

Université de Montréal

Les sondages électoraux français de 2002  
Influence de la méthodologie sur la qualité des estimations

par

Mylène Larochelle

Département de Sociologie  
Faculté des Arts et des Sciences

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

juin 2004

©Mylène Larochelle, 2004



HM

15

U54

2004

V.016

**Direction des bibliothèques**

**AVIS**

L'auteur a autorisé l'Université de Montréal à reproduire et diffuser, en totalité ou en partie, par quelque moyen que ce soit et sur quelque support que ce soit, et exclusivement à des fins non lucratives d'enseignement et de recherche, des copies de ce mémoire ou de cette thèse.

L'auteur et les coauteurs le cas échéant conservent la propriété du droit d'auteur et des droits moraux qui protègent ce document. Ni la thèse ou le mémoire, ni des extraits substantiels de ce document, ne doivent être imprimés ou autrement reproduits sans l'autorisation de l'auteur.

Afin de se conformer à la Loi canadienne sur la protection des renseignements personnels, quelques formulaires secondaires, coordonnées ou signatures intégrées au texte ont pu être enlevés de ce document. Bien que cela ait pu affecter la pagination, il n'y a aucun contenu manquant.

**NOTICE**

The author of this thesis or dissertation has granted a nonexclusive license allowing Université de Montréal to reproduce and publish the document, in part or in whole, and in any format, solely for noncommercial educational and research purposes.

The author and co-authors if applicable retain copyright ownership and moral rights in this document. Neither the whole thesis or dissertation, nor substantial extracts from it, may be printed or otherwise reproduced without the author's permission.

In compliance with the Canadian Privacy Act some supporting forms, contact information or signatures may have been removed from the document. While this may affect the document page count, it does not represent any loss of content from the document.

Université de Montréal  
Faculté des études supérieures

Ce mémoire intitulé:

Les sondages électoraux français de 2002  
Influence de la méthodologie sur la qualité des estimations

présenté par:

Mylène Larochelle

a été évalué par un jury composé des personnes suivantes:

Paul Bernard - président-rapporteur

Claire Durand - directrice de recherche

André Blais - codirecteur de recherche

Patrick Fournier - membre du jury

## RÉSUMÉ

---

Le mémoire présentera une analyse des sondages électoraux français, des méthodologies utilisées et de leur performance au premier tour de l'élection présidentielle française de 2002. Les données sont constituées des résultats de tous les sondages réalisés durant le mois précédant le premier tour de l'élection présidentielle française de 2002 de même que des informations méthodologiques fournies par les médias-écrits nationaux, les sites Internet des six principaux instituts de sondage et la Commission des sondages. Des analyses de séries chronologiques permettent a) de tracer le portrait de l'évolution des intentions de vote des seize candidats et des quatre regroupements de partis (extrême gauche, gauche modérée, droite modérée et extrême droite), b) d'observer les tendances à la hausse ou à la baisse de l'évolution des intentions de vote et c) de calculer l'écart entre la prédiction de la série et le résultat de l'élection. À partir des séries chronologiques des regroupements de partis, une analyse de la performance individuelle de chaque sondeur sera effectuée.

Les résultats des analyses indiquent que les principales erreurs de prédictions sont la surestimation de la gauche modérée, la sous-estimation de l'extrême droite et plus particulièrement la sous-estimation du candidat du Front national Le Pen. Cette recherche permettra de vérifier la qualité des estimations et d'examiner si l'on peut attribuer les erreurs de prédiction à la méthodologie de sondage.

**Mots clés:** sondage, méthodologie, élection présidentielle, France, prédiction, institut de sondage, estimation.

## ABSTRACT

---

This report will present an analysis of the French electoral surveys, the methodologies used, and their performances during the first round of the French presidential elections. The results of all the polls that were taken during the last month of the French presidential election of 2002, as well as the information given by the national media, and the internet sites of the six principal pollsters and of the *Commission des sondages*, constitute our database. Time series analysis does three things. First, it maps the evolution of the estimations of the sixteen candidates and of the four partisan groupings (extreme left, moderate left, moderate right, extreme right). Second, it tracks how the votes increase or decrease. Finally, it allows us to estimate the difference between the prediction of the votes and the final presidential results. A performance analysis is carried out for every candidate of each party.

The results show that the major errors during the predictions are the surestimation of the moderate left, the underestimation of the extreme right, and of the underestimate of Le Pen (Front National party). This analysis is to verify the quality of the estimation, and to examine if errors can be attributed to the methodology used by pollsters.

**Key words:** poll, survey, pollster, French election, prediction, methodology, estimation.

## TABLE DES MATIÈRES

---

La page de titre .....	i
L'identification du jury .....	ii
Le résumé en français et les mots clés français .....	iii
Le résumé en anglais et les mots clés anglais .....	iv
La table des matières .....	v
La liste des tableaux .....	vii
La liste des graphiques .....	viii
Les remerciements .....	ix
<b>INTRODUCTION</b> .....	<b>1</b>
<b>CHAPITRE 1 - Contexte socio-politique</b> .....	<b>4</b>
1.1 Le contexte social .....	4
1.1.1 Opinion publique .....	4
1.1.2 Science et démocratie .....	7
1.1.2.1 Critiques de Bourdieu en France .....	8
1.1.2.2 Contre-critiques .....	9
1.2 Origines et essor des sondages d'opinion .....	11
1.3 Le contexte français .....	13
1.3.1 La période pré-sondage .....	13
1.3.2 Le système politico-institutionnel .....	15
1.3.3 Les élections présidentielles de 1965 à 1995 .....	18
1.3.4 L'évolution du comportement électoral .....	19
1.3.5 La réglementation de la publication et diffusion des sondages .....	21
1.3.5.1 Loi n° 77-808 du 19 juillet 1977 .....	21
1.3.5.2 Loi n° 2002-214 du 19 février 2002 .....	23
1.3.6 Synthèse .....	23
<b>CHAPITRE 2 - Sondages et élections: quelques cas types</b> .....	<b>25</b>
2.1 Élections législatives .....	25
2.1.1 Grande-Bretagne .....	25
2.1.2 Irlande .....	28
2.1.3 Québec/Canada .....	29
2.2 Élections présidentielles .....	30
2.2.1 États-Unis .....	31
2.2.2 France .....	32

CHAPITRE 3 - Problématique .....	36
3.1 Explications des erreurs de prédiction .....	37
3.1.1 Explications psychosociologiques .....	38
3.1.2 Explications méthodologiques .....	40
3.2 Premier tour de l'élection présidentielle française de 2002 .....	42
3.2.1 Campagne électorale .....	42
3.2.2 Résultats du vote .....	44
3.3 Méthodologie des instituts de sondage français .....	46
3.4 Objectifs de recherche .....	53
 CHAPITRE 4. Méthodologie .....	 58
4.1 Données .....	58
4.2 Méthode d'analyse .....	60
4.2.1 Procédure d'analyse choisie .....	60
4.2.2 Justification de la procédure d'analyse .....	62
 CHAPITRE 5. Résultats .....	 66
5.1 Évolution des intentions de vote .....	66
5.1.1 Évolution des candidats .....	67
5.1.2 Évolution des regroupements de partis .....	74
5.2 Prédictions des séries chronologiques .....	78
5.3 Performance des instituts de sondage .....	80
 CHAPITRE 6. Discussion des résultats .....	 83
6.1 Hypothèses possibles .....	83
6.2 Principales découvertes .....	87
6.3 Redressements .....	91
 CONCLUSION .....	 94
BIBLIOGRAPHIE .....	99
ANNEXE A .....	104
ANNEXE B .....	107
ANNEXE C .....	111
ANNEXE D .....	114
ANNEXE E .....	116
ANNEXE F .....	118
ANNEXE G .....	121



## LISTE DES TABLEAUX

---

TABLEAU I	Les catégories socio-démographiques de base de l'échantillonnage par quotas . . .	47
TABLEAU II	Les dates de réalisation des derniers sondages de chaque institut . . . . .	49
TABLEAU III	Les écarts entre les résultats du vote et les prédictions des séries pour les candidats et pour les regroupements de partis . . . . .	78
TABLEAU IV	La comparaison des sondages présentant des écarts significatifs par rapport à l'estimation moyenne des intentions de vote des regroupements de partis . . . .	88
TABLEAU V	La comparaison des caractéristiques méthodologiques de l'institut Ifop et des autres instituts . . . . .	90
TABLEAU VI	Les sondages réalisés du 20 mars au 18 avril 2002 . . . . .	105
TABLEAU VII	La méthodologie du dernier sondage de chaque institut . . . . .	108
TABLEAU VIII	Les coefficients de régression des séries chronologiques des candidats de gauche . . . . .	112
TABLEAU IX	Les coefficients de régression des séries chronologiques des candidats de droite . . . . .	112
TABLEAU X	Les coefficients de régression des séries chronologiques des regroupements de partis . . . . .	113
TABLEAU XI	L'évolution des intentions de vote pour les candidats de gauche du 20 mars au 21 avril 2002 . . . . .	115
TABLEAU XII	L'évolution des intentions de vote pour les candidats de droite du 20 mars au 21 avril 2002 . . . . .	115
TABLEAU XIII	L'évolution des intentions de vote pour les regroupements de partis du 20 mars au 21 avril 2002 . . . . .	115
TABLEAU XIV	Les redressements effectués par l'institut Sofres pour le sondage réalisé les 17 et 18 avril 2002 . . . . .	117
TABLEAU XV	Les redressements effectués par l'institut Ipsos pour le sondage réalisé les 17 et 18 avril 2002 . . . . .	119
TABLEAU XVI	La reconstitution du vote aux élections législatives de 1997 par l'institut Sofres pour le sondage réalisé les 17 et 18 avril 2002 . . . . .	122
TABLEAU XVII	La reconstitution du vote au premier tour de l'élection présidentielle de 1995 par l'institut Sofres pour le sondage réalisé les 17 et 18 avril 2002 . . . . .	123
TABLEAU XVIII	La reconstitution du vote au deuxième tour de l'élection présidentielle de 1995 par l'institut Sofres pour le sondage réalisé les 17 et 18 avril 2002 . . .	123

