]; Triplett, 1997 [7]). Cette pratique entraîne habituellement un redressement des paramètres démographiques de l'échantillon (Triplett, 1997; O'Neil, 1979; Triplett, Blair, Hamilton et Kang, 1997; Chevalier et Durand, 1997; Groves et Lyberg, 1988). «Les refus du répondant peuvent avoir un effet - d'une ampleur inconnue - sur les estimés. Une façon de gérer ce problème consiste à minimiser le taux de non-réponse¹³» (O'Neil, 1979). Ainsi, cette pratique est de toute première importance si l'on veut obtenir un échantillon représentatif de la population à l'étude et des estimés fiables. «Il s'est avéré que la proportion des répondants ayant finalement consenti à compléter l'entrevue s'expliquait davantage par la récupération des refus que par la façon dont le sondage était présenté au début¹⁴» (Cantril, 1991: 100).

1.1.2.3 Nombre de tentatives de contact

Il y a consensus chez plusieurs chercheurs (Cantril, 1991; O'Neil, 1979; Traugott, 1987; Triplett, 1998; Triplett, Blair, Hamilton et Kang, 1996) pour dire que le nombre de tentatives de

proportion de ces refus du ménage pourront être récupérés et ainsi, les refus du ménage et les refus du répondant auront tendance à se diviser en deux parties égales.

¹³ «Respondent refusals [...] can have an effect - of unknown magnitude - on sample estimates. One way of dealing with this problem is to try to minimize the nonresponse rate».

¹⁴ «It has been shown that the proportion of individuals ultimately willing to consent to a telephone interview has more to do with attempts to convert refusals than with the way the survey is introduced to the respondent.»

contact affecte le taux de réponse et les paramètres démographiques de l'échantillon. Pour un court sondage pré-électoral, «Traugott a conclu qu'après quatre appels (trois rappels), l'échantillon présentait les caractéristiques très similaires à celles observées en contactant l'autre tiers de l'échantillon éventuel en faisant cinq appels ou plus¹⁵» (Cantril, 1991: 101). Par rapport à l'échantillon final, après trois appels pour compléter l'entrevue, 57 p. cent des entrevues sont complétées, après quatre, 68 p. cent et après 10, 92 p. cent (Traugott, 1987). Utilisant les données d'un long sondage (25 minutes en moyenne) portant sur les pratiques culturelles des Québécois, Chevalier et Durand (1997) considèrent «qu'il faut [...] jusqu'à sept ou huit tentatives pour joindre le ménage sélectionné pour s'assurer de minimiser efficacement les différences entre personnes non-rejointes et rejointes» (Durand, Vachon et Tanguay, 1998: 422). Après 4 tentatives pour joindre le ménage sélectionné, le taux de réponse atteint 54,1 p. cent, après 7, 58,2 p. cent et après 10, 58,9 p. cent.

1.1.2.4 Durée de la période de collecte

La durée de la période de collecte peut donner, dans certains cas précis, des indices quant aux efforts mis pour obtenir un bon taux de réponse (Sebold, 1988). Cependant, la durée de la période de collecte n'est pas en soi garante des efforts mis pour joindre les personnes plus difficiles à contacter et moins coopératives. Plusieurs autres facteurs entrent en ligne de compte: les pratiques de récupération des refus, le nombre de tentatives de contact, le nombre de lignes affectées à l'étude, etc. Pour De Heer, «la durée de la période de collecte des données est d'importance secondaire, la récupération des refus et la réaffectation occupent une place unique, indépendamment de la durée de

¹⁵ «Traugott concluded from the study that “by four calls (three call-backs) the sample assumes characteristics that are very close to those observed by contacting another one-third of the eventual sample using five or more calls.»

la période de collecte des données¹⁶» (De Heer, 1996: 10). Cependant, on peut penser qu'il y a une relative adéquation quand la seule règle qui prévaut en ce qui a trait au nombre de tentatives de contact est le nombre d'appels qui peuvent être réalisés dans le laps de temps dévolu à la réalisation de l'étude. «Pour plusieurs sondages commerciaux, la période de collecte est tellement réduite qu'il n'y a pas assez de temps disponible pour réessayer de joindre les répondants qui n'étaient pas disponibles au début de l'enquête¹⁷» (Traugott, 1987: 51).

Dans certaines situations, on peut penser qu'il y a une certaine influence de la durée de la période de collecte sur le taux de réponse. Cependant, il faut être très prudent dans l'interprétation de cet indicateur, plusieurs autres facteurs entrant en ligne de compte.

1.1.2.5 Substitution de numéros de téléphone

La substitution de numéros de téléphone consiste à mettre sur le terrain un nouveau numéro de téléphone quand il est impossible de joindre le ménage après un certain nombre d'appels ou quand il y a refus. Cette pratique n'aide pas à améliorer la représentativité de l'échantillon et contribue à faire diminuer le taux de réponse. De fait, «bien que cette méthode soit utile pour obtenir le nombre de cas désirés, elle tend à remplacer les non-répondants par des personnes qui ressemblent davantage aux répondants qu'aux non-répondants et elle n'apparaît pas être très utile pour contrôler les biais de la

¹⁶ «The length of the data collection period seems of secondary importance; refusal conversion and reassignment take place independent of the length of the data collection period.»

¹⁷ «[...] in many commercial surveys, the field period is so truncated that not enough elapsed time is available to pursue some respondents who were unavailable at the start of interviewing.»

non-réponse¹⁸» (Smith, 1983: 53). Selon Dillman (1978), cette procédure doit, autant que possible, être évitée.

1.1.3 Redressement

Il apparaît que certaines pratiques méthodologiques peuvent avoir des effets néfastes sur la représentativité de l'échantillon. Si certaines portions de la population, notamment les personnes moins coopératives et plus difficiles à joindre, sont mal représentées et qu'elles ont des comportements et des attitudes différentes du reste de la population, les résultats obtenus pourront être biaisés. Pour tenter de palier à ce problème, plusieurs firmes font du redressement, c'est-à-dire qu'elles pondèrent les résultats de leur sondage en fonction de la répartition de certaines variables démographiques dans la population: sexe, région, langue et groupe d'âge habituellement. Cette procédure est également utilisée pour les sondages par quota. Dans cette situation, on effectue un redressement sur la base de facteurs autres (langue, groupe d'âge, etc.) que ceux utilisés pour les quotas (sexe, région, etc.).

Cette procédure soulève un certain nombre de questions. Les biais dus à la non-réponse ne peuvent être corrigés que partiellement. Premièrement, les données de référence utilisées, celles du recensement, remontent à 1991 dans certains cas. Deuxièmement, le fait de redresser en fonction de certaines variables (sexe, âge, etc.) peut avoir des effets inattendus sur la distribution d'autres variables (revenu, profession, etc.). Troisièmement, par cette pratique, on se trouve à attribuer aux

¹⁸ «While this method is useful to achieve a prescribed sample size, it tends to replace nonrespondents with people who resemble respondents rather than nonrespondents, and this approach does not appear to be widely used to deal with nonresponse bias».

non-répondants les mêmes attitudes et comportements qu'aux répondants ayant un profil socio-démographique similaire. Par exemple, si dans un échantillon¹⁹, il y a 40 p. cent d'hommes et qu'en fait, pour être représentatif de la population à l'étude, il doit y en avoir 49 p. cent, on effectuera un redressement des résultats afin de tenir compte de ce fait (voir tableau I). Autrement dit, on attribuera aux 18,4 p. cent d'hommes non-joints (9 p. cent de l'échantillon) le même résultat (40 p. cent d'appui pour le parti X) qu'aux 81,6 p. cent d'hommes ayant répondu. Toutefois, les intentions de vote réelles des hommes non-joints peuvent être tout à fait différentes. Si l'intention de vote réelle des hommes non rejoints est de 50 p. cent d'appui pour le parti X et que l'on redresse en fonction du résultat imputé plutôt que du résultat réel, l'estimation des intentions de vote des hommes pour le parti X sera, dans notre exemple, sous-estimé de 1,8 points de pourcentage.

Tableau I. Redressement des résultats pour les hommes, données imputées et réelles

	% de l'échantillon	% des hommes	Résultats réels et imputés	Résultats après redressement
Hommes rejoints	40%	81,6% (40/49)	40%	32,6%
Hommes non-rejoints	9%	18,4% (9/49)	40% (imputé)	7,4%
			50% (réel)	9,2%
Hommes (total)			(imputé)	40%
			(réel)	41,8%

Cet exemple permet de comprendre les conséquences reliées à l'utilisation de cette procédure.

¹⁹ Ces chiffres sont fictifs. Ils ne sont utilisés que dans le but d'illustrer le propos.

Comme les personnes plus difficiles à joindre et celles moins coopératives ont des caractéristiques spécifiques (O'Neil, 1979; Chevalier et Durand, 1997; Traugott, 1987), on peut penser que plus le redressement est important, plus les biais dus à la non-réponse risquent d'être importants.

Comme pour la plupart des pratiques méthodologiques et comme ailleurs dans le monde (Taylor, 1994, 1997), il n'y a pas au Canada consensus sur l'utilisation du redressement. Entre autre, il n'y a pas de variables de redressement qui font l'unanimité. Selon Taylor (1997), peu de recherches ont été réalisées pour trouver de bonnes variables qui tiennent compte des différentes réalités nationales. De plus, l'auteur note certains problèmes reliés à l'utilisation du vote à l'élection précédente comme variable de redressement.

Après toutes les élections au Royaume-Uni et en Amérique, plus de personnes affirment avoir voté pour le gagnant qu'il n'y en a eu en réalité. Quand les gouvernements deviennent moins populaires, davantage d'électeurs tendent à cacher qu'ils ont voté pour eux. Si la race est un facteur, les gens qui ont voté contre un candidat noir affirment parfois avoir voté pour lui²⁰ (Taylor, 1994: 217).

Tous les éléments ci-haut mentionnés permettent de douter de la capacité de cette procédure à corriger adéquatement les biais entraînés par la non-réponse. Au-delà du fait d'être un homme, ou d'avoir toute autre caractéristique démographique contrôlée, les non-répondants peuvent avoir des attitudes et des comportements différents de ceux du reste de la population (O'Neil, 1979; Chevalier et Durand, 1997; Traugott, 1987). Ainsi, imputer les résultats des répondants aux non-répondants

²⁰ «After all election in the UK and America more people claimed to have voted for the winner than did so. When governments become more unpopular more voters than to deny they voted for them. If race is a factor, people who voted against black candidate sometimes claim to have voted for him.»

n'est pas sans risque de biais.

1.2 Représentativité de l'échantillon

Dans l'étude réalisée par Chevalier et Durand (1997), il apparaît que dans 9 p. cent des entrevues complétées, il y a eu précédemment refus du ménage et dans 10 p. cent, refus de la personne sélectionnée. Ainsi, dans 19 p. cent des entrevues finalement complétées, il y a eu un refus initial qui, subséquemment, a été récupéré. Sachant que les personnes moins coopératives ont des caractéristiques socio-démographiques, des comportements et des attitudes propres (Chevalier et Durand, 1997; Groves et Lyberg, 1988; O'Neil, 1979; Triplett, 1997; Triplett, Blair, Hamilton et Kang, 1996), il apparaît évident que la récupération des refus peut avoir un effet important sur la représentativité de l'échantillon et sur la qualité des estimés. Il faut noter que le profil des répondants moins coopératifs diffère selon le type de refus: refus du ménage ou refus de la personne sélectionnée (Chevalier et Durand, 1997).

Les répondants plus difficiles à joindre ont, eux aussi, un profil spécifique (Chevalier et Durand, 1997; Sebold, 1988; Traugott, 1987; Triplett, 1997). De plus, il apparaît que les répondants ayant initialement refusé ont un profil socio-démographique opposé à celui des personnes plus difficiles à joindre (Chevalier et Durand, 1997; O'Neil, 1979). Bref, «l'ampleur des biais de la non-réponse est fonction de la différence de caractéristiques entre les répondants et les non-répondants, et du taux de non-réponse lui-même²¹» (Groves et Lyberg, 1988 dans Drew, Choudry et Hunter, 1988: 235).

²¹ «The magnitude of respondents bias is a fonction both of the difference in characteristics of respondents and nonrespondents, and of the nonresponse rate itself».

1.2.1 Représentativité des personnes moins coopératives

Il est nécessaire de faire la distinction entre les refus du ménage et les refus de la personne sélectionnée. Dans un premier temps, il sera question du profil des répondants ayant initialement refusé de participer et dans un deuxième temps, de ceux demeurant dans un ménage où il y a eu initialement refus.

Le tableau II présente les caractéristiques des répondants selon leur degré de coopération.

Tableau II. Profil du répondant coopératif et du répondant moins coopératif

Variable socio-démographique	Répondant coopératif (aucun refus)	Répondant moins coopératif (ayant initialement refusé)
Occupation	Professionnel (O'Neil, 1979)	Travailleur qualifié (O'Neil, 1979)
Scolarité		Moins scolarisé (O'Neil, 1979; Triplett, 1997; Triplett et al., 1996; Chevalier et Durand, 1997; Groves et Lyberg, 1988)
Race	Noir (O'Neil, 1979)	Blanc (O'Neil, 1979) Noir (Triplett et al., 1997)
Ethnicité	Noir africain (O'Neil, 1979)	Allemand et Polonais (O'Neil, 1979)
Âge	25-34 ans (O'Neil, 1979)	65-74 ans (retraité) (O'Neil, 1979; Triplett, 1997; Triplett et al., 1996; Chevalier et Durand, 1997; Groves et Lyberg, 1988)
Ménage	Présence d'enfants (O'Neil, 1979)	Absence d'enfant (O'Neil, 1979; Triplett, 1997) Absence d'enfant de moins de 18 ans (Triplett, 1997) Une ou deux personnes (Chevalier et Durand, 1997)
Logement	Autre type de logement (O'Neil, 1979) Locataire (O'Neil, 1979)	Bungalow (O'Neil, 1979) Propriétaire (O'Neil, 1979) Édifice cinq logements et plus (Chevalier et Durand, 1997)
Revenu		Faible revenu (moins de 20 000\$) (Chevalier et Durand, 1997)
Région		Banlieue (ouest, est et nord de Montréal) (Chevalier et Durand, 1997) Plus en milieu urbain (p/r rural) (Groves et Lyberg, 1988; Triplett, 1997)

Malgré quelques divergences entre les auteurs, le profil des personnes moins coopératives est assez cohérent. Il s'agit le plus souvent de personnes âgées, ayant un faible niveau de scolarité, un faible revenu. Travailleurs qualifiés à la retraite, ils demeurent plus souvent en milieu urbain. Selon Groves et Lyberg (1988), ces personnes âgées refusent davantage de participer à cause de problèmes auditifs

et de leur désengagement social qui les amène à hésiter à interagir avec un étranger. La peur d'être harcelé et une socialisation différente envers le téléphone peuvent également être des facteurs explicatifs. Le profil socio-démographique de ces personnes est assez typé.

Les personnes demeurant dans un ménage où il y a eu initialement refus ont également un profil assez précis (tableau III).

Tableau III. Profil du répondant demeurant dans un ménage où il y a eu initialement refus (Chevalier et Durand, 1997)

Variable socio-démographique	Caractéristique du répondant
Âge	15-24 ans
Scolarité	Secondaire ou cégep (aux études)
Ménage	Deux à quatre personnes
Revenu du ménage	40 000 à 60 000\$
Logement	Maison unifamiliale
Langue	Anglophone, allophone
Région	Ouest de l'île de Montréal

Selon les auteurs, «le profil type du refus de ménage serait [...] celui où les parents, plus souvent un ménage de la banlieue de Montréal, bloquent l'accès au jeune sélectionné».

1.2.2 Représentativité des personnes plus difficiles à joindre

Il est important de ne pas confondre le nombre de tentatives de contact pour joindre le ménage et le nombre d'appels nécessaires pour compléter l'entrevue. Sur le plan conceptuel, ces indicateurs

sont différents: le premier permet d'obtenir des informations sur les caractéristiques des répondants demeurant dans un ménage plus difficile à joindre et le second, sur ceux dont l'entrevue a été plus difficile à compléter. Cependant, sur le plan des caractéristiques des ménages et des répondants, il y a une relative adéquation entre ces indicateurs.

Selon Sebold (1988), en utilisant la GANT (génération aléatoire de numéros de téléphone) pour générer l'échantillon, après 4 appels, 50% des numéros où il n'y a pas encore eu contact sont des ménages et après 20 appels, 5%. Ainsi, si l'on veut faire diminuer la non-réponse attribuable aux non-contacts, il apparaît important de faire un certain nombre de rappels pour joindre le ménage. De fait, selon plusieurs chercheurs (Cantril, 1991; O'Neil, 1979; Sebold, 1988; Traugott, Groves et Lepkowski, 1987; Traugott, 1987; Triplett, 1997; Triplett, Blair, Hamilton et Kang, 1996), le nombre de tentatives de contact affecte la représentativité de l'échantillon. Le tableau IV présente les caractéristiques des répondants demeurant dans un ménage plus difficile à joindre et de ceux pour qui l'obtention de l'entrevue a demandé plus de contacts.

Tableau IV. Profil des répondants plus difficiles à joindre

Variable socio-démographique	Caractéristiques des répondants demeurant dans un ménage plus difficile à joindre	Caractéristiques des répondants dont l'entrevue a été plus difficile à compléter
Âge	25-34 ans (Chevalier et Durand, 1997) Moins souvent âgé de 65 ans et plus (Sebold, 1988)	25-34 ans (Chevalier et Durand, 1997*; Traugott, 1987) 18-24 ans (Triplett, 1997)
Sexe		Homme (Chevalier et Durand, 1997*; Traugott, 1987; Triplett, 1997)
Race		Noir (Triplett, 1997)
Ménage	Une personne (Chevalier et Durand, 1997) Célibataire (Sebold, 1988)	Célibataire (Triplett, 1997)
Scolarité	Universitaire (Chevalier et Durand, 1997)	Universitaire (Chevalier et Durand, 1997*)
Statut	En emploi ou au chômage (Chevalier et Durand, 1997)	Pas retraité (Chevalier et Durand, 1997*)
Revenu du ménage		80 000\$ et plus (Chevalier et Durand, 1997*)
Logement	Locataire (Chevalier et Durand, 1997) Édifice de 5 logements et plus (Chevalier et Durand, 1997)	
Région	Centre de l'île de Montréal (Chevalier et Durand, 1997)	Centre et ouest de Montréal (Chevalier et Durand, 1997*)
Langue		Anglophone, allophone, 1997 (Chevalier et Durand, 1997*)

*: Ces informations ne figurent pas dans la version finale du texte de Chevalier et Durand (1997). Elles ont été tirées d'une version de travail communiquée par les auteurs.

Il y a une relative adéquation entre les deux indicateurs. Ainsi, il apparaît que les personnes plus difficiles à joindre sont typiquement plus souvent jeunes, assez scolarisées, célibataires et vivent en appartement en milieu urbain.

1.3 Qualité des estimés

Les différentes pratiques méthodologiques ne s'équivalent pas, certaines permettant d'obtenir un échantillon plus représentatif de la population à l'étude et par le fait même, des résultats plus fiables. Des études ont été menées afin de mesurer l'impact de certaines pratiques, notamment celles entraînant la sous-représentation des personnes plus difficiles à joindre et moins coopératives, sur la qualité des estimés. Cependant, il faut noter que peu de recherches analysent les effets de la sous-représentation de ces groupes sur la qualité de la mesure des intentions de vote. Pour ce qui est de la sous-représentation des personnes moins coopératives, aucune recherche n'a été trouvée. Pour ce qui est de la sous-représentation des personnes plus difficiles à joindre, trois recherches américaines traitent de la question (Traugott, 1987; firme Harris, 1988 et firme Gallup, 1988 dans Cantril, 1991). De plus, certains chercheurs (Chevalier et Durand, 1997; O'Neil, 1979; Triplett, 1997; Triplett, Blair, Hamilton et Kang, 1996) ont exploré les effets de la sous-représentation de ces groupes sur la qualité de certains estimés mais non sur la mesure des intentions de vote.

Les personnes ayant initialement refusé leur coopération sont en général moins impliquées socialement, plus isolées (O'Neil, 1979).

D'autres variables qui ont été examinées mais qui ne sont pas explicitement prises en compte dans cet article permettent de penser que les personnes moins coopératives sont moins intégrées socialement que les autres répondants. Elles sont significativement moins portées à participer à des activités ou à des organisations communautaires (activités et organisations interprétées au sens large) que les autres répondants. De plus, elles sont moins portées à indiquer soit leur volonté ou bien une occasion où elles ont appelé la police pour rapporter des crimes qu'elles ont observés dans leur quartier, celles qui ont

appelé l'ont fait moins fréquemment²².

Chevalier et Durand (1997) arrivent à des résultats qui, bien que ne portant pas sur le même sujet, sont assez congruents avec ces derniers. En effet, il apparaît que les répondants ayant initialement refusé²³ font moins de sorties culturelles²⁴ (2,34 en moyenne par semaine) que les répondants demeurant dans un ménage où il n'y a eu aucun refus (3,64 par semaine). Cependant, sur le plan de l'écoute télévisuelle, il n'y a pas de différence significative. Ainsi, en concordance avec l'étude de O'Neil (1979), les personnes ayant initialement refusé sortent moins, sont moins impliquées socialement. Elles fournissent également moins d'informations: plus fort taux de non-réponse partielle aux questions et entrevue plus courte (Triplett, 1997).

Les répondants demeurant dans un ménage où il y a eu initialement refus²⁵ fournissent des informations de qualité comparable à celles des autres (Triplett, 1997). De plus, ils écoutent

²² «Other variables for which data were examined but which were not explicitly considered in this paper provide grounds for belief that resisters may be less socially participating than other respondents. They are markedly less likely to indicate participation in the activities of neighborhood organizations (with both «activities» and «organizations» broadly interpreted) than their amenable counterparts. Furthermore they are less likely to indicate either an inclination or a history of calling the police to report observed crimes in their neighborhoods and those who have called are likely to have done so less frequently.»

²³ Il s'agit de répondants ayant initialement refusé de passer l'entrevue qui, après une ou plusieurs tentatives de récupération, ont accepté de collaborer.

²⁴ Nombre de sorties au cours des trois derniers mois incluant le théâtre, le cinéma, les spectacles de musique, de variétés et de danse.

²⁵ Il s'agit des répondants demeurant dans un ménage où il y a eu refus d'un membre quelconque du ménage avant sélection qui, après une ou plusieurs tentatives de récupération, ont passé l'entrevue.

davantage la télévision (31,1 heures en moyenne par semaine) que les autres (25,6 heures par semaine) alors que sur le plan des sorties culturelles, il n'y a pas de différence significative (Chevalier et Durand, 1997).

Pour ce qui est des répondants plus difficiles à joindre, il faut tout d'abord distinguer les recherches disponibles en fonction du type d'indicateur utilisé: le nombre de tentatives pour joindre le ménage ou le nombre d'appels nécessaires pour compléter l'entrevue. Le premier indicateur fournit des informations sur les comportements des répondants demeurant dans un ménage plus difficile à joindre et le second, sur ceux dont l'entrevue a été plus difficile à compléter. La recherche de Chevalier et Durand (1997) a utilisé ces deux indicateurs²⁶. Toutefois, que l'on utilise l'un ou l'autre des indicateurs, on peut constater que les personnes plus difficiles à joindre écoutent moins la télévision et font davantage de sorties.

Plus directement reliées au sujet de cette étude, les recherches de Traugott (1987), de la firme Harris (1988) et de la firme Gallup (1988) (dans Cantril, 1991) ont démontré qu'un faible nombre de tentatives de contact pour compléter l'entrevue entraîne un biais en faveur des Démocrates (voir tableaux V et VI), les répondants plus difficiles à joindre, c'est-à-dire typiquement les jeunes, les hommes, les célibataires et les personnes demeurant en milieu urbain (Chevalier et Durand, 1997; Sebold, 1988; Traugott, 1987; Triplett, 1997), étant davantage républicains (Traugott, 1987; Cantril,

²⁶ Les informations portant sur les pratiques culturelles des personnes plus difficiles à joindre, sur la base du nombre d'appels pour compléter l'entrevue, ne figurent pas dans la version finale du texte de Chevalier et Durand (1997). Ces informations proviennent d'une version de travail communiquée par les auteurs.

1991).

Tableau V. Nombre d'appels faits pour compléter l'entrevue et intentions de vote (élection américaine de 1988)

Nombre d'appels	Bush	Dukakis
1 ^{er} appel	47,8% (49,5%*)	44,8%
2 ^e appel	48,9%	43,6%
3 ^e appel	49,1%	43,4%
4 ^e appel	(50,9%*)	

Gallup, 1988; *: Harris, 1988

Tableau VI. Nombre d'appels faits pour compléter l'entrevue et intentions de vote (élection américaine de 1984)

Nombre d'appels	Reagan	Mondale
1 ^{er} appel	48%	45%
2 ^e appel	51%	45%
3 ^e appel	51%	42%
4 ^e appel	52%	40%

Traugott, 1987

Au premier appel, il y a surreprésentation des Démocrates par rapport à l'échantillon final. Après quatre appels, les paramètres socio-démographiques, notamment l'âge, le sexe et l'éducation, sont plus fidèles à la réalité et la proportion de Démocrates diminue, à l'avantage des Républicains (Traugott, 1987). Les personnes plus conservatrices sont plus difficiles à joindre. Bref, un plus grand nombre d'appels améliore la fidélité des paramètres démographiques de l'échantillon et modifie l'estimation des intentions de vote.

Transposées au contexte québécois, les conclusions de ces trois recherches permettent de formuler l'hypothèse suivante: les personnes plus difficiles à joindre sont plus conservatrices et par conséquent, appuieraient davantage le Parti conservateur et le Parti libéral du Canada aux élections fédérales et le Parti libéral du Québec aux élections provinciales. Autrement dit, un faible nombre d'appels aurait tendance à sous-estimer les intentions de vote pour les partis plus conservateurs. Cependant, cette hypothèse ne peut pas être retenue sans un questionnement plus approfondi. Premièrement, la dynamique électorale qui prévaut au Québec est fort différente de celle des États-Unis, notamment à cause de la question nationale. Elle s'articule davantage autour du continuum fédéraliste/souverainiste que du continuum conservateur/libéral (droite/gauche) (voir tableau VII).

Tableau VII. Intentions de vote des Québécois à l'élection fédérale de juin 1997 en fonction de leur choix lors du référendum de 1995 sur la souveraineté

Vote avoué lors du référendum de 1995	Intentions de vote				
	PLC	BQ	PC	Autres	NSP/NRP
OUI	8%	57%	14%	2%	19%
NON	44%	3%	31%	4%	18%

Résultats basés sur les sondages de SOM publié dans La Presse du 29 mai 1997, avant répartition des indécis.

Deuxièmement, les firmes de sondage du Québec font sept tentatives de contacts pour joindre le ménage en moyenne et dans le Reste du Canada, cinq ($p=,016$) (Durand, Vachon et Tanguay, 1998). Selon Traugott (1987), après 4 appels pour compléter l'entrevue, les caractéristiques des répondants ne diffèrent pas significativement de celles des non-répondants²⁷. Ainsi, étant donné que toutes les

²⁷ Il s'agit de personnes qui n'ont pas complété l'entrevue au quatrième appel, mais qui, après un certain nombre de rappels, ont finalement répondu.

firmes québécoises et canadiennes font plus de quatre appels, peut-on encore faire l'hypothèse que cette pratique va avoir une incidence sur les paramètres de l'échantillon et l'estimation des intentions de vote?