---

**LISTE DES GRAPHIQUES**

---

<b>GRAPHIQUE 1:</b>	La modélisation ARIMA	page 63
<b>GRAPHIQUE 2:</b>	Évolution des intentions de vote de Jospin	page 68
<b>GRAPHIQUE 3:</b>	Évolution des intentions de vote de Laguille	page 68
<b>GRAPHIQUE 4:</b>	Évolution des intentions de vote de Chevènement	page 69
<b>GRAPHIQUE 5:</b>	Évolution des intentions de vote de Mamère	page 69
<b>GRAPHIQUE 6:</b>	Évolution des intentions de vote de Besancenot	page 70
<b>GRAPHIQUE 7:</b>	Évolution des intentions de vote de Hue	page 70
<b>GRAPHIQUE 8:</b>	Évolution des intentions de vote de Taubira	page 70
<b>GRAPHIQUE 9:</b>	Évolution des intentions de vote de Gluckstein	page 70
<b>GRAPHIQUE 10:</b>	Évolution des intentions de vote de Chirac	page 71
<b>GRAPHIQUE 11:</b>	Évolution des intentions de vote de Le Pen	page 71
<b>GRAPHIQUE 12:</b>	Évolution des intentions de vote de Bayrou	page 72
<b>GRAPHIQUE 13:</b>	Évolution des intentions de vote de Saint-Josse	page 72
<b>GRAPHIQUE 14:</b>	Évolution des intentions de vote de Madelin	page 73
<b>GRAPHIQUE 15:</b>	Évolution des intentions de vote de Mégret	page 73
<b>GRAPHIQUE 16:</b>	Évolution des intentions de vote de Lepage	page 74
<b>GRAPHIQUE 17:</b>	Évolution des intentions de vote de Boutin	page 74
<b>GRAPHIQUE 18:</b>	Évolution des intentions de vote de l'extrême gauche	page 75
<b>GRAPHIQUE 19:</b>	Évolution des intentions de vote de la gauche modérée	page 75
<b>GRAPHIQUE 20:</b>	Évolution des intentions de vote de la droite modérée	page 76
<b>GRAPHIQUE 21:</b>	Évolution des intentions de vote de l'extrême droite	page 76
<b>GRAPHIQUE 22:</b>	Évolution des intentions de vote de l'extrême gauche et les sondages	page 80
<b>GRAPHIQUE 23:</b>	Évolution des intentions de vote de la gauche modérée et les sondages	page 80
<b>GRAPHIQUE 24:</b>	Évolution des intentions de vote de la droite modérée et les sondages	page 81
<b>GRAPHIQUE 25:</b>	Évolution des intentions de vote de l'extrême droite et les sondages	page 81

## REMERCIEMENTS

---

Sans l'aide et le soutien de certaines personnes, je n'aurais certainement pu actualiser ce projet de mémoire. Je dédie ce mémoire à toutes les personnes qui de près ou de loin auront été des sources d'encouragement, de motivation et d'inspiration.

Je n'exprimerai jamais assez ma gratitude envers Claire Durand qui m'a proposé d'entamer ce projet de recherche en mars 2002, lors de mon séjour à Lille. Sa grande expertise méthodologique, sa rigueur, sa générosité, son encadrement, son enthousiasme et sa disponibilité en sol outre-atlantique et montréalais auront été essentiels à la réalisation de mon mémoire.

Je tiens également à remercier mon codirecteur André Blais. Sa grande expertise politique, et particulièrement, ses connaissances du contexte électoral français m'auront été d'une grande aide.

Du plus profond de mon coeur, je tiens à remercier toute ma famille: mon père, ma mère, mon frère, sur qui j'ai toujours pu m'appuyer autant financièrement, moralement et affectueusement. D'autres personnes ont aussi été des sources inestimables d'encouragement; je pense, entre autres, à Annick, Annie, Brigitte, Bruno, Christine, Cindy, Caroline, Frédérick, Jean-Steve, Marc, Marie-Eve, Mélanie, Mylène, Pascale, Pauline, Sophie, Suzanne, Victor. Que ce soit par un simple mot ou un soutien quotidien, toutes ces personnes auront été très précieuses dans la réalisation de ce projet. Je vous aime!

Je souhaite également exprimer ma reconnaissance au département de sociologie pour m'avoir accordé une bourse de fin de rédaction. Cette bourse aura permis de me consacrer à plein temps à la réalisation de mon mémoire. De plus, je tiens à remercier la Commission des sondages, son président Monsieur Galabert et sa secrétaire Madame Volaine pour leur accueil et leur gentillesse lors de mes visites.

## INTRODUCTION

Les résultats du premier tour de l'élection présidentielle française de 2002 ont suscité une importante polémique à propos de la validité des sondages. Tous les sondages prédisaient un duel entre Chirac et Jospin au second tour de l'élection présidentielle. Cependant, le jour de l'élection, le candidat de l'extrême droite Le Pen devança Jospin de moins d'un point de pourcentage. Cet événement n'est pas unique dans l'histoire des sondages. Un écart systématique et semblable pour tous les sondeurs et la sous-estimation d'un parti ou d'un candidat, généralement de droite, se sont déjà produits entre autres en Grande-Bretagne (1992, 1997) et au Québec/Canada (1998, 2000). Toutefois, le cas français se caractérise par l'importance des conséquences du biais des sondages. Quelle est l'ampleur réelle du biais des sondages? Serait-il possible d'expliquer ce biais par la méthodologie de sondage?

L'analyse des sondages du premier tour de l'élection présidentielle française de 2002 nécessite une étude historique du contexte social, politique et intellectuel des sondages d'opinion. À ce propos, Blondiaux pose l'hypothèse justifiant tout à fait la logique adoptée dans le cadre du premier chapitre:

Pour comprendre l'instrument, il faut s'en dégager, apprendre à en éloigner le regard et s'intéresser non pas seulement à l'histoire des procédures, des méthodes et de leur transformation éventuelle, mais aussi, et peut-être surtout, au contexte social, politique, intellectuel dans lequel cette méthode a vu le jour et s'est diffusée (Blondiaux, 1998: 18).

Le postulat de l'acceptation du sondage comme instrument de mesure de l'opinion publique implique une sélection d'événements liés au concept d'opinion publique qui seront présentés au premier chapitre. La période précédant l'avènement des sondages d'opinion, leurs conditions d'apparition, leur évolution et leur état actuel seront constamment interrogés. Ensuite, l'analyse du contexte français permettra de poser les

bases nécessaires afin d'éclairer le problème posé, celui des erreurs de prédiction des sondages au premier tour de l'élection présidentielle française de 2002.

En fait, les situations électorales diffèrent selon les pays et les campagnes: les forces et les faiblesses des candidats, les enjeux de la campagne, les écarts plus ou moins importants entre les principaux candidats ou partis, etc. Dans une lutte serrée entre les candidats, ou les partis, l'estimation du vainqueur est certainement plus difficile que dans le cas d'un candidat ou d'un parti donné largement en avance. Et puis, contrairement aux échecs, "la plupart des nombreux succès enregistrés dans le monde par les sondages passent relativement inaperçus parce qu'ils s'inscrivent naturellement dans les orientations politiques du moment" (Max, 1981: 118). Toutefois, quelques cas types de grandes erreurs de prédiction électorales sont identifiées par les spécialistes en sondages. Au second chapitre, les principaux cas types ainsi que les raisons avancées pour les expliquer seront présentés. De plus, afin de bien comprendre le cas français de 2002, l'ensemble des performances des sondeurs lors des élections présidentielles françaises de 1965 à 1995 seront expliquées.

Le troisième chapitre permet de faire le point et d'exposer les explications possibles des erreurs de prédiction. Un retour sur la campagne électorale, les résultats du scrutin au premier tour et la méthodologie utilisée par les instituts français permettront de poser les explications possibles des erreurs de prédiction survenues au premier tour de l'élection présidentielle française de 2002. Enfin, les objectifs de recherche seront identifiés afin de traiter ce problème.

L'analyse de séries chronologiques est essentielle pour étudier les estimations et les prédictions de sondages. Au quatrième chapitre, elle sera présentée de même que les données utilisées dans le cadre de cette analyse. Les analyses de séries chronologiques ARIMA permettront de tracer un portrait de l'évolution des intentions de vote des candidats et des regroupements de partis, de calculer le nombre de points gagnés ou perdus entre le début et la fin de la série chronologique, et de calculer l'écart entre la prédiction finale de la série chronologique et le résultat de l'élection. De plus, cette

analyse servira à évaluer la qualité du travail de chaque sondeur.

Au cinquième chapitre, l'analyse de séries chronologiques assurera une prise en compte adéquate des tendances dans la prévision, par les sondages, du score des divers candidats lors du scrutin. En effet, elle présentera l'évolution moyenne des intentions de vote des candidats et des regroupements de partis (extrême gauche, gauche modérée, droite modérée et extrême droite) et permettra d'évaluer la qualité des estimations et des prédictions de sondages.

Suite à l'identification des erreurs de prédiction, les explications possibles de ces biais de sondages seront discutées. Le sixième chapitre portera sur trois points principaux afin de répondre à cette question: les hypothèses possibles qui ne peuvent être vérifiées faute de données adéquates, les principales découvertes ainsi que la remise en question de certaines pratiques "sondagières".

Enfin, la conclusion permettra de résumer les six chapitres afin de comprendre l'ensemble de la démarche effectuée pour traiter du problème posé. Ensuite, un bref retour sur les hypothèses non-vérifiées dans le cadre de ce travail ainsi que sur la remise en question des pratiques assez peu communes des sondeurs français, en particulier les redressements, seront présentées. Finalement, les principales découvertes de ce mémoire seront présentées.

## CHAPITRE 1 - Contexte socio-politique

L'avènement des sondages se fonde sur une demande constante de connaissances de la part des pouvoirs centraux, un développement de la démocratie et de ses composantes (élections, presse libre, etc.), de l'institutionnalisation des pratiques, de la diffusion des sondages, etc. Ce premier chapitre porte sur le contexte social de l'évolution du concept d'opinion publique, sur la naissance des sondages d'opinion aux États-Unis et particulièrement sur le contexte français. La section sur le contexte français pose les bases nécessaires afin d'éclairer le problème posé, celui des erreurs de prédiction des sondages au premier tour de l'élection présidentielle française de 2002. Pour étudier ce problème, il est primordial de comprendre comment la France s'est approprié les sondages d'opinion, de quelle façon le système politico-institutionnel a permis le développement et la légitimation des sondages aux élections présidentielles françaises à partir de 1965. L'évolution du comportement électoral ainsi que la réglementation de la publication et diffusion des sondages sont également des éléments essentiels au contexte français pour traiter le problème posé. Auparavant, il convient de faire un bref retour sur la période pré-sondage afin de saisir le contexte social ayant permis le développement des sondages d'opinion.