2 Hypothèse

Certaines pratiques méthodologiques peuvent affecter la représentativité de l'échantillon et la qualité des estimés. De fait, si les répondants moins coopératifs (Chevalier et Durand, 1997; Groves et Lyberg, 1988; O'Neil, 1979; Triplett, 1997; Triplett et al., 1997) et ceux plus difficiles à joindre (Chevalier et Durand, 1997; Harris, 1988 dans Cantril, 1991; Gallup, 1988 dans Cantril, 1991; Sebold, 1988; Traugott, 1987; Triplett, 1997) ont un profil socio-démographique, des attitudes et des comportements typés, les pratiques méthodologiques qui entraînent leur sous-représentation risquent de causer des biais dans les estimés. Ainsi, l'hypothèse générale de la présente recherche est que les pratiques méthodologiques moins strictes sont associées à une moins bonne estimation des intentions de vote.

Cette hypothèse peut apparaître assez triviale à première vue, mais en réalité, elle est loin de l'être. À ce propos, Durand, Vachon et Tanguay (1998) notent des différences significatives, tant sur le plan des pratiques méthodologiques que des attentes, entre les firmes du Québec et celles du Canada anglais. Ils notent que «les taux de réponse pour un sondage téléphonique auprès de la population se [situent] entre 50 et un peu plus de 60 p. cent pour les firmes du Québec alors que ceux qu'[obtiennent] les firmes du Reste du Canada [varient] généralement entre 20 et 40 p. cent» (Durand, Vachon et Tanguay, 1998: 419). Sur le plan des attentes, le taux de réponse jugé acceptable par les directeurs de firme de sondage du Québec est de 56 p. cent comparativement à 43 p. cent au Canada anglais. Ainsi, qu'il s'agisse du taux de réponse ou de certaines pratiques méthodologiques, notamment le mode de sélection des répondants, le nombre de tentatives pour les joindre et la récupération des refus, il y a des différences significatives dans les pratiques et ces différences sont fortement liées aux régions (Québec et Canada anglais) (Durand, Vachon et Tanguay, 1998). Il n'y

Il n'y a pas de consensus sur la meilleure façon de réaliser un sondage et les firmes ont tendance à justifier leurs pratiques, quelles qu'elles soient, par leurs résultats, à savoir une estimation de l'intention de vote qui serait toujours adéquate. Cette recherche vise à déterminer si les différences dans les pratiques méthodologiques entraînent des différences dans l'estimation des intentions de vote, et plus spécifiquement si des pratiques moins strictes entraînent une moins bonne estimation.

3 Méthodologie

3.1 Données

Les données utilisées pour la présente recherche proviennent des sondages et des strates de sondages, c'est-à-dire de portions de sondages pan-canadiens, mesurant les intentions de vote des résidents du Québec pendant la campagne électorale fédérale de 1997. Durant la campagne, soit du 2 au 28 mai, les résultats de 10 sondages réalisés au Québec et de 15 sondages effectués dans l'ensemble du Canada ont été publiés dans les médias. Dans le cadre de cette recherche, les 10 réalisés au Québec et 7 des 15 sondages pan-canadiens²⁸ sont à l'étude (voir Annexe 3). Pour les sondages, la taille moyenne des échantillons est de 1003, le plus petit comptant 980 entrevues complétées et le plus grand 1014, et pour les strates, elle est de 600, le plus petit comptant 270 entrevues complétées et le plus grand 970. Les informations portant sur ces sondages ont été recueillies par le biais de questionnaires aux firmes de sondage (voir Annexe 2), dans les principaux quotidiens (La Presse, The Gazette, The Globe and Mail) et sur internet (Chroniques de Sibérie, Élection Canada, sites des firmes de sondage). Le questionnaire a été élaboré de façon à couvrir l'ensemble des pratiques méthodologiques pouvant avoir une incidence sur la justesse de l'estimation des intentions de vote. Il comprend cinq sections: les résultats des sondages, le questionnaire, le devis de sondage, la gestion de terrain et la base administrative (voir Annexe 2).

Cette recherche vise à déterminer si des pratiques moins strictes entraînent une moins bonne estimation des intention de vote. Ce faisant, les résultats des sondages constituent les variables de base

²⁸ Les sondages exclus de l'analyse ont été réalisés par la firme américaine Zogby et par le Strategic Council. Dans le paysage canadien des firmes de sondage, elles sont de moindre importance. De plus, il a été impossible de recueillir toutes les informations nécessaires à leur examen.

de la présente étude. Parmi toutes les pratiques méthodologiques mesurées, plusieurs n'ont pu être analysées. Dans certains cas, il n'est pas possible de distinguer l'effet de firmes de celui des pratiques: par exemple, il y a autant de libellés de la question sur l'intention de vote que de firmes. Plusieurs pratiques, loin de se distribuer aléatoirement dans les différentes firmes, forment plutôt des amalgames. Autrement dit, elles sont fortement corrélées les unes avec les autres. Ainsi, si l'on voulait tester l'effet spécifique d'une de ces pratiques, il ne serait pas possible de déterminer si cet effet est causé par celle-ci ou par une des autres pratiques avec laquelle elle est corrélée. De plus, étant donné que seulement dix-sept sondages sont à l'étude, le nombre de variables contrôles qui peuvent être incluses dans les analyses est limité. Ces différentes contraintes ont fait apparaître la nécessité de classer les sondages.

3.2 Mesures

3.2.1 Mesures de l'intention de vote

Afin d'évaluer la fiabilité des sondages pré-électoraux, il est nécessaire d'élaborer des indicateurs de la qualité de l'estimation des intentions de vote. Pour ce faire, il faut une base de comparaison. Il faut un modèle de l'évolution des intentions de vote pour chaque parti tout au long de la campagne.

Les résultats de sondage ne peuvent pas être immédiatement analysés. Il faut d'abord élaborer une série de base pour chaque parti qui servira de variable dépendante, c'est-à-dire qu'il faut répartir les résultats des sondages afin d'obtenir des séries quotidiennes des intentions de vote. Pour ce faire, les résultats de chaque sondage réalisé durant la campagne électorale ont été répartis sur chacune des journées où il y a eu collecte d'information. Une fois répartis, les résultats ont été pondérés par la

taille de leur échantillon, permettant ainsi de tenir compte de l'importance des différents sondages. Finalement, des moyennes pondérées ont été calculées pour chaque jour de la campagne. Ce sont ces moyennes quotidiennes qui forment la série de base. Il est à noter que les résultats d'un sondage, ou plus précisément les résultats à une des journées où il y a eu collecte d'information, ont été retirés de la série de chaque parti: au jour 12 de la série (13 mai 1997), les résultats d'un seul sondage (intentions de vote pour chaque parti), largement déviants, servaient au calcul des résultats de cette journée²⁹.

Pour modéliser l'évolution des intentions de vote, il faut tenir compte des principaux événements de la campagne et de la dynamique des intentions de vote propre à chaque parti, ceux-ci constituant les variables indépendantes de l'analyse. Parmi tous les événements de la campagne, voici ceux qui ont joui d'une meilleure couverture médiatique et ainsi, qui auraient été plus susceptibles d'affecter les intentions de vote: Preston Manning, le Chef du Parti réformiste, dévoile le programme du Parti libéral du Canada quelques jours plus tôt que prévu (30 avril 1997); Michel Vastel publie, dans *Le Soleil*, des extraits du livre de Jacques Parizeau (7 mai 1997); le débat en anglais (12 mai 1997); le premier débat en français³⁰ (13 mai 1997); le deuxième débat en français (18 mai 1997); et finalement, Jean Chrétien donne une entrevue dans laquelle il affirme qu'il n'aurait jamais accepté une victoire serrée du oui au référendum de 1995 (26 mai 1997). Tous ces événements ont été testés. Seul le

²⁹ Des analyses réalisées pour chaque parti, avant et après le retrait de ces résultats déviants (remplacés par des interpolations linéaires) ont permis de constater que la variance résiduelle et l'erreur-type étaient beaucoup moindre sans ceux-ci.

³⁰ Il y a eu deux débats en français. Le premier (13 mai 1997) a été interrompu parce que la modératrice a eu un malaise.

premier débat en français est apparu comme ayant un effet significatif, c'est-à-dire influençant les intentions de vote. Ensuite, afin de modéliser la dynamique des intentions de vote propre à chaque parti, une variable temps et une variable temps au carré ont été testées. La première marque une évolution linéaire, c'est-à-dire une hausse ou une baisse constante et la seconde, une évolution quadratique, c'est-à-dire une descente suivie d'une montée ou inversement une montée suivie d'une descente.

Pour tester et modéliser l'effet de ces variables indépendantes sur la série de base de chaque parti, il faut utiliser l'analyse de régression. Cependant, étant donné que les séries chronologiques violent l'une des exigences de la régression ordinaire, soit l'indépendance des observations, il n'est pas possible d'utiliser ce type d'analyse. Si on le faisait, le coefficient de corrélation multiple et l'estimé de l'erreur-type seraient surestimés. Afin de tenir compte de la nature des données, il est nécessaire d'utiliser un autre type d'analyse de régression, l'analyse de séries chronologiques ARIMA. Trois analyses ARIMA ont donc été réalisées afin de modéliser les intentions de vote des résidents du Québec pour chacun des principaux partis fédéraux: le Parti libéral du Canada (PLQ), le Parti progressiste conservateur (PC) et le Bloc québécois (BQ).

3.2.2 Mesures de la mauvaise estimation des intentions de vote

Ces analyses ont permis d'élaborer des modèles de l'évolution des intentions de vote pour chaque parti tout au long de la campagne. Ce sont ces modèles qui constituent les bases de comparaison nécessaires au calcul des différents indicateurs de la mauvaise estimation: les erreurs absolues, c'est-à-dire les valeurs absolues des écarts entre les résultats des sondages et l'intention de

vote estimée, et les erreurs relatives, c'est-à-dire les écarts propres entre les résultats des sondages et l'intention de vote estimée (en tenant compte de la direction des erreurs). Ainsi, le premier indicateur permet de voir si des pratiques méthodologiques entraînent une moins bonne estimation des intentions de vote en général et le deuxième indicateur, si des pratiques sont associées à une surestimation ou à une sous-estimation des intentions de vote pour certains partis.

3.3 Classification des sondages

Les sondages ont été catégorisés, sur la base de la documentation consultée³¹, comme ayant un échantillon fortement ou faiblement maximisé. Un sondage ayant un échantillon faiblement maximisé est réalisé avec des pratiques ayant tendance à faire augmenter la non-réponse: l'absence de prise de rendez-vous, l'utilisation de quotas et la substitution de numéros de téléphone par exemple.

Dans le but de valider cette classification a priori, des analyses en classes hiérarchiques ont été réalisées. Ce type d'analyse permet d'identifier des groupes de sondages³² relativement homogènes en fonction des pratiques méthodologiques utilisées. Pour les fins de cette analyse, quatre variables ont été utilisées: le taux de réponse, la substitution de numéros de téléphone, l'utilisation de quotas et la récupération des refus. La méthode de Ward, dont l'objectif est de minimiser la variabilité intra-classe (Rapkin et Luke, 1993), a été employée pour effectuer le regroupement des sondages en

³¹ Les critères utilisés sont le taux de réponse, le pourcentage de non réponse attribuable aux non-contacts et aux refus, l'utilisation de quotas, le nombre d'appels faits pour joindre le ménage et la substitution de numéros de téléphone.

³² Dans la pratique, des firmes puisque celles-ci utilisent sensiblement les mêmes pratiques d'un sondage à l'autre.

fonction de la similitude des pratiques. La carré de la distance euclidienne a servi de mesure de distance entre les cas et les variables ont été standardisées afin de normaliser leur échelle³³. Les taux de réponse pour deux sondages étaient manquants. Tel que proposé par Rapkin et Luke (1993), la moyenne des taux de réponse leur a été imputée. Afin de vérifier si cette imputation avait un effet sur la classification, une analyse a été réalisée sans ces deux sondages. Cette procédure n'a pas entraîné une classification différente et donc, ces deux sondages ont été gardés dans l'analyse.

Après analyse, deux regroupements apparaissaient possibles: l'un comprenant trois classes et l'autre, deux. Étant donné que deux des trois classes étaient relativement similaires et que le nombre de sondages dans l'analyse est plutôt restreint, le regroupement comprenant deux classes apparaissait conceptuellement et empiriquement préférable. La première classe regroupent les sondages (12) ayant un échantillon fortement maximisé et la seconde, les sondages (5) ayant un échantillon faiblement maximisé.

3.4 Analyses des relations

De façon opérationnelle, l'hypothèse principale de la présente recherche se formule ainsi: les sondages ayant un échantillon faiblement maximisé sont associés à une moins bonne estimation des intentions de vote, soit des erreurs absolues ou relatives significativement plus importantes. Pour tester cette hypothèse, des analyses de variance ont été réalisées. Les erreurs absolues et les erreurs relatives constituent les variables dépendantes et la typologie de la maximisation de l'échantillon, la

³³ Cette normalisation était nécessaire étant donné les très grandes différences entre les échelles de mesure.

variable indépendante. Étant donné le faible nombre de cas, le seuil de significativité a été fixé à 0,10. Des analyses de covariance ont été réalisées afin de contrôler pour la taille des échantillons, cette variable permettant de contrôler, dans une certaine mesure, pour l'erreur d'échantillonnage des différents sondages. Ces analyses de covariance ont été réalisées de façon conservatrice. Il y a eu attribution hiérarchique des effets, c'est-à-dire que la taille de l'échantillon, la variable contrôle, a été entrée en priorité dans l'équation. Ainsi, elle s'est vu assigné, en plus de sa variance propre, toute celle qu'elle partageait avec la maximisation de l'échantillon, la variable indépendante. Ce faisant, il ne restait à la maximisation que sa variance propre. De plus, toutes ces analyses ont été réalisées en excluant les résultats d'un sondage réalisé par la firme Environnics (publié dans La Presse du 17 mai 1997) qui, par rapport aux autres, étaient largement déviants.

Pratiques méthodologiques et estimation des intentions de vote

par:
Sébastien Vachon, Claire Durand et André Blais,
Université de Montréal

Article soumis à
Cahiers canadiens de sociologie

© Vachon, S., Durand, C. et Blais, A., 1998

4 Pratiques méthodologiques et estimation des intentions de vote

Les sondages constituent un mode d'enquête largement utilisé par les pouvoirs publics et parapublics, les médias, les banques, les entreprises privées, etc. Parmi tous les sondages produits, les plus visibles sont, sans nul doute, les sondages pré-électoraux. De fait, on peut constater que leur nombre, que ce soit au Canada ou l'étranger, ne cesse d'augmenter (Lachapelle, 1991). C'est ainsi que, aujourd'hui plus que jamais, la qualité et la fiabilité de la méthodologie de sondage prennent toute leur importance. De fait, dans la mesure où les sondages peuvent affecter la perception que les électeurs se font de la force ou de faiblesse des différents partis et dans la mesure où cette perception peut affecter le vote (Johnston, Blais, Brady et Crête, 1992), il devient impératif d'estimer la fiabilité de ces sondages.

Cet article vise à mettre en lumière l'effet des pratiques méthodologiques, reliées à la réalisation des sondages, sur l'estimation des intentions de vote. L'article porte sur les sondages mesurant les intentions de vote des résidents du Québec pendant la campagne électorale fédérale de 1997.

4.1 Contexte

D'un point de vue théorique, pour s'appuyer sur la théorie des probabilités et pouvoir ainsi mesurer l'erreur d'échantillonnage, il est nécessaire que la génération de l'échantillon se fasse au hasard et que chaque membre de la population ait une chance connue d'être sélectionné. Si, en théorie, la meilleure façon de procéder à cette sélection semble relativement claire, dans la pratique, les choses ne sont pas si simples. En effet, il n'y a pas une façon de faire standardisée et codifiée, mais

une pléiade de pratiques méthodologiques différentes (Durand , Vachon et Tanguay, 1998; Taylor, 1994, 1997; Voss, Gelman et King, 1995).

Certaines pratiques méthodologiques peuvent affecter la représentativité de l'échantillon et à échéance, la qualité des estimés. De fait, si les répondants moins coopératifs (Chevalier et Durand, 1997; Groves et Lyberg, 1988; O'Neil, 1979; Triplett, 1997; Triplett et al., 1997) et ceux plus difficiles à joindre (Chevalier et Durand, 1997; Harris, 1988 dans Cantril, 1991; Gallup, 1988 dans Cantril, 1991; Sebold, 1988; Traugott, 1987; Triplett, 1997) ont un profil socio-démographique, des attitudes et des comportements typés, les pratiques méthodologiques qui entraînent leur sous-représentation risquent de causer des biais dans les estimés. Afin d'obtenir un échantillon représentatif et des résultats fiables, il est préférable d'utiliser des pratiques méthodologiques permettant de réduire le plus possible la non-réponse, qu'elle soit liée à des problèmes de coopération ou à la difficulté à joindre certains groupes de la population.

La réalisation d'un sondage demande un nombre important d'opérations. La génération de l'échantillon constitue la première étape. Il est nécessaire d'avoir une base de sondage qui comprend tous les membres de la population à l'étude. Dans la pratique, pour un sondage téléphonique, peu importe le mode de génération utilisé, seules les personnes demeurant dans un ménage privé ayant le téléphone ont une chance d'être sélectionnées. L'annuaire téléphonique, qu'il soit sur papier ou sur cd-rom, permet le recours à des méthodes simples de sélection. L'utilisation de cette base a comme avantage de maximiser le nombre de ménages valides dans l'échantillon de départ. Cependant, elle n'inclut pas les numéros de téléphone non inscrits (numéros confidentiels, récents, déménagements,

etc.), soit environ 25% des ménages privés (Taylor, 1997; Tremblay, 1981). Pour pallier à ce problème, il est possible d'ajouter un nombre aux numéros de téléphone générés à partir de l'annuaire³⁴. Cette méthode permet d'accéder aux numéros non publiés et de diminuer le nombre de numéros non valides (Landon et Banks, 1977). Cependant, pour que cette méthode soit efficace, il faudrait que les numéros non inscrits dans l'annuaire se distribuent de façon aléatoire dans les différents échanges (Cantril, 1991; Lepkowski, 1988 dans Lavrakas, 1993). Par rapport à la génération aléatoire de numéros de téléphone (GANT), «[...] cette méthode ne donne pas une chance égale à tous les numéros attribués d'être choisis. De plus, de nouveaux échanges mis en service après la publication de l'annuaire ne pourront pas être sélectionnés»³⁵ (Landon et Banks, 1977:297). La GANT constitue un groupe de techniques d'échantillonnage probabiliste. Schématiquement, elle consiste à générer au hasard des numéros de téléphone, en fonction des informations disponibles sur les échanges et les préfixes, les 3 ou 4 derniers chiffres des numéros, parmi tous ceux susceptibles d'être attribués. Pour Taylor (1994: 217), l'utilisation de cette méthode, malgré les coûts supplémentaires qu'elle entraîne, est une nécessité.

Dans un sondage téléphonique, un échantillon à deux degrés doit être utilisé: la sélection ou la génération de numéros de téléphone permet de sélectionner un ménage; la sélection dans le ménage permet de choisir un répondant. Parmi les différents modes de sélection, les plus utilisés au Canada sont la sélection en fonction de la date de naissance, les quotas basés sur le sexe ou l'âge et la grille

³⁴ On ajoute, par exemple, 3 à tous les numéros de sorte que le 354-3590 (numéro sélectionné dans l'annuaire) devient le 354-3593.

³⁵ «[...] the design does not give all working numbers a chance to be included. For instance, new sets of numbers placed in service after directory publication will not be sampled».

de Kish (Durand et al., 1998). La grille de Kish, utilisée par 47% des firmes québécoises et par aucune firme du Canada anglais (Durand, Vachon et Tanguay, 1998), est la plus exigeante en temps et ferait augmenter les risques de refus de la personne contactée, car elle exige l'énumération complète de tous les membres du ménage par leur âge, demande qui peut être perçue comme intrusive (Cantril, 1991; Groves et Lyberg, 1988; Lavrakas, 1993). Cependant, elle a l'avantage d'être strictement probabiliste, c'est-à-dire que chaque membre du ménage a une chance connue d'être sélectionné.

La sélection en fonction de la date d'anniversaire la plus rapprochée, à venir ou passée, est employée par les deux tiers des firmes de sondage canadiennes³⁶ (Durand, Vachon et Tanguay, 1998). Elle est rapide, peu intrusive et permet d'obtenir un bon taux de réponse (Salmon et Nichols, 1983; Lavrakas, 1993). Cependant, cette méthode augmente les risques d'auto-sélection, ce qui a pour effet de contribuer à la sous-représentation des personnes moins coopératives ou plus difficiles à joindre. De fait, selon Lavrakas et al. (1993 dans Lavrakas, 1993), dans 20%-25% des ménages, il y a une autre personne dont la date d'anniversaire est plus récente que celle du répondant. Autrement dit, une fois sur quatre, la mauvaise personne est sélectionnée.

Une autre façon de faire très utilisée par les firmes privées consiste à mener des entrevues auprès des gens répondant à certains critères: sexe, âge, région, langue, etc. «Il n'y a pas alors de sélection dans le ménage jusqu'à ce que certains des quotas soient remplis; la sélection se fait ensuite en recherchant des personnes dont les caractéristiques socio-démographiques permettent de remplir les

³⁶ La sélection en fonction de la date d'anniversaire est utilisée, au moins à l'occasion, par 60% des firmes québécoises et par 70% des firmes du Canada anglais (Durand, Vachon et Tanguay, 1998).

quotas visés» (Durand, Vachon et Tanguay, 1998: 5). Selon Cantril (1991), le pur «quota sampling», utilisé, au moins à l'occasion, par 82% des firmes du Canada anglais et par aucune firme québécoise³⁷ (Durand, Vachon et Tanguay, 1998), entraîne une surreprésentation des personnes plus coopératives. L'utilisation des quotas a, depuis les débuts des sondages, fait l'objet de débats (Blondiaux, 1991; Dreesbeke, Fichet et Tassi, 1987; March et Scarbrough, 1990; Stephenson, 1979). Sur le plan théorique, l'utilisation des quotas ne se justifie pas. Sur le plan pratique, les avis sont plus partagés. Des auteurs (Taylor, 1997; Bradburn, 1992) affirment que l'utilisation des quotas permet, dans bien des cas, d'obtenir des résultats plus fiables tandis que d'autres (Stephenson, 1979; March et Scarbrough, 1990) notent certains biais sur le plan de la représentativité de l'échantillon. Il demeure que le principal problème, outre l'absence de base théorique, est l'absence de critères permettant de juger de la qualité de l'échantillon.

Une fois le devis de sondage finalisé, d'autres décisions doivent être prises. Il faut planifier les moyens à prendre pour joindre tous les ménages et les individus sélectionnés. Ces moyens ont trait à la prise de rendez-vous, les tentatives de récupération des refus, le nombre et la variété de tentatives de contact, la durée de la période de collecte, etc. La prise de rendez-vous, devant un refus potentiel ou en l'absence de la personne sélectionnée, permet de faire augmenter le taux de réponse, surtout quand le questionnaire est long (Collins, Wilson et Blackshaw, 1988). Cette pratique permet notamment de faire diminuer le nombre de refus circonstanciels: personnes occupées, fatiguées, etc (Cantril, 1991).

³⁷ Ces informations sont basées sur l'auto-déclaration des firmes.

La récupération des refus vise à faire diminuer les biais introduits par la sous-représentation des personnes moins coopératives. S'il y a récupération, les refus du ménage et les refus du répondant devraient se diviser en deux parties égales (Chevalier et Durand, 1997; Collins, Wilson et Blackshaw, 1988). Cette pratique entraîne une augmentation significative du taux de réponse (Chevalier et Durand, 1997; O'Neil, 1979; Triplett, 1997) et habituellement, redresse les paramètres démographiques de l'échantillon (Chevalier et Durand, 1997; Groves et Lyberg, 1988; O'Neil, 1979; Triplett et al., 1996). Le nombre de tentatives de contact affecte aussi, selon plusieurs chercheurs, le taux de réponse et la représentativité de l'échantillon (Cantril, 1991; O'Neil, 1979; Sebold, 1988; Traugott et al., 1987; Traugott, 1987; Triplett, 1997; Triplett et al., 1996). Il faut au moins quatre tentatives de contact réparties à divers moments de la journée et de la semaine (Traugott, 1987) et jusqu'à sept ou huit (Chevalier et Durand, 1997) pour s'assurer de réduire significativement les différences entre personnes non rejointes et celles qui l'ont été (Durand, Vachon et Tanguay, 1998).

La durée de la période de collecte peut donner, dans certains cas, des indices quant aux efforts mis pour obtenir un taux de réponse acceptable. Cependant, la durée de la période de collecte n'est pas en soi garante des efforts mis pour joindre les personnes moins coopératives ou plus difficiles à joindre. Plusieurs autres facteurs entrent en ligne de compte: les pratiques de récupération des refus, le nombre de tentatives de contact, le nombre de lignes affectées à l'étude, etc. «La durée de la période de collecte des données est d'une importance secondaire, la conversion des refus et la réaffectation occupent une place unique, indépendamment de la durée de la période de collecte des données»³⁸ (De

³⁸ «The length of the data collection period seems of secondary importance; refusal conversion and reassignment take place independent of the length of the data collection period».

Heer, 1996: 10). La substitution de numéros de téléphone consiste à mettre sur le terrain un nouveau numéro quand il est impossible, après un certain nombre d'appels, de joindre le ménage ou quand il y a refus. Cette pratique n'aide pas à améliorer la représentativité de l'échantillon, car «elle tend à remplacer les non-répondants par des personnes qui ressemblent plus aux répondants qu'aux non-répondants»³⁹ (Smith, 1983: 53). Le biais causé par la non-réponse ne diminue pas (Smith, 1983; Groves, 1989). Selon Dillman (1978), cette procédure doit, autant que possible, être évitée.

En résumé, il apparaît donc que certaines pratiques peuvent avoir des effets néfastes sur la représentativité de l'échantillon. Pour corriger ces biais, et pour tenir compte des fractions de sélection et des taux de réponse différenciés selon les régions, plusieurs firmes font du redressement, c'est-à-dire qu'elles pondèrent les résultats de leur sondage en fonction de la répartition de certaines variables démographiques (sexe, région et langue habituellement). Cette procédure soulève plusieurs questions: les données de référence utilisées remontent, dans certains cas, à 1991; le fait de redresser en fonction de certaines variables peut avoir des effets inattendus sur la distribution d'autres variables; et finalement, par cette pratique, on se trouve à attribuer aux non-répondants les mêmes attitudes et comportements qu'aux répondants ayant un profil socio-démographique similaire.