### 1.1 Le contexte social

#### 1.1.1 Opinion publique

Les gouvernements démontrent depuis longtemps de l'intérêt pour connaître l'état d'esprit et les caractéristiques démographiques de la population pour des fins de contrôle financier (impôts) et de contrôle politique (révoltes). Les connaissances quantifiées sur la population ont d'abord été assouvies par les recensements, puis ensuite sont apparus les rapports périodiques sur l'état de l'opinion publique, les semeurs de bruits recueillant les propos des habitants, etc. Au début du XIX<sup>e</sup> siècle, une nouvelle source

d'information sur l'opinion publique s'est présentée aux États-Unis: les "votes de paille" (Stoetzel et Girard, 1973: 40). Selon Robinson (1932), l'origine de ce nouveau procédé correspondrait aux deux "votes de paille" publiés dans la presse à l'occasion de l'élection présidentielle américaine de 1824. Cette nouvelle méthode utilisée pendant un peu plus d'un siècle consistait à recueillir des bulletins découpés dans les journaux et retournés après réponse, des cartes envoyées par la poste et des réponses à des bureaux de vote simulés dans certains quartiers.

Les débats associés au "parler au nom de l'opinion publique" ont dicté la signification du concept d'opinion publique. Sans prétendre présenter l'ensemble des sens attribués au concept d'opinion publique, il semble primordial d'en présenter certaines formes. L'article de Cabin (1999) permet d'identifier quatre étapes historiques des visions de l'opinion publique. Selon cet auteur, l'apparition de l'expression "opinion publique", attribuée à Rousseau, date du milieu du XVIII<sup>e</sup> siècle et désigne "les idées et sentiments partagés par un peuple ou une communauté" (Cabin, 1999: 13). L'expression a toutefois une connotation péjorative et elle fait référence au domaine du contrôle social. En effet, cette vision se rattache davantage au rôle de censure, de conformité des moeurs et de pression sociale exercée sur les conduites individuelles qu'à l'idée de sentiments partagés.

À la fin du XVIII<sup>e</sup> siècle, la "Raison" est triomphante. Selon Cabin (1999), Necker serait le premier à associer la formation de l'opinion publique à la catégorie des gens éclairés et informés. Cette période se caractérise entre autre par la rencontre des intellectuels dans les cafés, la lecture des journaux et les discussions sur les sujets d'actualité. Une nouvelle sphère publique se constitue pour reprendre l'idée d'Habermas (1978): un espace de discussion, d'émancipation individuelle et d'usage public de la raison individuelle. De plus, la première partie du XX<sup>e</sup> siècle serait également marquée par la théorie de Bryce (1911) et sa définition de l'opinion publique comme une "opinion informée et rationnelle". De cette approche découle la réflexion sur l'aptitude à faire partie du public. Parmi les nombreuses définitions de l'opinion publique se référant à ce questionnement, Champagne (1990) propose comme solution alternative aux



sondages, le recours à l'avis de l' "opinion éclairée" - ceux qui "savent".

En plus de proposer le recours à l'opinion éclairée, Champagne (1990) suggère l'étude des manifestations directes des minorités agissantes. Cette vision de l'opinion publique, celle du XIX<sup>e</sup> siècle, serait celle de l'âge de l'expression populaire et du mouvement social de Cabin (1999). Des gens se mobilisent, des groupes syndicaux et des partis politiques sont créés afin de défendre les intérêts, manifester, protester, etc. Le gouvernement porterait un intérêt particulier aux mouvements de masse et à leurs opinions à des fins de contrôle et de prévision, de façon à connaître la population, obtenir son soutien et parler en son nom.

Selon Cabin (1999), le XX<sup>e</sup> siècle fut celui de l'opinion "sondagière". La mise au point du sondage d'opinion par les statisticiens et les psychosociologues débute vers les années 30. Au moment de l'apparition des sondages, l'étude des pratiques des acteurs sociaux, le sens qu'ils accordent à la notion d'opinion publique et l'usage qu'ils en font auraient permis de comprendre le contenu sémantique accordé à l'opinion. De près ou de loin, les élites politiques, administratives et économiques, aux États-Unis, cherchaient à produire, manipuler et faire parler l'opinion publique.

Quel est le sens moderne du concept d'opinion publique? L'opinion publique équivaut-elle tout simplement à la définition attribuée à Gallup (1901-1984), c'est-à-dire à "ce que mesurent les sondages" (Cabin, 1999: 13)? Certains auteurs tels Cabin (1999) ainsi que Stoetzel et Girard (1973) affirment que les sondages d'opinion correspondent à des expressions parmi d'autres et aucunement à un moyen exclusif de mesurer l'opinion publique.

Blondiaux spécifie:

Les sondages réalisent l'opinion publique au sens où "réaliser" signifie faire exister à titre de réalité concrète ce qui n'existait que dans l'esprit ou à faire correspondre une chose à une idée, à un mot. Ils fournissent une formulation tangible, chiffrée, univoque, tout à la fois susceptible de faire l'accord des parties en présence et d'être appropriée dans des stratégies discursives particulières, capable de dégager du magma des impressions confuses, des demi-mesures, des représentations en conflit de l'opinion publique qui prévalait avant leur apparition (Blondiaux, 1998: 15-16).

Dans cette optique, le postulat de départ établissant le sondage d'opinion comme un instrument de mesure de l'opinion publique est justifiable et justifié. Mais que représentent les sondages d'opinion en science et en démocratie?

### 1.1.2 Science et démocratie

Selon Brûlé, tout aurait fort bien commencé entre les sondeurs et les universitaires: "enfants du calcul des probabilités et de la psychologie sociale, les sondages d'opinion étaient dès leur naissance voués aux échanges avec le monde universitaire" (Brûlé, 1988: 186). Dans les années 70, des débats et des discours tenus sur les sondages provenant de milieux politique, intellectuel et social ont porté à la fois sur la science et sur la démocratie. Comme le souligne Blondiaux:

Si les premiers entreprennent dès l'origine de démontrer la *justesse* de l'instrument (discours sur la science), ils tiennent parallèlement à en établir la *justice* au regard d'une norme politique établie (discours sur la démocratie). À l'inverse, leurs adversaires, au nom d'une autre définition de la démocratie (ou d'une interprétation différente de la même norme), ont entrepris de dénoncer la menace que ferait peser l'instrument sur le bon fonctionnement des régimes politiques (Blondiaux, 1998: 187).

Ces débats sur le caractère démocratique des sondages peuvent s'analyser comme une lutte pour la légitimation de cette pratique scientifique. Par contre, l'utilisation de ce

concept par de nombreux acteurs nuit au processus de conceptualisation de l'opinion publique. Comme l'indique Blondiaux:

L'opinion publique constitue un objet universitairement pertinent, digne d'être étudié et pour lequel un effort de clarification conceptuelle s'impose. Mais dans le même temps, et parce qu'il s'agit d'une notion largement répandue, galvaudée par les acteurs politiques, les usages vulgaires de la notion en ont altéré le sens au point de parasiter le travail de définition du spécialiste de sciences sociales (Blondiaux, 1998: 126).

Les multiples motivations des acteurs et leurs constructions du concept d'opinion publique rendent le travail des scientifiques difficile. Lorsqu'une notion a plusieurs sens, comme dans ce cas-ci, s'entendre et créer une science légitime devient une tâche lourde. Les débats aux États-Unis des années 40, dans l'Allemagne des années 50 ou en France des années 70 seront récurrents.

### **1.1.2.1 Critiques de Bourdieu en France**

Bourdieu occupe une place prépondérante dans les débats intellectuels des années 70 en France. Dans son article "L'opinion publique n'existe pas", Bourdieu indique: "les sondages peuvent apporter une contribution utile à la science sociale à condition d'être traités de façon rigoureuse, c'est-à-dire avec des précautions particulières" (Bourdieu, 1973: 1292). Pour parvenir à cette rigueur, Bourdieu (1973) conteste trois postulats: 1) celui selon lequel tout le monde a une opinion, 2) le fait que toutes les opinions se valent, et enfin 3) le postulat implicite du consensus sur les problématiques imposées par les questions de sondage.

Selon Bourdieu (1973), l'effet consensuel selon lequel "tout le monde a une opinion" en vient à ignorer les gens "sans opinion" à certaines questions dans les sondages. Les gens n'ayant pas d'opinion ont un portrait particulier et leurs non-réponses doivent être traitées et analysées. La signification accordée à la non-réponse est essentielle à la compréhension de l'ensemble des réponses, de la problématique.

De plus, les répondants n'ont pas tous une "compétence politique" équivalente; la production d'une opinion n'est pas à la portée de tous. Bourdieu (1973) critique aussi le postulat selon lequel toutes les opinions se valent. Selon cet auteur, "le fait de cumuler des opinions qui n'ont pas du tout la même force réelle aboutit à une distorsion très profonde" (Bourdieu, 1973: 1293).

Un autre biais des enquêtes par sondage est lié à la formulation des questions, aux problématiques obligées. Comme il le note: "les problématiques [...] sont profondément liées à la conjoncture et dominées par un certain type de demande sociale. Autrement dit, les problèmes qui se posent sont des problèmes qui s'imposent comme problèmes politiques" (Bourdieu, 1973: 1294). Cet auteur déplore que les problématiques ne correspondent pas aux intérêts des gens, mais bien à ceux des hommes politiques ou des commanditaires. De plus, le fait de poser la même question à tout le monde présuppose l'existence de l'hypothèse d'un consensus sur les questions méritant d'être posées. Les gens interrogés sur des questions qu'ils ne se posent pas se retrouvent à formuler un avis pouvant être sans enjeu réel pour eux. Les questions ne doivent pas servir à produire une opinion, mais bien à la recueillir.

Finalement, le sondage d'opinion est un instrument d'action politique, et l'imposition de l'existence de l'opinion publique se fait par la somme des opinions individuelles. Les problématiques guident "la signification des réponses ainsi [que] la signification qui est donnée à la publication des réponses" (Bourdieu, 1973: 1295). L'existence de l'opinion publique viendrait "légitimer une politique et renforcer les rapports de force" (Bourdieu, 1973: 1295).

### **1.1.2.2 Contre-critiques**

Dans ses "Réflexions sur la "sondomanie"", Jaffré (1985) note bien que les sondages agacent, inquiètent et irritent, mais il accorde à cette "sondomanie" des raisons profondes et certains avantages. Selon Jaffré, "Dans les sociétés contemporaines où n'existe plus la connaissance interpersonnelle, le sondage est l'un des moyens qui permet de connaître

l'opinion des autres et de situer sa propre opinion" (Jaffré, 1985: 16). Les sondages informent non seulement les hommes politiques, mais également les citoyens.