Au Canada, comme dans les autres pays industrialisés, on constate une absence de consensus sur tous les aspects de la méthodologie de sondage (Taylor, 1994, 1997; Durand et al., 1998; Voss, Gelman et King, 1995). Les pratiques des différentes firmes forment des structures complexes et

³⁹ « [...] it tends to replace nonrespondents with people who resemble respondents rather than nonrespondents».

différenciées. Il apparaît évident que les différentes pratiques méthodologiques ne s'équivalent pas, certaines permettent d'obtenir un échantillon plus représentatif de la population à l'étude et par le fait même, des résultats plus fiables. Des études ont été menées afin de mesurer l'impact de certaines pratiques sur la qualité des estimés: Traugott (1987), la firme Harris (1988) et la firme Gallup (1988) ont ainsi démontré qu'un faible nombre de tentatives de contact entraîne un biais en faveur des Démocrates, les répondants plus difficiles à joindre, c'est-à-dire les jeunes hommes célibataires, scolarisés et demeurant en milieu urbain (Chevalier et Durand, 1997; Sebold, 1988; Traugott, 1987; Triplett, 1997), étant davantage républicains. L'étude de Lau (1994), portant sur les élections américaines de 1992, montre qu'un faible nombre de jours de terrain et des opérations de collecte uniquement les soirs de semaine entraînent une moins bonne estimation des intentions de vote en général; à l'inverse, l'utilisation des sondages à panel-roulant⁴⁰ produit des résultats plus fiables. Bref, à la lumière de la documentation consultée, il apparaît que les pratiques méthodologiques moins strictes sont associées à une moins bonne estimation des intentions de vote. L'absence de prise de rendez-vous, l'utilisation des quotas, la substitution de numéros de téléphone, etc. sont quelques unes des pratiques qui risquent de faire augmenter les biais dus à la non-réponse. Presque toutes les recherches analysant la méthodologie des sondages pré-électorales ont été réalisées aux États-Unis. Ainsi, par l'examen de la situation canadienne, il sera possible de réexaminer les conclusions de certaines recherches américaines dans un contexte social différent.

⁴⁰ Ces sondages sont réalisés sur une période plus ou moins longue (par exemple: une campagne électorale). Un faible nombre d'appels sont réalisés chaque jour mais en calculant des moyennes mobiles sur trois, quatre ou cinq jours, il est possible d'obtenir des estimés quotidiens comparables à ceux des sondages ponctuels.

4.2 Méthodologie

4.2.1 Données

Les données utilisées pour la recherche proviennent des sondages et des strates de sondages, c'est-à-dire de portions de sondages pan-canadiens, mesurant les intentions de vote des résidents du Québec pendant la campagne électorale fédérale de juin 1997. Durant la campagne, soit du 2 au 28 mai, les résultats de 10 sondages réalisés au Québec et de 15 sondages effectués dans l'ensemble du Canada ont été publiés dans les médias. Dans le cadre de la recherche, les 10 réalisés au Québec et 7 des 15 sondages pan-canadiens⁴¹ sont à l'étude (voir Annexe 3). La taille moyenne des échantillons pour les sondages est de 1000 et celle des strates, de 600. Les informations portant sur ces sondages ont été recueillies par le biais de questionnaires aux firmes de sondage (voir Annexe 2), dans les principaux quotidiens (La Presse, The Gazette, The Globe and Mail) et sur internet (Chroniques de Sibérie, Élection Canada, sites des firmes de sondage).

4.2.2 *Modèle d'analyse*

Afin de déterminer quelles pratiques méthodologiques sont reliées à une moins bonne estimation des intentions de vote, il est nécessaire d'avoir une base de comparaison. Il faut donc estimer un modèle de l'évolution de l'intention de vote tout au long de la campagne électorale.

Dans un premier temps, il s'agit d'élaborer une série de base, c'est-à-dire qu'il faut répartir les

⁴¹ Les sondages exclus de l'analyse ont été réalisés par la firme américaine Zogby et par Strategic Council. Dans le paysage canadien des firmes de sondage, elles sont de moindre importance. De plus, il a été impossible de recueillir toutes les informations nécessaires à leur examen.

résultats des sondages de façon à obtenir une série quotidienne des intentions de vote. Les résultats de chaque sondage sont répartis sur chacune des journées où il y a eu collecte d'information. Ainsi, pour un sondage réalisé entre le 19 et le 23 mai, les résultats sont affectés aux cinq jours où il y a eu collecte. Pour chaque jour, les résultats des sondages ont été pondérés par la taille de leur échantillon, ce qui permet de tenir compte de l'importance relative de chaque sondage. Des moyennes pondérées ont ainsi été calculées pour chaque jour de la campagne. Ce sont ces moyennes qui constituent la série brute des intentions de vote.

Dans un deuxième temps, il faut modéliser la dynamique des intentions de vote au fil de la campagne. Si l'on veut connaître l'effet net d'une variable indépendante, d'un événement par exemple, il faut contrôler pour l'auto-corrélation des erreurs de la variable dépendante, l'erreur de chaque donnée étant corrélée avec la ou les erreurs précédentes. Afin de répondre à cette exigence, l'analyse de série chronologique ARIMA, développée à l'origine par Box et Jenkins (1976), a été utilisée. Ce type d'analyse permet de contrôler l'auto-corrélation des erreurs, de mesurer et de modéliser efficacement l'effet net de variables indépendantes sur la série.

Dans un troisième temps, prenant ces modèles de l'évolution des intentions de vote pour chaque parti comme bases de comparaison, il est possible d'élaborer, pour chaque parti, différents indicateurs de la mauvaise estimation des intentions de vote: la somme des écarts relatifs (erreurs relatives) et celle des écarts absolus (erreurs absolues) entre les résultats des sondages et ceux issus des modèles. Les erreurs relatives permettent de voir si des pratiques méthodologiques entraînent une sur ou une sous-estimation des intentions de vote pour certains partis et les erreurs absolues si des pratiques sont

associées à une moins bonne estimation des intentions de vote en général.

Afin de mesurer les pratiques, une classification des sondages a été réalisée en fonction de la plus ou moins grande maximisation de l'échantillon, c'est-à-dire des efforts mis pour obtenir une entrevue auprès des personnes moins coopératives et plus difficiles à joindre. La maximisation peut être opérationnalisée, entre autre, par le taux de réponse⁴². Pour obtenir un bon taux de réponse, différentes pratiques méthodologiques doivent être employées: nombre et répartition sur les différentes périodes de la semaine (jour semaine, soir semaine et fin de semaine) des tentatives de contact des ménages, récupération des premiers refus de collaborer, prise de rendez-vous, etc. Des analyses de classes hiérarchiques ont été réalisées. Ce type d'analyse permet une «identification empirique des groupes de cas ayant des profils similaires sur un certain nombre de variables»⁴³ (Rapkin et Luke, 1993: 254).

Enfin, l'impact de la maximisation des échantillons sur l'estimation des intentions de vote est examiné au moyen d'analyses de variance et de covariance. Ce type d'analyse permet d'étudier, pour chaque variable testée, les différences de moyennes et d'écart-type. Le test de Levene sur l'homogénéité de la variance, permettant d'établir la plus ou moins grande robustesse des tests F sur les moyennes, est également calculé. De plus, grâce aux analyses de covariance avec attribution hiérarchique des effets, il est possible de contrôler pour l'effet d'une covariable et ainsi, de mesurer

⁴² Bien qu'il n'y ait pas unanimité sur la manière de calculer ce taux, on s'entend généralement pour dire qu'il s'agit du rapport entre le nombre d'entrevues complétées et le nombre de numéros dans l'échantillon (valides et éligibles).

⁴³ «[...] empirically identifying groups of cases with similar profiles across variables».

l'effet net de la variable indépendante.

4.3 Les résultats

4.3.1 Mesures

4.3.1.1 Intentions de vote

Pour modéliser la dynamique des intentions de vote pour chaque parti tout au long de la campagne, il a été nécessaire de tenir compte du fait qu'il y a auto-corrélation au premier degré entre les données. Afin de dégager les grandes tendances, des fonctions linéaires et quadratiques ont été testées et modélisées. De tous les événements de la campagne, seul le premier débat en français a eu un effet significatif sur l'intention de vote au Québec. Afin de vérifier la justesse des modèles élaborés, des projections⁴⁴ ont été réalisées jusqu'au jour de l'élection. Finalement, des tests d'auto-corrélation et d'auto-corrélation partielle des erreurs ont permis de vérifier si le processus d'auto-corrélation au premier degré était bien contrôlé par l'analyse de série chronologique ARIMA.

Le modèle des intentions de vote pour le Bloc québécois, présenté à la figure 1, montre une fonction linéaire non-significative, le parti terminant la campagne où il l'avait commencé, et une fonction quadratique en U significative, le Bloc québécois perdant des points de pourcentage en milieu de terrain, principalement à cause du premier débat en français (3,5 p. cent) et récupérant graduellement le terrain perdu par la suite. Finalement, la projection des intentions de vote jusqu'au

⁴⁴ Lors de l'élection canadienne de 1997, une loi fédérale interdisait la publication de sondages pré-électorales 48 heures avant le scrutin. Ainsi, entre le 28 juin, dernière journée où il y a eu collecte d'informations, et le 2 juin, jour de l'élection, des projections ont été réalisées. En 1998, contestée devant la Cour suprême du Canada, cette loi a finalement été abrogée.

jour de l'élection démontre, étant donné la justesse de celle-ci (1,3 point de pourcentage près), la fiabilité de cette modélisation.

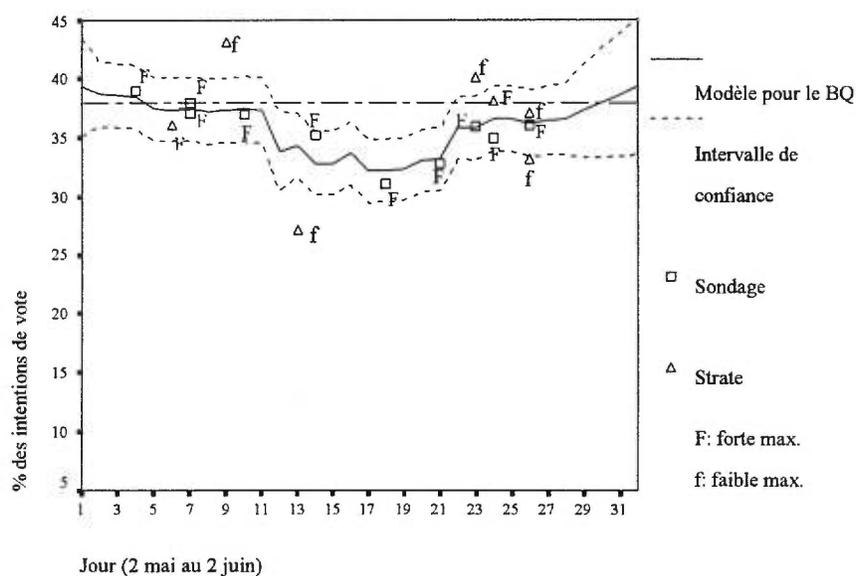


Tableau VIII. Analyse de série chronologique ARIMA pour le Bloc québécois

Variable	B	Erreur-type	Valeur de T	p
Auto-régression au premier degré	0,566	0,159	3,565	0
Premier débat en français	-3,52	1,355	-2,597	0
Fonction linéaire	-0,46	0,267	-1,738	0
Fonction quadratique	0,02	0,008	2,092	0
Constante	39,74	1,502	26,46	0
Erreur-type	1,25			
Variance résiduelle	1,562			

Pour le Parti Conservateur, la série chronologique, présentée à la figure 2, montre une fonction linéaire positive significative, une fonction quadratique en \cap et une hausse importante de 8,4 p. cent de l'intention de vote lors du premier débat en français. Du début au milieu de la campagne, le parti a connu une hausse importante de 17,8 p. cent, près de la moitié de cette hausse pouvant être attribuée au débat, mais progressivement, cet «effet Charest» s'est estompé, les gains du Parti Conservateur se limitant finalement à une hausse totale de cinq points de pourcentage. Il est à noter finalement la grande justesse de la prédiction (0,1 point de pourcentage près).

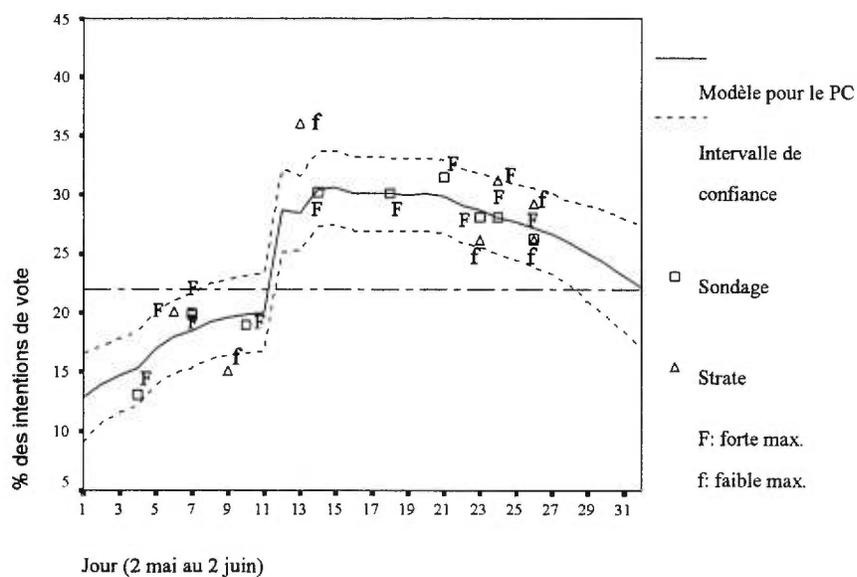


Figure 2. Modèle pour le PC, projection, sondages et strates, maximisation de l'échantillon

Tableau IX. Analyse de série chronologique ARIMA pour le Parti conservateur

Variable	B	Erreur- type	Valeur de T	p
Auto-régression au premier degré	0,276	0,174	1,585	0,1
Premier débat en français	8,387	1,388	6,043	0
Fonction linéaire	1,184	0,237	4	0
Fonction quadratique	0	0,007	-4,968	0
Constante	11,65	1,271	9,166	0
Erreur-type	1,443			
Variance résiduelle	2,083			

Le modèle des intentions de vote pour le Parti libéral (figure 3) montre une fonction linéaire significative, le parti ayant, au total, perdu près de cinq points de pourcentage entre le début et la fin de la campagne, et une fonction quadratique en \cup significative, le Parti libéral du Canada perdant des points au début de la campagne et de façon légèrement plus prononcée lors du premier débat en français (2 p. cent) et récupérant graduellement une partie du terrain perdu par la suite. Finalement, comme pour les autres partis, la projection des intentions de vote jusqu'au jour de l'élection est juste (1,1 point de pourcentage près).