Selon Stoetzel et Girard (1973), la publication des sondages en période électorale participerait à la démocratisation de l'information. Ces auteurs ajoutent qu'un grand nombre de personnes seraient en déficit d'éléments de renseignements si les commanditaires demeuraient les seuls détenteurs de l'information. En plus de fournir des informations sur les mentalités du temps présent, le grand nombre de sondages publiés permet la comparaison des sondages. En effet, Jaffré (1985) indique qu'un changement important dans l'opinion pourra être éclairé par un sondage supplémentaire portant sur le sujet. L'abondance des sondages réduit les risques d'intoxication, de rétention de l'information ou de mauvaise interprétation. Plusieurs informations sur un sujet permettent nécessairement d'en dresser un portrait plus judicieux et une analyse supérieure. D'après Lancelot (1984), le sondage participe à la "culture de liberté". Dans le pluralisme et le relativisme de cette source d'information, les sondages contribuent à maximiser la démocratie, offrant ainsi cette "culture de liberté".

De plus, l'apport des sondages à la démocratie s'entrevoit selon Lancelot (1984) par la sélection et le contrôle des gouvernants. Le sondage joue un rôle de preuve comme le souligne Bogart (1972): "une majorité silencieuse est par définition sans effet. Elle ne peut devenir une force politique que lorsque les sondages la rendent consciente d'elle-même ou lorsque des leaders politiques revendiquent son soutien" (cité dans Blondiaux, 1998: 14). En effet, le recours à l'opinion publique par le sondage permet de justifier les actions gouvernementales. Le respect des droits de l'opposition correspond également à un apport des sondages à la démocratie: le sondage ne représente pas uniquement la majorité. "L'opinion publique n'est pas plus homogène qu'elle n'est jamais unanime" (Stoetzel et Girard, 1973:178). Le sondage laisse apparaître des proportions, donne le droit de parole à une minorité. La complexité de la structure de l'opinion se retrouve dans l'interprétation du sondage. Au sujet de la complexité de l'opinion, Cayrol contredisant Bourdieu ajoute que "l'objet opinion n'est pas donné par la seule addition de réponses isolées. Il faut élaborer des structures, interpréter les résultats, poser des

hypothèses d'analyse, établir des plans d'expérience, confronter les données aux hypothèses, retourner sur le terrain, analyser et vérifier encore..." (Cayrol, 2000: 129).

## 1.2 Origines et essor des sondages d'opinion

Aux États-Unis, deux traditions sont à l'origine des sondages d'opinion: l'étude de marché et l'enquête journalistique. Les études de marché ont débuté lorsque les entreprises industrielles et commerciales ont cherché à connaître les habitudes et les préférences de leur clientèle. En 1935, Roper, un spécialiste des études de marché, publie les résultats d'une enquête portant sur les attitudes des Américains. Cette enquête publiée dans le magazine *Fortune* présente les résultats obtenus à l'aide d'un questionnaire mené auprès d'un échantillon de 3000 personnes. Roper sera amené à devenir l'un des pères fondateurs de l'industrie des sondages aux États-Unis.

En 1935, Gallup crée son institut, *l'American Institute of Public Opinion* (AIPO) et signe des accords avec plusieurs journaux américains. Gallup réussit plus que ses compétiteurs Roper et Crossley à incarner la nouvelle pratique, entre autres, par son statut de professeur de journalisme et de publicité à l'université Drake (Iowa) et à la Northwestern University (Illinois). Gallup, qui compléta une *M. A.* en psychologie (1925) et un *Ph. D.* en journalisme (1928), notera dans le cadre de sa thèse de doctorat l'idée de représentativité d'un groupe "obtenu par un choix convenable de sujets d'après leur appartenance aux diverses catégories professionnelles, et que le nombre de sujets à choisir peut être déterminé par un calcul théorique" (Stoetzel et Girard, 1973: 53).

Les journalistes jouent un rôle important dans le processus de redéfinition de l'opinion publique en commandant des enquêtes et en représentant l'opinion. La presse sera la première à fournir des résultats d'enquêtes et ainsi, elle participera à la circulation de cette nouvelle source d'information. Les médias permettront à tous d'accéder à l'information et ils financeront les instituts.

Dès les années 20 et 30, l'adhésion à la technique est quasi immédiate chez les

spécialistes en sciences sociales aux États-Unis, principalement ceux issus de la psychologie et de la sociologie. Les méthodes traditionnelles d'analyse laissent place aux interrogations sur le statut de la science et à la formation de croyances. La technique du sondage d'opinion permet la connaissance et la mesure objective par sa théorie statistique et par la disparition de l'observateur. De plus, l'opérationnalisation du concept en peu de temps a permis de transformer la notion d'opinion publique de "concept ambigu" en "construit mesurable", pour reprendre l'idée de Cantrell (1992), et ce, aux États-Unis, vers le milieu des années 40, et un peu plus tard dans les autres démocraties occidentales.

Pourquoi les États-Unis sont-ils le lieu de naissance et de prédilection des sondages d'opinion? Lors de l'élection présidentielle américaine de 1936, le vote de paille connaît un échec retentissant à l'avantage des sondages d'opinion. Blondiaux (1998) note qu'avant même que le vote de paille soit lancé et que les résultats soient publiés dans la revue *Literary Digest*, Gallup prédit que la revue annoncera la victoire de Landon avec 56% des intentions de vote contre 44% pour Roosevelt. Les résultats de la revue provenant de 2 400 000 bulletins donneront effectivement, à tort, Landon vainqueur avec 57,4% et 42,6% à Roosevelt. Quant aux résultats de l'élection, ils seront tout autres, 61% pour Roosevelt et 39% pour Landon. Pourtant, cette tendance aura été saisie par les instituts de sondage. L'AIPO de Gallup annonçait bien Roosevelt vainqueur à 54% et 46% pour Landon (Blondiaux, 1998: 255-256). Selon Stoetzel et Girard (1973), la revue *Literary Digest* aurait sous-estimé le résultat pour Roosevelt en utilisant la liste des abonnés au téléphone et des propriétaires d'automobiles correspondant à la partie la plus riche de la population. Tandis que les instituts auraient complété leur échantillon par des sondages postaux et face à face avec des gens défavorisés - la proportion de gens qui votent davantage pour Roosevelt - et ainsi obtenir un échantillon représentatif selon la logique des quotas qui prévaut à cette époque. Cette victoire constitue non seulement celle des instituts, mais aussi et surtout celle de la méthode.

Revue scientifique, centres de recherche, organismes et instituts de sondage se succéderont et occuperont une place incontestable aux États-Unis à partir des années 50.

Blondiaux trace le parallèle suivant: “l’opinion publique est ainsi passée en quelques années du statut de croyance socialement fondée, d’idéal politiquement instable et d’objet scientifiquement insaisissable à celui de notion politiquement légitime, scientifiquement établie et socialement efficace” (Blondiaux, 1998: 278). L’extension de la méthode à d’autres pays se présente par la reproduction des conditions de l’apparition et du développement des sondages aux États-Unis. Et puis, l’histoire se répète:

Mêmes expériences limitées d’abord, même intérêt du public, même opposition des mêmes milieux, succès répétés, difficultés parfois, implantation solide, résurgence des mêmes méfiances à l’occasion des aléas politiques, diffusion par la presse en dépit de réticences persistantes, labeur des spécialistes et essor d’une technique qui fait ses preuves (Stoetzel et Girard, 1973: 56).

En 1938, la Grande-Bretagne fut le premier pays européen à bénéficier de la nouveauté américaine et des travaux de Gallup. La même année en France, le premier institut de sondages d’opinion, l’Institut français d’Opinion publique (Ifop), fait son apparition et son premier sondage porte sur les accords de Munich. L’Australie et le Canada ont suivi dès 1941 (Blondiaux, 1998: 290). Au lendemain de la guerre, les pays scandinaves, l’Italie et l’Allemagne se dotent des techniques, suivis par une autre vague de pays: l’Espagne, la Grèce et la Yougoslavie, les pays de l’Est, la Pologne, la Hongrie et la Roumanie (Stoetzel et Girard, 1973: 60).

## **1.3 Le contexte français**

### **1.3.1 La période pré-sondage**

En France, l’opinion publique se présente tout d’abord comme “une opinion sans public” (Blondiaux, 1998: 581). Les groupes susceptibles de s’intéresser aux sondages, soit les journalistes, les hommes politiques et les scientifiques, sont peu nombreux et surtout peu intéressés. Dans les années 30, l’importation de l’entreprise est pratiquement vouée à l’échec. La statistique dont cet instrument fait usage fascine dans un premier temps, mais en vient rapidement à être remise en question. Est-ce que les journalistes, les gens, sont



en mesure de s'approprier cette nouvelle information? Une fois la technique acceptée, elle doit être comprise, ce qui ne semble pas être le cas. Qu'est-ce qui explique l'échec en France des sondages d'opinion dans un premier temps et une réussite tardive qui en fait actuellement l'un des pays dans lequel il y a le plus de sondages d'opinion? La "sondomanie" actuelle en France n'a pas toujours été (Cowans, 2002: 72). C'est plutôt l'extrême lenteur des sondages d'opinion à s'implanter qui attire l'attention.

Alfred Max et Jean Stoetzel ont occupé une place prépondérante dans l'apparition des sondages en France, mais seul le second a réussi à se faire "reconnaître comme le découvreur des sondages" (Blondiaux, 1998: 290). En effet, la création de l'Ifop en 1938, l'investissement continu ainsi que son statut de professeur à La Sorbonne à partir de 1955 de Stoetzel lui permettront de se mériter le titre de père fondateur des sondages d'opinion en France.

Pour Stoetzel, l'étude de l'opinion publique passe nécessairement par la psychologie et par la sociologie. Les conduites d'un individu et les opinions formulées dépendront de facteurs institutionnels, interpersonnels et individuels. Selon Stoetzel, l'opinion privée relève des mécanismes psychologiques formant les attitudes profondes et l'opinion publique correspond au stéréotype auquel l'individu s'associe comme membre du corps social. En exprimant son opinion, l'individu laisse transparaître son attitude, ce qu'il pense, valorise, etc. La psychologie sociale permettra de théoriser l'affectif de l'homme en société.

Cowans (2002) qualifie de "sondophobie" les trente années où les hommes politiques et journalistes refusent de citer les résultats de sondage comme représentant de l'opinion publique. Cependant, il semblerait que la Libération de la France en 1944-45 a créé un espoir de croissance pour les sondages. À la fin de la Deuxième Guerre Mondiale, l'Ifop est joint par un second institut, le Service de sondages et statistiques (SSS). Cet institut publie ses résultats dans sa propre revue et parfois affiche une petite section dans les journaux. Malgré quelques ententes entre les sondeurs et la presse lors de la IV<sup>e</sup> République (1946-1958), il semble que la presse n'ait pas beaucoup fait usage des

sondages. Brûlé (1988) note diverses explications du sentiment d'ambivalence des journalistes à l'égard des sondages en France. Il évoque que les fonds qu'un service politique consacre à un sondage ne pourront servir à faire déplacer des collaborateurs, à étoffer l'équipe de journalistes, etc. Ensuite, Brûlé (1988) mentionne la possibilité que les journalistes politiques aient vu les sondages réalisés lors d'une campagne électorale comme des éléments limitant leur liberté d'expression sur ce qui se passe pendant la campagne. De plus, les intellectuels adoptent une attitude réticente envers la recherche empirique en sciences sociales. Bref, l'appropriation de la technique des sondages fut retardée par le manque d'intérêt de la part des journalistes, des intellectuels ainsi que des hommes politiques.