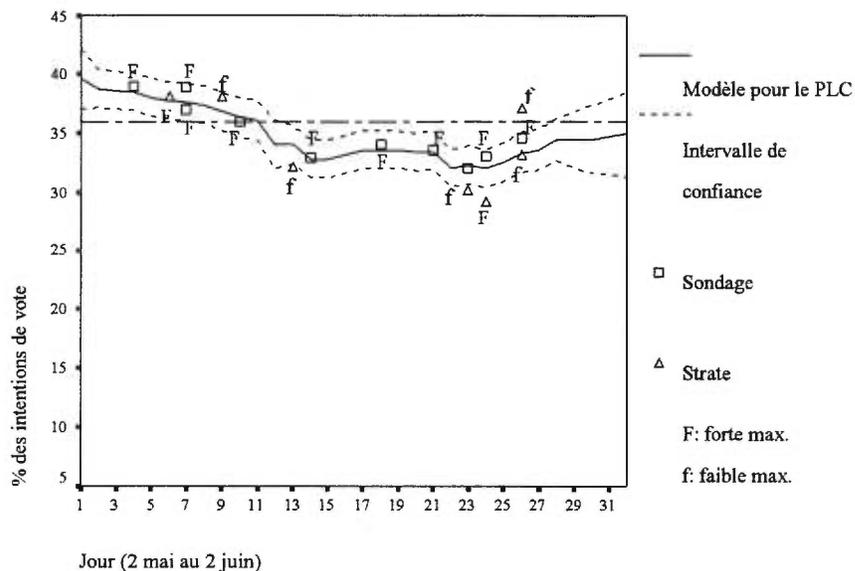


Tableau X. Analyse de série chronologique Arima pour le Parti libéral du Canada

Variable	B	Erreur- type	Valeur de T	p
Auto-régression au premier degré	0,608	0,182	3,34	0
Premier débat en français	-1,98	0,805	-2,46	0
Fonction linéaire	-0,51	0,165	-3,065	0
Fonction quadratique	0,126	0,005	2,398	0
Constante	40,1	0,938	42,76	0
Erreur-type	0,736			
Variance résiduelle	0,542			

4.3.1.2 Caractérisation des pratiques

Grâce à une analyse en classes hiérarchiques (figure 4), il a été possible de valider une première classification au jugé des pratiques méthodologiques. La méthode de Ward, dont l'objectif est de

minimiser la variabilité intra-classe (Rapkin et Luke, 1993), a été employée pour effectuer l'agrégation des cas. Le carré de la distance euclidienne a servi de mesure de distance entre les cas et étant donné la nécessité de normaliser les échelles des variables⁴⁵, les «scores» Z ont été utilisés. Le choix des variables s'est fait sur la base de premières analyses univariées, les variables choisies ayant un effet significatif sur l'un ou l'autre des indicateurs de la mauvaise estimation. En classant les différents sondages en fonction de la plus ou moins grande maximisation de leur échantillon, il apparaît possible de dégager deux groupes distincts: les sondages à échantillon fortement maximisé et ceux à échantillon faiblement maximisé. Pour les premiers, il y a, dans la totalité des cas, récupération des refus et aucune substitution de numéros de téléphone. Pour la moitié de ces sondages, les quotas sont employés.

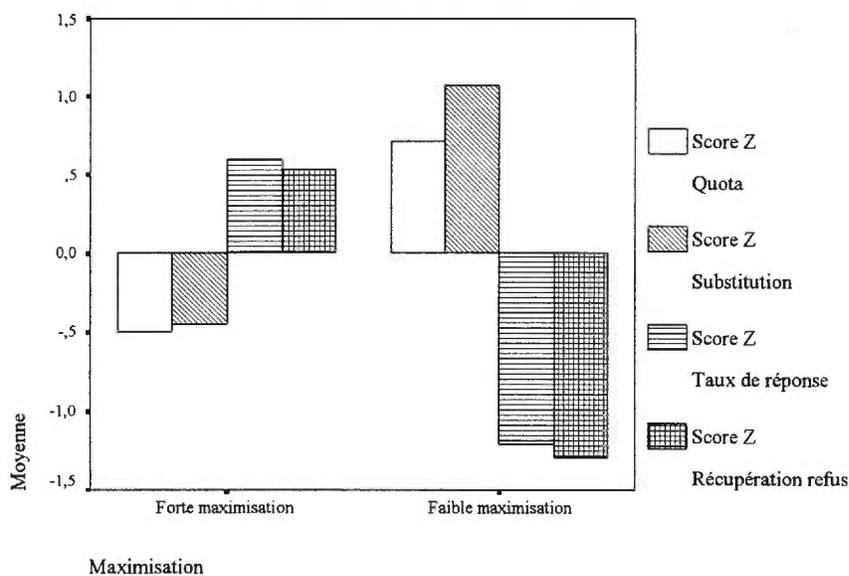


Figure 4. Analyses en classes hiérarchiques - catégorisation des pratiques méthodologiques

⁴⁵ Cette normalisation était nécessaire étant donné les très grandes différences entre les échelles de mesure.

Le taux de réponse moyen est de 60%. Ainsi, on peut constater que les pratiques méthodologiques employées, qui sont relativement strictes, permettent d'obtenir un bon taux de réponse. Pour les seconds, il y a, dans trois cas sur quatre, aucune récupération des refus et substitution de numéros de téléphone. Les quotas sont employés dans tous les cas et le taux de réponse moyen est de 30%. Bref, il y a concordance entre les pratiques méthodologiques employées et le taux de réponse obtenu.

4.3.2 Effet de la maximisation de l'échantillon sur l'estimation des intentions de vote

Il est maintenant possible d'examiner l'effet de la maximisation sur les différents indicateurs de la mauvaise estimation. Pour ce faire, des analyses de variance et des analyses de covariance avec attribution hiérarchique des effets ont été réalisées. Étant donné le faible nombre de cas, le seuil de significativité a été fixé à 0,1. Toutes les analyses ont été réalisées en excluant les résultats d'un sondage réalisé par la firme Environnics (publié dans La Presse du 17 mai 1997) qui, par rapport aux autres, étaient largement déviants⁴⁶. Ainsi, cette précaution permet d'être plus confiant de la robustesse des tests F sur les moyennes.

4.3.2.1 Erreurs relatives

Les erreurs relatives, c'est-à-dire la somme des écarts relatifs entre les résultats des sondages et les estimations des modèles, permettent d'examiner si la plus ou moins grande maximisation de l'échantillon entraîne une sur ou une sous-représentation des intentions de vote pour certains partis. Le tableau XI montre qu'une faible maximisation de l'échantillon est associée à une sous-estimation

⁴⁶ Ce sondage, ou pour être plus précis cette strate de sondage, comptent moins de 500 répondants dans son échantillon. De plus, son taux de réponse est très faible.

des intentions de vote pour le Parti conservateur ($F= 3,598$, $p= ,079$). Si l'on contrôle pour la taille de l'échantillon, cette association demeure significative ($F= 3,395$, $p= ,088$). Ainsi, après contrôle, la sous-estimation est de 2,6 p. cent. De plus, le test de Levene⁴⁷ n'est pas significatif, indiquant que les variances de groupe sont homogènes. Il apparaît que ces erreurs de prédiction ne sont pas expliquées par la taille des échantillons ($F= ,911$, $p= ,357$).

Tableau XII. Analyse de variance et de covariance mesurant l'effet de la maximisation de l'échantillon sur l'estimation des intentions de vote - erreurs relatives

Variable (n)	Maximisation de l'échantillon			Taille échantillon n F (p)	Maximisation n F (p)	Levene (p)
	Forte Moyenne (é.t.)	Faible Moyenne (é.t.)	Total Moyenne (é.t.)			
Erreurs rel. BQ (16)	-,02 (1,15)	1,87 (3,97)	,453 (2,201)		2,421 (,142)	14,068 (,002)
	-0,272	2,627	0,453	,677 (,425)	2,030 (,178)	9,384 (,008)
Erreurs rel. PC (16)	,354 (1,513)	-1,655 (2,705)	-,148 (1,987)		3,598 (,079)	6,097 (,212)
	0,658	-2,566	-0,148	,911 (,357)	3,395 (,088)	8,828 (,297)
Erreurs rel. PLC (16)	,053 (1,129)	,535 (2,503)	,174 (1,495)		,297 (,594)	1,714 (,058)
	0,4521	-0,661	0,174	,637 (,439)	,637 (,439)	1,174 (,403)

Il y a un effet net attribuable à la maximisation. Dans la réalisation de ces sondages, peu d'efforts ont été mis pour joindre les personnes moins coopératives et plus difficiles à joindre. Ces deux groupes, ayant des caractéristiques socio-démographiques, des attitudes et des comportements différents, sont sous-représentés. Pour le Bloc québécois et le Parti libéral du Canada, aucun effet significatif n'est

⁴⁷ Le test de Levene indique qu'il n'y a pas de différence significative entre les variances de ces deux catégories: forte et faible maximisation. Ce faisant, la robustesse du test F est assurée.

observé. Toutefois, pour le Bloc québécois, le test de Levene est significatif, indiquant des résultats moins constants pour les sondages à échantillon faiblement maximisé. Cependant, étant donné que le test F sur les moyennes n'indique pas de différence significative, il n'est pas possible de déterminer si cet effet est dû à la taille de l'échantillon (erreur d'échantillonnage) ou à la maximisation.

4.3.2.2 *Erreurs absolues*

Les erreurs absolues, c'est-à-dire la somme des écarts absolus entre les résultats des sondages et les estimations des modèles, permettent de voir s'il y a association entre la plus ou moins grande maximisation de l'échantillon et la qualité générale de l'estimation des intentions de vote. Pour le Bloc québécois (tableau XII), il apparaît qu'une faible maximisation de l'échantillon entraîne des erreurs de prédiction significativement plus importantes ($F= 15,521$, $p= ,001$). De plus, même en contrôlant pour la taille des échantillons, l'association demeure significative ($F= 4,058$, $p=,065$), soit 3,2 p. cent d'erreurs comparativement à 0,9 p. cent pour les sondages à échantillon fortement maximisé. Certes, il y a hétérogénéité des variances ($F= 8,470$, $p= ,010$), mais étant donné que le test F est assez fort ($F=4,058$, $p=,065$), il est possible d'affirmer la robustesse de la relation (voir Lindman, 1992).

Tableau XIII. Analyse de variance et de covariance mesurant l'effet de la maximisation de l'échantillon sur l'estimation des intentions de vote - erreurs absolues

Variable (n)	Maximisation de l'échantillon			Taille échantillon F (p)	Maximisation F (p)	Levene (p)
	Forte Moyenne (é.t.)	Faible Moyenne (é.t.)	Total Moyenne (é.t.)			
Erreurs abs. BQ (16)	,872 (.706)	3,48 (2,076)	1,524 (1,609)		15,521 (.001)	,998 (.027)
	0,96	3,214	1,524	10,694 (.006)	4,058 (.065)	8,470 (.010)
Erreurs abs. PC (16)	1,214 (.904)	2,565 (1,482)	1,552 (1,845)		4,919 (.044)	4,262 (.335)
	1,469	1,799	1,552	6,831 (.021)	,115 (.740)	,744 (.011)
Erreurs abs. PLC (16)	,748 (.816)	1,86 (1,427)	1,026 (1,069)		3,864 (.069)	2,570 (.131)
	0,788	1,742	1,026	2,654 (.127)	,989 (.338)	3,550 (.080)

Pour le Parti conservateur, la relation n'est plus significative quand on contrôle pour la taille de l'échantillon. De fait, il y a une forte relation entre la maximisation et la taille des échantillons ($F=23,569$, $p=,000$). Dans cette situation, il n'est pas possible de départager la part attribuable à la maximisation et celle attribuable à la taille des échantillons, ou plus précisément à l'erreur d'échantillonnage. Pour le Parti libéral du Canada, la relation est significative avant contrôle pour la taille des échantillons ($F=3,864$, $p=,069$). Contrairement aux deux autres partis, il n'y a pas, pour ce parti, de corrélation entre la taille des échantillons et les erreurs absolues ($R=-,399$, $p=,125$). Ainsi, il n'est pas pertinent de contrôler pour cette variable et seul le premier test F doit être interprété. Ainsi, il apparaît qu'une faible maximisation de l'échantillon entraîne une moins bonne estimation des intentions de vote pour ce parti, soit 1,86 p. cent comparativement à ,748 p. cent pour les sondages à échantillon fortement maximisé.

4.4 Conclusion

Pour le Parti libéral du Canada et pour le Bloc québécois, une faible maximisation de l'échantillon entraîne une moins bonne estimation des intentions de vote. Pour le Parti conservateur, elle est associée une sous-estimation des intentions de vote. Ces résultats s'expliquent, du moins en partie, par le peu d'effort qui a été mis pour réduire la non-réponse, qu'elle soit liée à des problèmes de coopération ou à la difficulté à joindre certains groupes de la population. Cependant, à ce moment-ci, il n'est pas possible de déterminer l'effet spécifique de chacun de ces deux aspects de la non-réponse sur la qualité de l'estimation des intentions de vote. D'autres études seront nécessaires. Plus précisément, il serait pertinent d'examiner des bases administratives de sondages pré-électorales afin d'étudier les caractéristiques socio-démographiques des personnes moins coopératives et plus difficiles à joindre. Ce faisant, il serait possible de connaître l'effet de la maximisation sur la représentativité de l'échantillon et sur la qualité des estimés.