### **1.3.2 Le système politico-institutionnel**

En France, l'avènement de la V<sup>e</sup> République en 1958 entraîne éventuellement la légitimation des sondages d'opinion. Jusqu'en 1962, le président de la République était désigné au suffrage universel indirect, c'est-à-dire qu'il était élu par le corps électoral constitué d'élus et de délégués d'élus. Le dernier président nommé par ce mode d'élection fut le général de Gaulle. Il décrocha la victoire avec 78,5% des suffrages exprimés des 80 000 élus et délégués d'élus (Bréchon et Denni dans Bréchon (dir.), 2002: 19).

Lors de son premier mandat, De Gaulle exerça un pouvoir plus important que celui mentionné dans la Constitution. Dans l'hypothèse d'un second mandat, il souhaite légitimer le rôle du Président, qui est le sien, ainsi que limiter les pouvoirs des parlementaires contestataires de ses initiatives. En 1962, le nouveau mode de scrutin pour élire le président de la République est adopté par référendum et il est appliqué pour la première fois en 1965. Dorénavant le président de la République sera désigné au suffrage universel direct. D'après l'article 7 de la Constitution du 4 octobre 1958, approuvé le 6 novembre 1962:

Le Président de la République est élu à la majorité des suffrages exprimés. Si celle-ci n'est pas obtenue au premier tour de scrutin, il est procédé le deuxième dimanche suivant à un second tour. Seuls peuvent s'y présenter les deux candidats qui, le cas échéant après retrait de candidats plus favorisés, se trouvent avoir recueilli le plus grand nombre de suffrages au premier tour (Passelecq (dir.), 1995: 25).

L'avènement des sondages électoraux en France a lieu lors de l'élection présidentielle de 1965. C'est à ce moment précis que les sondages auraient gagné en légitimité. Gerstlé explique la situation: "Entre le début octobre et le début décembre 1965, les sondages montrent la baisse continue des intentions de vote (de 69 à 43%) pour le général de Gaulle et donc "préparent" l'opinion au ballottage" (dans Bréchon (dir.), 2002: 88). De plus, l'opération "estimation" - reposant sur le dépouillement de votes réellement émis - présentée sur Europe 1 participe à la légitimation de l'instrument en prédisant bien les résultats de l'élection. Pour donner une estimation fiable du résultat de l'élection dès 20h, les instituts prélèvent à l'avance un échantillon de bureaux de vote fermant à 18h et à 19h. Les responsables des instituts transmettent par téléphone les résultats des premières centaines de bulletins de vote et si l'opération se déroule bien, vers 18h40, les chiffres ne bougent plus beaucoup (Bréchon et Salomon dans Bréchon (dir.), 2002: 120). Alors que les journalistes politiques les plus chevronnés annonçaient que De Gaulle serait élu dès le premier tour, les sondeurs avaient vu juste: le général aurait à se présenter au second tour avant d'être nommé président de la République. Cette élection, la première au suffrage universel direct depuis la brève expérience de 1848 plus d'un siècle auparavant, favorisa le recours aux sondages d'opinion pour la mesure des intentions de vote.

Sous la IV<sup>e</sup> République (1946-1958), Ifop réalise des travaux fiables, mais sa place demeure très limitée. Comment les sondages forment-ils de véritables outils fonctionnels sous la V<sup>e</sup> République? Les sondages fournissent la possibilité de sélectionner les candidats qui se présenteront avant même l'élection, de mesurer la popularité des candidats, etc. Cet instrument permet également de prendre connaissance des rapports de force entre coalitions et à l'intérieur de celles-ci. En effet, les sondages d'opinion

procurent diverses informations permettant aux partisans d'effectuer des choix éclairés. Les sondages serviront également aux gouvernants qui pourront "mesurer le degré d'adhésion que leurs actes rencontrent, s'ils ne veulent pas laisser se creuser entre eux et le peuple un fossé qui minerait leur autorité" (Stoetzel et Girard, 1973: 236).

C'est en effet à partir de la V<sup>e</sup> République que les sondages deviennent de plus en plus importants et qu'ils alimentent les débats politiques. Leur visibilité n'est plus discutable, ils font maintenant la une de journaux tels *L'Express*, *France-Soir* et *Le Figaro*. La place faite aux sondages serait le reflet des démocraties occidentales selon Brûlé (1988). Il semblerait que plus les institutions donnent au corps électoral une prise directe sur le choix des dirigeants et sur la désignation des majorités, plus il y a de sondages. La France aurait à l'heure actuelle le record du monde des publications des sondages politiques. D'après Brûlé, "grâce à la combinaison de l'élection directe du chef de l'État, du scrutin majoritaire pour la désignation des députés et du recours possible au référendum" (Brûlé, 1988: 226), la France serait le pays offrant le plus d'emprise aux citoyens.

"L'élection du président de la République au suffrage universel est aujourd'hui un trait essentiel du système politique français, qui fait très largement consensus dans la classe politique et dans l'opinion" (Bréchon dans Bréchon (dir.), 2002: 9). Le premier tour de l'élection présidentielle permet de trancher entre les différents candidats qui s'opposent (gauche/droite) et parfois entre les prétendants à la direction d'une même coalition<sup>1</sup>. "Alors que le système institutionnel de la IV<sup>e</sup> République favorisait l'émergence de majorités parlementaires centristes, la V<sup>e</sup> République lamine le centre: celui-ci a été contraint de choisir entre les deux principales coalitions, qui se sont progressivement structurées autour de l'élection présidentielle. Cette dernière introduit dans le système politique de puissants effets de bipolarisation" (Bréchon dans Bréchon (dir.), 2002: 14). Lors du premier tour de l'élection présidentielle, des alliances se créent entre les partis. Ce système permettra à l'un des deux candidats au second tour de rassembler une

---

<sup>1</sup> Voir entre autres les luttes à droite entre Chirac et Barre en 1988 et entre Chirac et Baladur en 1995.

coalition pour diriger le pays et ainsi obtenir un pouvoir important.

### **1.3.3 Les élections présidentielles de 1965 à 1995**

En 1965, moment où eut lieu la première élection présidentielle au suffrage universel direct, le général de Gaulle, n'obtenant pas la majorité absolue au premier tour, dut se présenter au second tour contre le candidat Mitterrand. De Gaulle remporta l'élection avec 55,2 points de suffrage. Comme il a été mentionné auparavant, les sondages gagnèrent en légitimité en prédisant bien la tendance à la baisse du général de Gaulle au premier tour de l'élection présidentielle de 1965. De Gaulle démissionna en 1969 suite à l'échec du référendum sur la régionalisation et la réforme du Sénat tenue le 27 avril.

En 1969, Pompidou remporta la victoire au second tour avec 58,2 points. C'est à partir de ce moment que l'utilisation des sondages en campagne électorale devient plus intensive. La Sofres réalise sept sondages entre le référendum du 27 avril et l'élection de Pompidou le 15 juin (Gerstlé dans Bréchon (dir.), 2002: 89).

En 1974, suite au décès prématuré de Pompidou, c'est au tour de Mitterrand et Giscard de se disputer la présidence de la République. La lutte du second tour de l'élection présidentielle est serrée et Giscard devient président avec 50,8 points. En 1974, entre le 9 avril et le 5 mai, quinze sondages électoraux sont réalisés avant le premier tour, neuf sondages entre les deux tours, et ce, en plus des sondages réalisés tout au long de la campagne sur des sujets politiques (Bréchon et Salomon dans Bréchon (dir.), 2002: 115). Suite à cette lutte serrée entre Mitterrand et Giscard, un débat sur la réalisation et sur la publication des sondages amène la promulgation d'une loi visant entre autres à réglementer la publication et la diffusion des sondages en période électorale (voir la section 1.3.5.1).

Malgré les restrictions que la loi apporte, les sondages se multiplient. Une nouvelle instance, la Commission de sondages, exerce pour la première fois lors de l'élection de 1981, un contrôle sur la publication et la diffusion de 111 sondages présidentiels

(Bréchon et Salomon dans Bréchon (dir.), 2002: 116). Après 23 ans de pouvoir de la majorité présidentielle et parlementaire de la droite, l'élection présidentielle est remportée par le socialiste Mitterrand. Depuis le début des élections présidentielles au suffrage universel, cette victoire est la première de la gauche. Bréchon et Denni qualifient l'élection présidentielle de 1981 de "véritable rupture" (dans Bréchon (dir.), 2002: 20) avec les élections précédentes et identifient deux caractéristiques à cette élection. Selon eux, la "victoire" est non seulement celle de la gauche, mais ce mandat amènera également une première expérience de cohabitation<sup>2</sup> entre 1986 et 1988 (Bréchon et Denni dans Bréchon (dir.), 2002: 20).

En 1988, pour la première fois, le président sortant et le Premier ministre en exercice s'affrontent lors d'une élection présidentielle (Bréchon et Denni dans Bréchon (dir.), 2002: 46). Le 24 avril 1988 Mitterrand obtient 34,1% des suffrages exprimés au premier tour. Les principaux candidats de droite, Chirac et Barre recueillent respectivement 19,9% et 16,5% des suffrages exprimés (Bréchon et Denni dans Bréchon (dir.), 2002: 51). Au second tour, Mitterrand remporte l'élection avec 54,0% des suffrages exprimés, tandis que Chirac obtient 46,0% (Passelecq (dir.), 1995: 34). Les sondages se multiplient: la Commission des sondages contrôle 153 sondages en 1988.

En 1995, la droite reviendra au pouvoir suite aux deux mandats de la gauche, avec l'élection de Chirac qui recueille 52,6 points contre Jospin (47,4 points) (Passelecq (dir.), 1995: 34). La Commission des sondages contrôle 157 sondages durant cette campagne (Bréchon et Salomon dans Bréchon (dir.), 2002: 116).

#### **1.3.4 L'évolution du comportement électoral**

Plusieurs signes témoignent de l'évolution du comportement électoral "depuis le début des années 1990 sur fond de crise du système politique et des identifications partisans et idéologiques" (Perrineau et Ysmal dans Perrineau et Ysmal (dir.), 2003: 15). D'après

---

<sup>2</sup> La cohabitation correspond à la présence simultanée d'un gouvernement, d'une majorité parlementaire et d'un chef de l'État de tendances politiques opposées.

ces auteurs, le premier élément de l'évolution du comportement électoral est la montée de l'abstentionnisme<sup>3</sup> à partir des élections législatives de 1993 et de 1997, qui présentaient des taux d'abstention respectifs de 30,7% et de 31,5%. Ensuite, ce sont les élections présidentielles qui ont connu une baisse de participation. Les taux de participation des premiers tours présidentiels ont été de 82,0% en 1988 et 79,5% en 1995 (Perrineau et Ysmal dans Perrineau et Ysmal (dir.), 2003: 15). La même tendance se retrouve dans les élections régionales et européennes. La montée de l'abstentionnisme signifie que "d'une manière générale, un nombre croissant d'électeurs manifestaient leur éloignement par rapport au système politique et la structuration de celui-ci par les grandes forces politiques" (Perrineau et Ysmal dans Perrineau et Ysmal (dir.), 2003: 16).

Selon Perrineau et Ysmal (dans Perrineau et Ysmal (dir.), 2003), l'émergence des partis extrémistes fait également partie de l'évolution du comportement électoral. Dès 1984, le Front national s'installe et il accroît son influence au cours des années suivantes. Lors des élections législatives de 1988, le Front national obtient moins de 10% des suffrages exprimés. Les élections présidentielles de 1988 et 1995 ont présenté respectivement 14,4% et 15,3% des suffrages exprimés. Selon Perrineau et Ysmal (dans Perrineau et Ysmal (dir.), 2003), le recul du Front national lors des élections européennes est grandement associé à l'affrontement de deux candidats de l'extrême droite, Le Pen et Mégret, et de la présence du candidat souverainiste Pasqua, qui a attiré une partie de l'électorat du Front national.

Dans les années 90, ce fut au tour de l'extrême gauche de se développer. Perrineau et Ysmal notent que "jusqu'alors, les différents partis qui la composaient ne dépassaient pas 2% des suffrages exprimés" (dans Perrineau et Ysmal (dir.), 2003: 16). Toutefois, la candidate de Lutte ouvrière, Arlette Laguiller, recueillit 5,4% des suffrages exprimés lors de l'élection présidentielle de 1995. Les élections régionales de 1998 avec 4,2% et européennes de 1999 avec 5,2% confirmaient cette même tendance (Perrineau et Ysmal dans Perrineau et Ysmal (dir.), 2003: 16).

---

<sup>3</sup> Nous vous référons à Muxel qui traite de l'abstentionnisme (dans Perrineau et Ysmal (dir.), 2003 : p. 125-160).

Pendant cette période, des mouvements “catégoriels<sup>4</sup>” naissent, notamment le parti Chasse, pêche, nature et tradition. Ce parti a recueilli 4% aux élections européennes de 1989 et de 1994 et 6,9% en 1999 (Perrineau et Ysmal dans Perrineau et Ysmal (dir.), 2003: 17).

“Tous ces phénomènes traduisaient une instabilité et une fragmentation du système partisan ainsi qu’une tendance des électeurs à délaisser les partis de gouvernement” (Perrineau et Ysmal dans Perrineau et Ysmal (dir.), 2003: 17). Afin de comprendre les élections de 2002, Chiche et Reynié (2002) analysent l’évolution des forces électorales d’avril 1995 à décembre 2001. Selon eux aussi, “les caractéristiques actuelles de la participation électorale française révèlent un profond malaise civique” (Chiche et Reynié, 2002: 35) et ils en identifient trois traits majeurs: désengagement des électeurs, montée du refus de l’offre des partis et augmentation du vote pour les partis extrémistes.

### **1.3.5 La réglementation de la publication et diffusion des sondages**

#### **1.3.5.1 Loi n° 77-808 du 19 juillet 1977**

Au cours des années 70, les instituts et les sondages se multiplient. Leur présence s’avère appréciée, mais aussi critiquée par certains acteurs et une partie de l’opinion publique. Suite à des critiques portant sur le caractère privé du vote, sur l’influence éventuelle des sondages sur le vote des électeurs, sur les mauvaises prévisions des intentions de vote des candidats, etc., le Conseil constitutionnel et la Commission de contrôle de la campagne présidentielle recommanderont certaines mesures afin d’encadrer la réalisation et la publication des sondages. Le 19 juillet 1977, le Parlement vote une loi qui interdit la publication des sondages préélectoraux la dernière semaine avant le jour du scrutin. De plus, cette loi crée la Commission des sondages, chargée de surveiller la publication et la diffusion des sondages d’opinion. Le mandat de la

---

<sup>4</sup> Ce type de mouvement naît d’une thématique particulière. Par exemple, le parti Chasse, pêche, nature et tradition est né de l’enjeu de la chasse et de la pêche. Ensuite, le parti s’est orienté sur les enjeux plus généraux de la “ruralité” et de l’“antimondialisation” (Bussi dans Perrineau et Ysmal (dir.), 2003 : 313).



Commission des sondages consiste à faire respecter cette loi et à vérifier les pratiques méthodologiques des instituts. En théorie, tous les sondages en lien direct ou indirect avec une élection demeurent susceptibles d'être contrôlés. Dans la pratique, la Commission des sondages contrôle les sondages sur les intentions de vote, en tout temps, et les sondages portant sur les questions politiques, pendant les trois mois précédant le scrutin. La deuxième section de la loi n° 77-708 du 19 juillet 1977 relative à la publication et à la diffusion de certains sondages d'opinion stipule que :

Art. 2. La publication et la diffusion de tout sondage tel que défini à l'article 1<sup>er</sup> doivent être accompagnées des indications suivantes, établies sous la responsabilité de l'organisme qui l'a réalisé:

Le nom de l'organisme ayant réalisé le sondage;

Le nom et la qualité de l'acheteur du sondage;

Le nombre de personnes interrogées;

La ou les dates auxquelles il a été procédé aux interrogations.

Art. 3. - À l'occasion de la publication ou de la diffusion de tout sondage tel que défini à l'article 1<sup>er</sup>, l'organisme qui l'a réalisé doit procéder au dépôt auprès de la commission des sondages instituée en application de l'article 5 de la présente loi d'une notice précisant notamment:

L'objet du sondage;

La méthode selon laquelle les personnes interrogées ont été choisies, le choix et la composition de l'échantillon;

Les conditions dans lesquelles il a été procédé aux interrogations;

Le texte intégral des questions posées;

La proportion des personnes n'ayant pas répondu à chacune des questions;

Les limites d'interprétation des résultats publiés;

S'il y a lieu, la méthode utilisée pour en déduire les résultats de caractère indirect qui seraient publiés.

La commission des sondages peut ordonner la publication par ceux qui ont procédé à la publication ou à la diffusion d'un sondage tel que défini à l'article 1<sup>er</sup> des indications figurant dans la notice qui l'accompagne ou de certaines d'entre elles (Articles 2 et 3 de la loi n° 77-708 du 19 juillet 1977).

### 1.3.5.2 Loi n°2002-214 du 19 février 2002

À un certain moment, la loi n° 77-708 du 19 juillet de 1977 est apparue dérisoire puisque les citoyens accédaient, grâce à Internet et à la presse étrangère, aux sondages préélectoraux pendant la semaine précédant le scrutin. Suite à ces événements, l'interdiction française perdit tout son sens. C'est à partir du 19 février 2002 que la loi est modifiée. L'interdiction de publication se limitera dorénavant au jour précédant chaque tour du scrutin et au jour même du scrutin. De plus, le droit de toute personne à consulter la notice est prévu par l'article 3 de la loi n° 2002-214 du 19 février 2002. En 2002, les citoyens auront accès aux sondages plus tardivement dans la campagne électorale et donc, normalement, à plus de sondages. De plus, les citoyens ont théoriquement accès à l'information méthodologique des sondages à la Commission des sondages. En pratique, l'accès aux fiches techniques demeure limité<sup>5</sup>.

### 1.3.6 Synthèse

Au cours des années 20 et 30, la définition de la notion d'opinion publique ne fait pas consensus parmi les scientifiques. L'étude de l'opinion publique connaît un progrès aux États-Unis dès les années 20. Certains intellectuels américains, plus particulièrement les psychologues et les sociologues, étudient les attitudes et la théorisation de l'affectif de l'homme. Les statisticiens participeront également à la légitimation des sondages d'opinion. Par contre, les intellectuels français demeureront réfractaires à ce courant sociologique jusqu'aux années 50. Les sondages d'opinion ayant connu leur coup d'envoi en 1936 aux États-Unis, devront attendre l'élection présidentielle de 1965, en France.

Contrairement à la presse française, la presse américaine publie et diffuse des sondages d'opinion aux États-Unis à partir de 1935 et cette nouvelle méthode devient une forme de journalisme à part entière (Blondiaux, 1998: 265-266). Selon Brûlé, les rapports entre les journalistes et les sondeurs trouvèrent leur équilibre lors de "l'arrivée d'une nouvelle

---

<sup>5</sup> Nous reviendrons plus tard sur l'explication de l'accès aux données à la Commission des sondages.

génération de journalistes, qui, loin de voir dans les chiffres des sondages une concurrence, surent y puiser des éléments pour étayer leurs propres analyses et pour aider leurs lecteurs à mieux déchiffrer la situation” (Brûlé, 1988: 113). Selon cet auteur, il semblerait que la “bataille” entre les journalistes et les sondeurs soit révolue puisque tous les grands médias ont recours aux sondages à un moment ou à un autre.

Plusieurs éléments ont retardé la légitimation des sondages d’opinion en France: la quasi-absence de spécialistes et d’intellectuels s’intéressant aux sondages, le faible apport empirique dans l’enseignement des sciences sociales, le peu d’intérêt de la presse et des hommes politiques, et ce, pendant une trentaine d’années après l’apparition des sondages en 1938. Ce fut lors de l’élection présidentielle de 1965 que les sondages ont obtenu une légitimité. À partir de 1965, les hommes politiques, les scientifiques et les médias ont commandé des sondages, permis le développement de la méthode et lui ont accordé une place de choix dans les débats publics. De plus, la situation politique française a favorisé le recours aux sondages afin de mesurer les rapports entre les coalitions et à l’intérieur de celles-ci, les cotes de popularité des gouvernants, les intérêts des citoyens, etc. C’est ainsi que la France au départ réfractaire à la méthode est devenue un pays recourant fortement aux sondages.

En France, la baisse de participation, l’émergence des partis extrémistes de droite comme de gauche et la montée des mouvements catégoriels sont des traits majeurs du comportement électoral. Avant de procéder à l’analyse des sondages du premier tour de l’élection présidentielle de 2002, il s’avère nécessaire d’examiner de plus près les principaux cas types d’erreurs de prédiction électorales connues dans différents pays.