Références

Blondiaux, Loïc

1991 "L'invention des sondages d'opinion: expériences, critiques et interrogations méthodologiques (1935-1950)." *Revue française de science politique*: 391-97.

Box, George E.P. et Gwilym M. Jenkins

1976 *Time series analysis: forecasting and control*. Oakland, CA: Holden-Day.

Bradburn, Norman M., et Seymour Sudman

1988 *Polls and surveys*. San Francisco: Jossey-Bass.

Buchanan, William

1986 "Elections predictions: An empirical assessment." *Public opinion quarterly* 50: 222-227.

Cantril, Albert H.

1991 *The opinion connection*. Washington, D.C.: Congressional Quarterly.

Chevalier, Serge et Claire Durand

1997 *Le taux de réponse dans les sondages téléphoniques, déterminants et conséquences*. 65^e Congrès de l'ACFAS, Montréal, Université du Québec à Trois-Rivières.

Collins, M., W. Sykes, P. Wilson, et N. Blackshaw

1988 Nonresponse: The UK experience." dans Robert M., Groves, Paul P. Biemer, Lars E. Lyberg, James T. Massey, W.L. Nicholls II, et J. Wasberg (dirs), *Telephone survey methodology*. New York: Wiley: 213-231.

De Heer, Wim

1996 *International response trends*. 4th International social science methodology, Colchester England, University of Essex.

Dillman, Don A.

1978 *Mail and telephone surveys*. New York: Wiley.

Droesbeke, Jean-Jacques, Bernard Fichet, et Philippe Tassi

1987 "La place des sondages en statistique." dans Jean-Jacques Droesbeke, Bernard Fichet et Philippe Tassi, *Les sondages*. Paris: Economica: 3-28

Durand, Claire, Sébastien Vachon, et Isabelle Tanguay

1997 "La gestion de la méthodologie de sondage au Québec et dans le reste du Canada, une question de sociétés distinctes?" *Recherches sociographiques* 39(2 et 3): 417-438.

Groves, Robert M.

1989 "Probing the causes of nonresponse and efforts to reduce nonresponse." dans Robert M. Groves, *Survey errors and survey costs*. New York: Wiley.

Groves, Robert M., Paul P. Biemer, Lars E. Lyberg, James T. Massey, William L. Nicholls II, et Joseph Waksberg

1988, eds. Telephone survey methodology. New York, Wiley.

Johnston, Richard, André Blais, Henry E. Brady et Jean Crête

1992 *Letting the people decide. Dynamics of a canadian election*. Montréal: McGill-Queen's University Press.

Lachapelle, Guy

1991 *Les sondages et les médias lors des élections au Canada. Le pouls de l'opinion*. Approvisionnement et Services Canada, 16, Ottawa: Dundurn Press Limited & Wilson & Lafleur Limitée.

Landon, E. Laird et Sharon K. Banks

1977 "Relative efficiency and bias of plus-one telephone sampling." *Journal of marketing research* 14: 294-99.

Lavrakas, Paul J.

1993 *Telephone survey methods: sampling, selection and supervision*. Applied social research methods series, 7, Newbury Park: Sage.

Lindman, Harold R.

1992 *Analysis of variance in experimental design*. New York, Springer-Verlag, 531 pages.

March, Catherine et Elinor Scarbrough

1990 "Testing nine hypotheses about quota sampling." *Journal of the market research society* 32(4): 485-506.

O'Neil, Michael J.

1979 "Estimating the non response bias due to refusals in telephone surveys." *Public opinion quarterly* 43(2): 218-32.

Rapkin, Bruce D. et Douglas A. Luke

1993 "Cluster analysis in community research: epistemology and practice." *American journal of community psychology* 21(2): 247-277.

Salmon, Charles, et John Spicer Nichols

1983 "The next-birthday method of respondent selection." *Public opinion quarterly* 47:270-76.

Sebold, Janice

- 1988 "Survey period length, unanswered numbers, and nonresponse in telephone surveys." dans Robert M. Groves, Paul P. Biemer, Lars E. Lyberg, James T. Massey, William L. Nicholls II, et Joseph Waksberg (dirs), *Telephone survey methodology*. New York: Wiley: 247-256.

Smith, Tom W.

- 1983 "The hidden 25 percent: An analysis of nonresponse on the 1980 general social survey." *Public opinion quarterly* 47(3):50-68.

Stephenson, C. Bruce

- 1979 "Probability sampling with quotas: An experiment." *Public opinion quarterly* 43(4): 477-96.

Taylor, Humphrey

- 1994 "Horses for courses: How survey firms in different countries measure public opinion with Very different methods." *Journal of the market research society* 13(3):211-19.

Taylor, Humphrey

- 1997 "The very different methods used to conduct telephone surveys of the public." *Journal of the market research society* 39(3):421-32.

Traugott, Michael W.

- 1987 "The importance of persistence in respondent selection for preelection surveys." *Public opinion quarterly* 51:48-57.

Traugott, Michael W., Robert M. Groves, et James M. Lepkowski

- 1987 "Using dual frame designs to reduce nonresponse in telephone surveys." *Public opinion quarterly* 51: 522-39.

Tremblay, Victor

- 1988 *La sélection dans le bottins téléphoniques: ampleur et conséquences de la non-inscription*. Montréal, Centre de sondage, Université de Montréal.

Triplett, Timothy

- 1997 *What is gained from additionnal call attempts and refusal conversion and what are the cost implication?* Maryland, Survey research center, University of Maryland.

Triplett, Timothy, Jhonny Blair, Teresa Hamilton, et Yun Chiao Kang

- 1996 *Initial cooperators vs converted refusers: Are there response behavior differences?*. Proceedings of the American association of public opinion research conference, Social statistics section.

Voss, D.Stephen, Andrew Gelman, et Gary King

1995 "The polls - A review. Preelection survey methodology: details from eight polling organizations, 1988 and 1992." *Public opinion quarterly* 59: 98-132.

5 Discussion

Grâce aux modèles des intentions de vote pour chaque parti, il est possible de juger de la fiabilité de tous les sondages d'une campagne électorale. D'autres études, notamment celles de Buchanan (1986) et de Crespi (1988), utilisent le vote effectif à l'élection comme base de comparaison. Dans ce cas, il est possible d'examiner la fiabilité que des tous derniers sondages de la campagne uniquement. Comment pourrait-on en effet utiliser le vote effectif comme base de comparaison pour mesurer la qualité des sondages réalisés 30 jours plus tôt? Si la recherche porte sur une élection particulière, il apparaît évident que cette méthode impose des limitations importantes. De fait, ce n'est qu'en évaluant la qualité des sondages de plusieurs élections qu'il est possible d'obtenir un échantillon valable. Crespi (1988) a recueilli des informations sur 430 élections et Buchanan (1986) sur 155.

Il y a de nombreuses façons d'élaborer les modèles des intentions de vote (Fox, Andersen et Dubonnet, 1997; Lau, 1994; Pinard, Bernier et Lemieux, 1997). Il ne semble pas y avoir de consensus sur la meilleure méthode à employer. De plus, il est malheureux de constater que ces auteurs sont très avares de précisions sur leur démarche. Ainsi, il est difficile de commenter dans le détail les différentes approches employées. Malgré tout, quelques remarques s'imposent. Les méthodes élaborées par Pinard, Bernier et Lemieux (1997) et Fox, Andersen et Dubonnet (1997) ne tiennent pas compte de la taille des échantillons, ce qui a pour effet de surestimer la fiabilité des strates de sondage et des sondages à petits échantillons. Lau (1994) effectue une pondération des résultats de sondage. Dans l'élaboration de sa série de base, après répartition des résultats de chaque sondage sur chacune des journées où il y a eu collecte d'information, il pondère les résultats par N/j , où N est la taille de l'échantillon et j le nombre de jours de terrain. Cette pondération pose problème puisque si la taille des échantillons est égale, plus un sondage est réalisé rapidement, plus son poids dans la série de base

sera grand. Sachant que la durée de la période de collecte donne, bien qu'imparfaitement, certains indices quant aux efforts mis pour obtenir un bon taux de réponse⁴⁸, cette pondération apparaît discutable. Bref, il apparaît préférable d'effectuer une simple pondération par la taille des échantillons, ce qui permet de tenir compte de l'erreur d'échantillonnage de chaque sondage.

Afin d'analyser des données qui varient au fil du temps, il faut tenir compte du fait que le postulat d'indépendance des résultats ne tient plus. Ainsi, il apparaît curieux de constater qu'aucune des recherches consultées (Pinard, Bernier et Lemieux, 1997; Fox, Andersen et Dubonnet, 1997; Lau, 1994), ne fournit d'informations à ce propos. Sachant qu'il est nécessaire, afin de connaître l'effet d'une variable indépendante sur la série, de tenir compte du fait que les mesures d'une série chronologique sont corrélées les unes avec les autres, il faut s'interroger sur la qualité des tests effectués. De fait, s'il n'y a pas eu de contrôle, «les variances estimées [sous-estiment] sérieusement [...] les variances réelles. De plus, les variances estimées sont extrêmement importantes pour mesurer les intervalles de confiance, tester les hypothèses et calculer les valeurs de t »⁴⁹ (Ostrom, 1990: 22).

Le nombre de sondages réalisés pendant la campagne fédérale canadienne de 1997 est relativement limité comparativement aux dernières élections américaines. La présente recherche analyse la fiabilité de dix-sept sondages tandis que celle de Lau (1994), qui examine les sondages de la campagne présidentielle aux États-Unis, en compte cinquante-six. De plus, de ces dix-sept

⁴⁸ Plus la durée est courte, plus le taux de réponse est bas.

⁴⁹ «[...] the estimated variances will seriously underestimate the true variances. Furthermore, the estimated variances are extremely important in constructing confidence intervals, testing hypotheses, and computing t-ratios.»

sondages, sept sont des sondages, c'est-à-dire des portions de sondages pan-canadiens, tous réalisés par des firmes du Canada anglais. Dans les analyses, afin de mesurer l'effet net attribuable à la maximisation, la taille des échantillons a servi de variable contrôle. Ce faisant un problème est rapidement apparu. Les sondages réalisés au Québec (échantillon moyen: 1000), en conformité avec les résultats de Durand, Vachon et Tanguay (1998) ont tendance à être faits de façon plus stricte que dans le reste du Canada (échantillon moyen des strates: 600). Plusieurs pratiques «non orthodoxes» entraînent une moins bonne estimation des intentions de vote mais étant donné leur forte corrélation avec la taille des échantillons, il apparaît difficile de départager la part de l'explication attribuable à la pratique elle-même de celle attribuable à la taille des échantillons ou plus précisément à l'erreur d'échantillonnage. Des analyses séparées pour les sondages et les strates de sondages auraient permis de régler ce problème, mais étant donné le faible nombre de cas disponibles, il était impossible d'emprunter cette voie.

6 Conclusion

Il apparaît qu'une faible maximisation de l'échantillon entraîne une moins bonne estimation des intentions de vote pour le Parti libéral du Canada et pour le Bloc québécois et une sous-estimation des intentions de vote pour le Parti conservateur. Ces résultats s'expliquent, du moins en partie, par le peu d'effort qui a été mis pour réduire la non-réponse, qu'elle soit liée à des problèmes de coopération ou à la difficulté à joindre certains groupes de la population. L'absence de prise de rendez-vous, l'utilisation des quotas, la substitution de numéros de téléphone, etc. sont quelques unes des pratiques qui risquent de faire augmenter les biais dus à la non-réponse. En ce sens, il apparaît que les pratiques méthodologiques ne s'équivalent pas, certaines permettent d'obtenir un échantillon plus représentatif de la population à l'étude et par le fait même, des résultats plus fiables. Au Canada, comme dans les autres pays industrialisés, on constate une absence de consensus sur tous les aspects de la méthodologie de sondage (Taylor, 1994, 1997; Durand, Vachon et Tanguay, 1998; Voss, Gelman et King, 1995), malgré que la recherche en méthodologie de sondage soit relativement bien développée.

«Plusieurs partagent [...] la conviction que les sondages exercent une influence sur les campagnes électorales, contribuent au vote stratégique, créent un effet d'entraînement et, surtout, influent sur les campagnes de financement des partis, sur le moral des militants et sur la couverture médiatique» (Lachapelle, 1991: xvii). Ainsi, la fiabilité de la méthodologie de sondage prend toute son importance. Le public est en droit de s'attendre à une information juste et de qualité. Dans cette perspective, il apparaît plus que jamais nécessaire d'évaluer et d'analyser les différentes pratiques méthodologiques. D'autres recherches devront être menées. Par exemple, dans la présente recherche, il apparaît que les différentes pratiques méthodologiques, loin de se distribuer aléatoirement, ont

plutôt tendance à se regrouper, à former des amalgames. Ce faisant, il n'a pas été possible de déterminer quelles pratiques spécifiques, sous-représentant les personnes plus difficiles à joindre et/ou moins coopératives et entraînent ces erreurs d'estimation. La maximisation de l'échantillon est un indicateur général. Il serait pertinent d'examiner des bases administratives de sondages pré-électorales afin d'étudier les attitudes politiques et le comportement électoral des personnes moins coopératives et plus difficiles à joindre, ces deux aspects de la non-réponse pouvant être étudiés de façon distincte.