## CHAPITRE 2 - Sondages et élections: quelques cas types

Les sondages électoraux présentent un avantage certain comparativement aux études de marché; ils offrent la possibilité de vérifier la qualité des estimations par rapport aux résultats. Comme les sondages électoraux suscitent beaucoup d'intérêt, leurs performances sont largement discutées. Les spécialistes des sondages identifient quelques cas types de grandes erreurs de prédiction électorales. Ce chapitre en présentera les principaux cas types ainsi que les raisons qui ont été avancées pour les expliquer. Deux types d'élections seront abordées: les élections législatives - Grande-Bretagne, Irlande et Canada (Québec) - et les élections présidentielles - États-Unis et France. La section sur la France toutefois abordera l'ensemble des performances des sondeurs des élections présidentielles de 1965 à 1995 puisque ce portrait est essentiel à la compréhension du cas français de 2002.

### 2.1 Élections législatives

L'élection législative correspond à désignation des députés de l'Assemblée nationale par le suffrage universel. Selon le résultat du vote, le chef de parti ayant le plus de députés deviendra le Premier ministre, le chef de la majorité parlementaire, le responsable devant le Parlement. Les principales erreurs de prédiction documentées sont survenues lors des élections législatives de la Grande-Bretagne (1970, 1992, 1997), de l'Irlande (2002) ainsi que du Québec (1998) et du Canada (2000).

#### 2.1.1 Grande-Bretagne

Lors des élections générales de Grande-Bretagne en 1970, "Gallup Poll sous-estime les conservateurs de 4,5%, sur-estime les travaillistes de 4,9%, au total une méprise de 9,4% sur l'écart qui sépare les deux formations" (Bon, 1974: 150). Tous les instituts de sondage se sont trompés sur le parti vainqueur. Selon Bon (1974), l'explication de

l'erreur se trouve chez les "sans réponse". Cette catégorie de répondants qui paraissait stable et faible (8 à 10% selon les enquêtes) depuis le début de la campagne a été attribuée en part égale à chaque parti. Et puis, la réalité fut tout autre, l'abstention élevée (28,5%) et la démobilisation d'un électorat travailliste trop certain de sa victoire entraîne la victoire des conservateurs sur les travaillistes selon Bon (1974: 164).

Lors des élections générales britanniques de 1992, "les sondages d'intentions de vote accordaient en effet en moyenne une légère avance de 1,5 points aux travaillistes, avec la perspective d'un Parlement sans majorité, alors que les conservateurs ont en fait gagné avec une marge de 7,6 points et une majorité absolue aux Communes de 21 sièges" (Jadot, 2001: 111). Tous les sondages électoraux ont inversé le parti vainqueur, y compris les sondages menés la journée même du scrutin, à la sortie des urnes, et pourtant ces sondages mesurent les comportements passés et sont donc "*a priori* insensibles à l'impact d'un changement massif "de dernière minute"" (Jadot, 2001: 112). Ces élections se sont avérées un véritable échec pour les sondeurs.

Suite à cet événement, la *Market Research Society* (MRS) a décidé d'étudier les causes de cet échec avec l'aide de l'ensemble des instituts. Cette enquête s'échelonnant sur deux ans a débouché sur un ensemble de conclusions: un certain nombre d'explications possibles ont été invalidées, d'autres, considérées peu importantes, et des explications majeures ont été validées par la MRS (Jadot, 2001). Parmi les éléments étudiés, la taille des échantillons, le vote des expatriés (droit de vote accordé par les Conservateurs en 1997), les effets des techniques d'enquête (face à face, téléphonique, etc.) ainsi que le mensonge massif et délibéré aux instituts de sondage ont été invalidés.

La MRS a retenu deux facteurs peu importants dans l'explication de l'échec des sondages en 1992. Le premier facteur mis en cause a été le choix des points d'enquête, "4 des 5 instituts opérant dans des circonscriptions légèrement plus travaillistes que la moyenne (à cause de l'évolution démographique depuis leur sélection)" (Jadot, 2001: 113). Toutefois, Jadot (2001) explique que ce biais est au maximum de 1% et en plus le cinquième institut enquêtant dans des circonscriptions non-biaisées n'a pas obtenu de

meilleurs résultats que les quatre autres sondeurs. Le second facteur peu important dans l'explication de l'échec des sondages de 1992 est l'interrogation de personnes non inscrites sur les listes électorales. "La non-inscription, volontaire ou non, n'aurait eu un fort impact que si les citoyens concernés et interrogés à tort avaient été disproportionnellement enclins à déclarer aux sondeurs vouloir voter pour les travaillistes" (Jadot, 2001:114). Il semblerait que plusieurs indices infirment cette théorie dont le fait même que les instituts qui filtraient les non-inscrits n'ont pas obtenu de résultats différents des instituts ne filtrant pas.

Parmi les trois explications majeures et d'importance équivalente retenues par la MRS, la première tient aux défauts des plans de sondages et des quotas. D'une part, les caractéristiques de la structure de l'échantillon choisi ne représentent pas celle de la population et ce biais tend à réduire la proportion d'électeurs se rapprochant des conservateurs (Jadot, 2001; Curtice et al., 1997 et Jowell et al., 1993). D'autre part, les pondérations effectuées étant faiblement corrélées avec le vote n'ont pas permis de corriger les biais de représentation politique. La seconde explication est celle du changement d'opinion de dernière minute (*late swing*) et elle englobe plusieurs aspects peu favorables au parti travailliste : les courbes de popularité et les enjeux, les changements d'avis de dernier moment suite aux effets de campagne et l'importance des électeurs flottants ou indécis (Jadot, 2001; Curtice et al., 1997 et Jowell et al., 1993). D'après Jadot (2001), l'hypothèse de la participation différentielle, c'est-à-dire les répondants déclarant leur intention de vote qui finalement s'abstiennent le jour du scrutin ou au contraire ceux disant s'abstenir au moment du sondage qui voteront, n'a pas été validée.

La troisième et dernière explication concerne les refus différentiels selon la préférence partisane. Selon Jadot, la théorie de la "spirale du silence" de Noëlle-Neumann (1984) permet d'expliquer la non-réponse aux sondages. Cette théorie permet de comprendre l'écart entre les estimations enregistrées par les sondeurs et les résultats de l'élection. Dans le cas présent, la perception est que les électeurs conservateurs ont eu tendance à dissimuler leur intention de voter pour ce parti. Les conservateurs n'étaient plus "à la

page”, et cette situation renforçait le silence de leurs sympathisants (Jadot, 2001: 115). Donc, une sous-estimation des votes conservateurs a été produite par le refus de répondre à la question d'intention de vote, par le refus de participer à l'enquête et par la tendance de ces électeurs à se réfugier dans la réponse “ne sais pas”. La sous-déclaration et la sous-représentation sont des éléments clés de la compréhension de l'échec de l'élection britannique de 1992.

L'analyse de l'échec de l'élection de 1992 de la MRS a entraîné certains changements méthodologiques. À moins d'un an de l'élection générale de 1997, les instituts utilisent des méthodologies diversifiées, et les différences de résultats sont marquées et systématiques, ce qui devrait permettre l'évaluation des méthodes. Selon le premier bilan de 1997, “certes, tout au long de la campagne, les sondages ont rempli leurs deux “missions”: dire qui remporterait l'élection et signaler un raz-de-marée électoral travailliste” (Jadot, 2001: 123). Toutefois, les instituts signalent des marges de victoire des travaillistes de 10 à 22 points, et le résultat de l'élection sera de 13 points (Jadot, 2001: 123). Contrairement à l'élection générale de 1992 où les conservateurs étaient sous-estimés, l'élection de 1997 se caractérise par le départage du “vote anti-conservateur entre travaillistes et libéraux-démocrates dans une situation politique atypique” (Jadot, 2001: 123) et la surestimation du vote travailliste demeure. Ni l'élection générale de 1997 ni son analyse “n'ont permis d'établir avec certitude laquelle des méthodes serait la plus adéquate” (Jadot, 2001:124).

Curtice (1997) explique que l'élection générale de 1997 a permis de rétablir la réputation des sondages d'opinion, mais elle laisse tout de même certaines questions. Il soutient que l'analyse des refus est nécessaire à l'amélioration des estimations, puisque selon lui, il est impossible d'assumer que les intentions de vote des répondants refusant de dévoiler les intentions de vote soient proportionnelles à celles des répondants les déclarant.

### **2.1.2 Irlande**

En 2002, les derniers sondages réalisés par les deux principaux instituts irlandais, *Irish*

*Marketing Surveys* (IMS) et *Market Research Bureau of Ireland* (MRBI), ont surestimé le *Fianna Fail* de 8 et 3 points respectivement et sous-estimé le *Fine Gael* de 5 et 2 points (McElroy et Marsh, 2003). Les sondeurs ont bien prédit le parti gagnant lors des élections législatives irlandaises, mais ils ont surestimé la marge de victoire du *Fianna Fail*. McElroy et Marsh (2003) notent que la sous-estimation systématique pour les partis de droite liée aux refus d'être interviewé ainsi qu'aux répondants difficiles à joindre ne s'applique pas dans le cas de l'Irlande 2002 puisque les deux partis en tête sont à droite de l'échiquier et que l'un est surestimé (*Fianna Fail*) et l'autre, sous-estimé (*Fine Gael*). Parmi les sources d'erreurs identifiées par les auteurs, il semblerait que le changement de dernier moment (*late swing*), les différents changements d'idée et les préférences différentielles des indécis expliquent faiblement les erreurs de prédiction. Selon McElroy et Marsh (2003), les erreurs de prédictions s'expliquent par l'échantillonnage. L'amélioration de la qualité des prédictions serait possible grâce à la pratique de rappels afin d'augmenter le taux de réponse. Les petits échantillons ainsi que le faible nombre de points d'enquête sont également des éléments exigeant une certaine attention.

### 2.1.3 Québec/Canada

En 1998, les sondages préélectorales lors des élections législatives québécoises ont suscité une importante polémique. Les sondages prédisaient une victoire facile du Parti Québécois (PQ), parti de centre-gauche souverainiste, sur le Parti Libéral (PLQ), parti de centre-droit fédéraliste. Finalement les résultats de l'élection furent bien différents. Des six derniers sondages donnant 5 points d'avance au Parti Québécois, cinq prédisent la victoire du Parti Québécois par 5 à 10 points et un sondage prévoit la victoire du Parti Libéral (Durand et al., 2001: 108). Parmi l'ensemble des sondages publiés durant la campagne, onze sondages sur quatorze indiquent une avance de 1 à 10 points pour le PQ, mais les résultats du vote furent tout autres (Durand et al., 2001: 108). Le PQ gagne la majorité des circonscriptions, mais le PLQ remporte le vote avec 44% contre 43% des suffrages exprimés pour le PQ.