Au cours des prochaines élections, au Québec comme dans les autres provinces, l'exercice sera repris: élection municipale de Montréal à l'automne 1998, élection québécoise qui aura lieu au plus tard à l'automne 1999 et élection ontarienne qui aura lieu en 1999 ou 2000. D'une élection à l'autre, les indicateurs de la mauvaise estimation pourront être cumulés. Ainsi, il sera possible d'accroître progressivement la taille de l'échantillon disponible et ainsi, de permettre des analyses beaucoup plus poussées. D'autres variables contrôles pourront être incluses dans les analyses, permettant ainsi d'examiner séparément les sondages et les strates de sondage. Ces analyses plus fines permettront de mieux cerner la contribution unique de certaines pratiques méthodologiques à l'explication de la mauvaise estimation des intentions de vote.

Références

- Blondiaux, Loïc
1991 "L'invention des sondages d'opinion: expériences, critiques et interrogations méthodologiques (1935-1950)." *Revue française de science politique*: 391-97.
- Bradburn, Norman M., et Seymour Sudman
1988 *Polls and surveys*. San Francisco: Jossey-Bass.
- Buchanan, William
1986 "Elections predictions: An empirical assessment." *Public opinion quarterly* 50: 222-227.
- Cantril, Albert H.
1991 *The opinion connection*. Washington, D.C.: Congressional Quarterly.
- Chevalier, Serge et Claire Durand
1997 *Le taux de réponse dans les sondages téléphoniques, déterminants et conséquences*. 65^e Congrès de l'ACFAS, Montréal, Université du Québec à Trois-Rivières.
- Collins, M., W. Sykes, P. Wilson, et N. Blackshaw
1988 Nonresponse: The UK experience." dans Robert M., Groves, Paul P. Biemer, Lars E. Lyberg, James T. Massey, W.L. Nicholls II, et J. Wasberg (dirs), *Telephone survey methodology*. New York: Wiley: 213-231.
- Crespi, Irving
1988 *Pre-election polling. Sources of accuracy and error*. New York: Russell Sage Foundation.
- De Heer, Wim
1996 *International response trends*. 4th International social science methodology, Colchester England, University of Essex.
- Dillman, Don A.
1978 *Mail and telephone surveys*. New York: Wiley.
- Drew, J.D., G.H. Choudhry et L.A. Hunter
1988 "Nonresponse issues in government telephone surveys" dans dans Robert M., Groves, Paul P. Biemer, Lars E. Lyberg, James T. Massey, W.L. Nicholls II, et J. Wasberg (dirs), *Telephone survey methodology*. New York: Wiley: 233-246.
- Droesbeke, Jean-Jacques, Bernard Fichet, et Philippe Tassi
1987 "La place des sondages en statistique." dans Jean-Jacques Droesbeke, Bernard Fichet et Philippe Tassi, *Les sondages*. Paris: Economica: 3-28

Durand, Claire, Sébastien Vachon, et Isabelle Tanguay

1997 "La gestion de la méthodologie de sondage au Québec et dans le reste du Canada, une question de sociétés distinctes?" *Recherches sociographiques* 39(2 et 3): 417-438.

Fox, John, Robert Andersen et Joseph Dubonnet

1997 *The polls, the press, and the 1995 Quebec referendum*, Hamilton, McMaster University.

Groves, Robert M.

1989 "Probing the causes of nonresponse and efforts to reduce nonresponse." dans Robert M. Groves, *Survey errors and survey costs*. New York: Wiley.

Groves, Robert M., Paul P. Biemer, Lars E. Lyberg, James T. Massey, William L. Nicholls II, et Joseph Waksberg

1988, eds. *Telephone survey methodology*. New York, Wiley.

Lachapelle, Guy

1991 *Les sondages et les médias lors des élections au Canada. Le pouls de l'opinion*. Approvisionnements et Services Canada, 16, Ottawa: Dundurn Press Limited & Wilson & Lafleur Limitée.

Landon, E. Laird et Sharon K. Banks

1977 "Relative efficiency and bias of plus-one telephone sampling." *Journal of marketing research* 14: 294-99.

Lau, Richard

1994 "An analysis of the accuracy of "trial heat" polls during the 1992 presidential election." *Public opinion quarterly* 58:2-20.

Lavrakas, Paul J.

1993 *Telephone survey methods: sampling, selection and supervision*. Applied social research methods series, 7, Newbury Park: Sage.

March, Catherine et Elinor Scarbrough

1990 "Testing nine hypotheses about quota sampling." *Journal of the market research society* 32(4): 485-506.

O'Neil, Michael J.

1979 "Estimating the non response bias due to refusals in telephone surveys." *Public opinion quarterly* 43(2): 218-32.

Pinard, M., R. Bernier, et V. Lemieux

1997 *Un combat inachevé* Montréal: Presses de l'Université du Québec.

Salmon, Charles, et John Spicer Nichols

1983 "The next-birthday method of respondent selection." *Public opinion quarterly* 47:270-76.

Sebold, Janice

1988 "Survey period length, unanswered numbers, and nonresponse in telephone surveys." dans Robert M. Groves, Paul P. Biemer, Lars E. Lyberg, James T. Massey, William L. Nicholls II, et Joseph Waksberg (dirs), *Telephone survey methodology*. New York: Wiley: 247-256.

Smith, Tom W.

1983 "The hidden 25 percent: An analysis of nonresponse on the 1980 general social survey." *Public opinion quarterly* 47(3):50-68.

Stephenson, C. Bruce

1979 "Probability sampling with quotas: An experiment." *Public opinion quarterly* 43(4): 477-96.

Sudman, Seymour

1976 *Applied sampling*, New York, Academic Press.

Taylor, Humphrey

1994 "Horses for courses: How survey firms in different countries measure public opinion with very different methods." *Journal of the market research society* 13(3):211-19.

Taylor, Humphrey

1997 "The very different methods used to conduct telephone surveys of the public." *Journal of the market research society* 39(3):421-32.

Traugott, Michael W.

1987 "The importance of persistence in respondent selection for preelection surveys." *Public opinion quarterly* 51:48-57.

Traugott, Michael W., Robert M. Groves, et James M. Lepkowski

1987 "Using dual frame designs to reduce nonresponse in telephone surveys." *Public opinion quarterly* 51: 522-39.

Tremblay, Victor

1988 *La sélection dans le bottins téléphoniques: ampleur et conséquences de la non-inscription*. Montréal, Centre de sondage, Université de Montréal.

Triplet, Timothy

1997 *What is gained from additional call attempts and refusal conversion and what are the cost implication?* Maryland, Survey research center, University of Maryland.

Triplett, Timothy, Jhonny Blair, Teresa Hamilton, et Yun Chiao Kang

1996 *Initial cooperators vs converted refusers: Are there response behavior differences?*.
Proceedings of the American association of public opinion research conference, Social
statistics section.

Voss, D.Stephen, Andrew Gelman, et Gary King

1995 "The polls - A review. Preelection survey methodology: details from eight polling
organizations, 1988 and 1992." *Public opinion quarterly* 59: 98-132.

Annexe 1



Montréal, le 1er septembre 1998

Madame Susan A. McDaniel,
rédactrice en chef,
Cahiers canadiens de sociologie,
a/s Département de sociologie,
[REDACTED]

Madame,

Par la présente, je désire vous soumettre un article rédigé en collaboration avec Claire Durand et André Blais.

Cet article met en lumière l'effet de certaines pratiques méthodologiques reliées à la réalisation des sondages sur l'estimation des intentions de vote. Plus précisément, il évalue les pratiques associées aux sondages mesurant les intentions de vote des résidents du Québec pendant la campagne électorale fédérale de 1997.

J'espère que cet article saura intéresser votre revue. N'hésitez pas à me contacter au besoin. Vous pouvez me joindre par courrier électronique [REDACTED]

Veuillez agréer, Madame, l'expression de mes sentiments les meilleurs.

[REDACTED]
Sebastien Vachon

[REDACTED]
[REDACTED]

Annexe 2

RÉALISATION DU SONDAGE

NOM DE LA FIRME: _____

NOM DE LA PERSONNE RESSOURCE: _____

NUMÉRO DE TÉLÉPHONE _____

1. QUESTIONS SPÉCIFIQUES AU SONDAGE	
Dates de terrain	Début: _____ Fin: _____
Commanditaire(s) du sondage	_____ _____ _____
Référence de publication (média, date, page(s))	_____ _____

3. Répartition détaillée des réponses aux questions sur l'intention de vote:

3.1 Question filtre (le cas échéant):

3.2 Première question (sur l'intention de vote):

3.3 Question de relance (le cas échéant):

3.4: Résultats publiés:

3.5 Mode de répartition des répondants discrets:
(Qui est considéré discret?):

4. GESTION DE LA COLLECTE		
4.1	Type de sondage	Ponctuel 1 Tracking 2 Panel 3 Autre, précisez: _____ 4
4.2	Mode de génération des numéros de téléphone	RDD 1 Cd-rom 2 Cd-rom + (no. de téléphone + X) 3 Telephone book 4 Other, précisez: _____ 5
	Génération de l'échantillon	Générée sur place 1 Par une firme extérieure 2
4.3	Mode de sélection dans le ménage	Grille de Kish (âge seulement) 1 Aléatoire (âge + sexe) 2 Date d'anniversaire la plus rapprochée 3 Responsable du ménage 4 Autre, précisez: _____ 5
	Utilisation des quotas	Oui 1 Non 2
	Si oui, basés sur...	_____ _____ _____
4.4	Nombre de rappels faits pour joindre le ménage sélectionné	NO _____
4.5	Récupération refus du ménage	Tous 1 Une partie seulement 2 Aucun 3
	du répondant	Tous 1 Une partie seulement 2 Aucun 3
4.6	Prise de rendez-vous	Oui 1 Non 2

4.7	Substitution de numéros de téléphone	Oui 1 Non 2
4.8	Pourcentage de validation	% _____ Validation en début de terrain seulement 1 Tout au long 2
4.9	Sondage réalisé à partir de combien de centres d'appels?	NO _____
	Où sont situés ces centres d'appels?	_____ _____ _____
	Appel à la sous-traitance pour ce sondage (collecte)	Oui 1 Non 2

5. RAPPORT ADMINISTRATIF DE CONTACTS	
5.1	Numéros générés _____
5.2	Numéros non valides
	Hors service _____
	Non résidentiel _____
	Résidence secondaire _____
5.3	Numéros hors échantillon
	Problème de langue _____
	Âge, maladie _____
	Duplicata (2 lignes au même ménage) _____
	Hors quota _____
5.4	Numéros dans l'échantillon
	Pas de réponse _____
	Absence prolongée _____
	Incomplets _____
	Refus ménage _____
	Refus répondant _____
	Autre, précisez: _____
	Entrevues complétées _____

6. STRATIFICATION, PONDÉRATION, REDRESSEMENT	
6.1	<p>L'échantillon a été stratifié</p> <p>Oui 1</p> <p>Non 2</p> <p>Sur quelle base?</p> <p>Région, précisez: _____</p> <p>_____</p> <p>Autre, précisez: _____</p> <p>_____</p>
6.2	<p>Les résultats ont été pondérés...</p> <p>En fonction du nombre de personnes dans le ménage _____</p> <p>En fonction de la fraction de sélection (par strate s'il y a lieu) _____</p> <p>En fonction du taux de réponse (par strate s'il y a lieu) _____</p>
6.3	<p>Les résultats ont été pondérés (redressés) selon les données...</p> <p>De Statistique Canada, 19 ____ _____</p> <p>Autre, spécifiez: _____</p> <p>En fonction des strates (régions) _____</p> <p>En fonction de l'âge _____</p> <p>En fonction du sexe _____</p>

Annexe 3

Intentions de vote des résidents du Québec pendant la campagne électorale fédérale de 1997

Firme de sondage	Date du début et de la fin de la période de collecte	Taille de l'échantillon	Intentions de vote					Source		
			PLC	BQ	PC	RP	NPD	Autre	Indécis	
Som	2 au 7 mai 1997	980	39	39	13	-	3	6	20	La Presse, 8 mai 1997
Angus Reid	5 au 8 mai 1997	796	38	36	20	-	4	2	-	Site de la firme, 9 mai 1997
Crop	6 au 10 mai 1997	1014	37	38	20	2	3	-	17	La Presse, 13 mai 1997
Léger et Léger	6 au 9 mai 1997	1008	38,8	37,1	19,8	,5	2	1,8	18,7	Journal de Mtl, 10 mai 1997
Gallup	7 au 12 mai 1997	270	38	43	15	-	2	1	-	La Presse, 17 mai 1997
Som	9 au 12 mai 1997	1002	36	37	19	2	1	5	22	Le Soleil, 14 mai 1997
Environnics	13 au 15 mai 1997	487	32	27	36	1	4	-	-	The Globe and Mail, 17 mai 1997
Léger et Léger	14 au 16 mai 1997	1004	32,9	35,2	30,1	,5	,8	,5	18,1	The Globe and Mail, 17 mai 1997
Som	16 au 21 mai 1997	1002	34	31	30	1	3	1	25	La Presse, 23 mai 1997
Léger et Léger	20 au 23 mai 1997	1004	33,5	32,8	31,4	-	-	2	13,8	Journal de Mtl, 24 mai 1997
Crop	22 au 26 mai 1997	1007	32	36	28	1	2	-	19	Le Soleil, 29 mai 1997
Ekos Research	22 au 25 mai 1997	970	30	40	26	-	1	4	26	The Globe and Mail, 27 mai 1997
Som	23 au 27 mai 1997	1002	33	35	28	1	2	1	19	La Presse, 29 mai 1997
Angus Reid	22 au 27 mai 1997	810	29	38	31	-	2	-	-	Site de la firme, 30 mai 1997
Léger et Léger	26 au 28 mai 1997	1005	34,6	36,1	26,3	-	-	3	12,7	Journal de Mtl, 30 mai 1997
Gallup	25 au 28 mai 1997	389	37	33	29	-	1	-	-	The Gazette, 30 mai 1997
Environnics	25 au 28 mai 1997	472	33	37	26	1	4	1	-	The Globe and Mail, 30 mai 1997