Durand et al. (2001) mènent un sondage post-électoral auprès des répondants. Ils concluent à certaines explications des erreurs de prédiction de 1998. D'une part, ils soutiennent que l'exclusion des gens n'apparaissant pas sur les listes de téléphone (personnes ayant un numéro confidentiel) ou vivant dans un logement collectif, entraîne le retrait substantiel d'une proportion de la population ayant des caractéristiques spécifiques relatives au vote. Les gens vivant en institution voteraient plus à droite et donc seraient davantage en faveur du PLQ. D'autre part, les erreurs de prédiction sont dues en partie aux répondants qui refusent de révéler leur intention de vote ou qui disent ignorer pour qui ils voteront. En fait, ces derniers ne se répartissent pas de la même façon que les répondants acceptant de dévoiler leur intention et ils sont plus souvent à droite de l'échiquier. En résumé, les biais de l'échantillon et de la non-réponse dans le cas des élections québécoises de 1998 entraînent systématiquement une sous-estimation du PLQ et une surestimation du PQ.

En 2000, les sondages préélectoraux auraient bien performé à l'échelle nationale lors des élections canadiennes. Toutefois, Durand (2002) note une sous-estimation du Parti Libéral au Québec, semblable à celle observée lors des élections québécoises de 1998. Au Canada, les six derniers sondages prédisaient pour les cinq principaux partis des écarts de plus ou moins 0,2 à 0,5 points avec les résultats de l'élection, tandis qu'au Québec, les cinq derniers sondages sous-estiment de 4,2 points le Parti Libéral et les prédictions des autres partis se trouvent surestimées de 1,4 points pour l'Alliance Canadienne, de 1,5 points pour le Bloc Québécois ainsi que de 1,2 points pour le Nouveau Parti Démocratique. Le Parti Conservateur fut exactement estimé. La sous-estimation du Parti Libéral au Québec en 2000 semble une réplique de celle des élections provinciales de 1998.

## **2.2 Élections présidentielles**

L'élection présidentielle permet de désigner le président de la République, le chef de l'État et le chef de gouvernement au suffrage universel direct. Les sondages électoraux contestés de certaines élections présidentielles ont attiré l'attention. En premier lieu, les

élections de 1948, 1996 et 2000 aux États-Unis seront présentées. En second lieu, l'ensemble des élections présidentielles françaises de la V<sup>e</sup> République seront exposées.

### 2.2.1 États-Unis

Lors de l'élection présidentielle américaine de 1948, le président sortant, Truman, affronte le républicain Dewey, donné vainqueur par les principaux instituts américains de sondage Gallup, Crossley et Roper. Cependant, le jour du scrutin, Truman l'emporte avec 49,5% des suffrages exprimés et Dewey obtient 45,9% des suffrages exprimés (Passelecq (dir.), 1995: 46). Les sondeurs ayant cessé d'interroger l'électorat environ trois semaines avant l'élection ont en moyenne minimisé le score de Truman de 4,5% (Bon, 1974: 152). Afin d'éviter un échec semblable, une enquête sur les causes des erreurs est entreprise et certaines recommandations sont faites par le *Social Science Research Council* (SSRC). La première faute soulevée est liée à la date hâtive des derniers sondages publiés. Les sondeurs ont négligé les changements d'opinion dans les dernières semaines de la campagne. La seconde faute est la négligence des enjeux des dernières semaines de la campagne électorale qui se sont avérés en faveur de Truman. La troisième faute: "les instituts, persuadés à tort de la fermeté des intentions de vote, ont négligé les abstentionnistes, les indécis, les fluctuants" (Bon, 1974: 154). Les sondeurs ont ainsi mesuré les intentions de vote des répondants ne sachant pas dans quelle proportion ceux-ci allaient véritablement voter ou ne connaissant pas la fermeté des intentions de vote. Finalement, la dernière faute soulevée par le SSRC concerne la constitution de l'échantillon. L'utilisation des quotas a entraîné une sous-représentation des gens les moins faciles à rejoindre (paysans et gens à bas revenus ou peu instruits) et ces catégories se sont avérées plus favorables à Truman.

Plus récemment, toujours aux États-Unis, certaines élections présidentielles ont attiré l'attention. En 1996, tous les sondeurs prédisent bien la victoire du démocrate Clinton. Sept sondeurs sur neuf surestiment la marge de victoire<sup>6</sup> (Mitofsky, 1998: 242). L'écart

---

<sup>6</sup> Un institut sous-estime l'écart entre Clinton et Dole d'un point et un institut prédit bien la marge de victoire de Clinton.

entre Clinton et Dole étant de 9 points à l'élection, huit instituts sur neuf donnent des écarts de 0 à 5 points de la marge de victoire de Clinton sur Dole et un institut surestime cet écart de 9 points (Mitofsky, 1998: 242).

L'élection présidentielle américaine de 2000 présente une situation différente de l'élection de 1996. Face à une lutte serrée entre le candidat démocrate Gore et le républicain Bush, les derniers sondages donnaient 14 fois sur 19 Bush gagnant, 2 fois sur 19 Gore gagnant et 3 fois sur 19 les candidats à égalité (Traugott, 2001: 392). Le jour du vote, Gore gagne le vote populaire avec 48,4 points contre 47,9 points pour Bush. Toutefois, la lutte demeure serrée pour gagner les circonscriptions. Bush devint le président des États-Unis en gagnant le vote dans l'État de la Floride. Selon Traugott (2001), les sondages ont généralement bien estimé les résultats du scrutin malgré un vainqueur mal prédit.

### **2.2.2 France**

En France, la première élection présidentielle au suffrage universel direct de 1965 - depuis la brève expérience de 1848 - a entraîné la légitimation des sondages d'opinion. Puisque l'élection présidentielle est le moment fort de la politique française, celui qui mobilise le plus d'électeurs, et que ces derniers exercent le suffrage universel direct depuis 1965, c'est à partir de cette période que seront étudiés les sondages d'intention de vote. Comme la présente étude porte sur le premier tour présidentiel de 2002, seulement les élections présidentielles seront présentées. De plus, étant donné que les sondages sur les seconds tours n'ont jamais eu de grands écarts avec les résultats des élections, ce sont les sondages des premiers tours qui seront présentés. En disposant des résultats du premier tour, le choix dichotomique du second tour est beaucoup plus simple à prédire que la prévision des intentions de vote des premiers tours (Bréchon et Salomon dans Bréchon (dir.), 2002: 126). Les résultats de sondages obtenus à partir de redressements opérés sur le souvenir du vote au premier tour permettent en effet d'obtenir de bonnes prédictions (Gazier et Abraham, 1995: 144).

En 1965, les sondages présidentiels auraient prédit les intentions de vote à 1 point près pour le général de Gaulle, mais le candidat de gauche Mitterrand est sous-estimé de 4,7 points et le candidat du centre Lecanuet surestimé de 4 points dans le dernier sondage (Bréchon et Salomon dans Bréchon (dir.), 2002: 126).

En 1969, il existe des écarts pour presque tous les candidats dans le dernier sondage Ifop; le gaulliste Pompidou et le communiste Duclos sont sous-estimés de 3 et 3,5 points tandis que le candidat de centre Poher et le candidat de l'extrême gauche Rocard sont tous deux surestimés de 1,5 points et le candidat de gauche Defferre de 4 points (Bréchon et Salomon dans Bréchon (dir.), 2002: 126).

En 1974, élection anticipée de deux ans suite au décès de Pompidou, Bréchon et Salomon notent une légère surestimation du socialiste Mitterrand et une sous-estimation du candidat de centre Giscard par rapport au dernier sondage publié à deux jours du scrutin (dans Bréchon (dir.), 2002: 127). Le pronostic et la réalité sont plus similaires qu'en 1969.

En 1981, "l'écart n'est [...] pas très fort avec la réalité des urnes, sauf en ce qui concerne François Mitterrand, sous-estimé d'environ 4 points alors que Georges Marchais [le candidat communiste] était surestimé de 3" (Bréchon et Salomon dans Bréchon (dir.), 2002: 127). Stoetzel (1982) mentionne que les instituts de sondages ont fait un bon travail lors des élections françaises de 1981. Cependant, il semblerait que le public français et étranger fut surpris de l'élection du candidat socialiste Mitterrand au poste de Président de la République "après 23 ans de pouvoir sans partage de la droite puisque celle-ci disposait à la fois de la majorité présidentielle et parlementaire" (Bréchon et Denni dans Bréchon (dir.), 2002: 20). "Ce qui venait de se produire aurait pu être lu six mois à l'avance dans les résultats des sondages si le public français, les milieux politiques et les personnels des organes de l'information collective avaient su lire" (Stoetzel, 1982: 3). Les sondages d'intention de vote précédant le premier tour de l'élection présidentielle, en sous-estimant le candidat de gauche Mitterrand, prévoient un plus grand écart entre les intentions de vote du candidat de centre Giscard et

Mitterrand, les candidats au second tour. Quant aux sondages portant sur le second tour, ils prédisaient bien le vainqueur Mitterrand, et ce, environ trois semaines avant le vote effectif<sup>7</sup>. Face à la surprise de 1981, Stoetzel (1982) mentionne deux types d'explication. D'une part, il reproche au public de ne pas avoir perçu les tendances. Pourtant, Stoetzel (1982) explique que de multiples graphiques et tableaux ont été publiés sur des questions politiques montrant le déclin de Giscard dans l'opinion publique durant la campagne électorale. Selon Stoetzel (1982), si le public avait observé les sondages dans leur ensemble, il aurait été possible pour lui de saisir le déclin de Giscard à l'avantage de Mitterrand et au contraire, le fait de voir les sondages un à un a pu égarer le public. D'autre part, l'attention du public a pu être détournée de la réalité par la préférence de certains médias pour Giscard (Stoetzel, 1982: 8). La Commission des sondages note également que les médias dans leur course au sensationnalisme titrent sur le classement des candidats et négligent la marge d'erreur inhérente aux sondages (Gazier et al., 1995: 142).

En 1988, depuis sa percée lors des élections européennes en 1984, l'extrême droite de Le Pen prend de l'importance pour la première fois sur l'issue du scrutin en recueillant 14,4% des suffrages exprimés. La surestimation d'un peu plus de 3,5 points de Chirac et la sous-estimation de 2 points de Le Pen dans le dernier sondage Sofres non-publié attirent l'attention (Bréchon et Salomon dans Bréchon (dir.), 2002: 127).

En 1995, durant la campagne électorale précédant le premier tour de l'élection présidentielle, les sondeurs ont surestimé les intentions de vote du candidat de droite Chirac et ils ont sous-estimé les candidats de gauche Jospin et de droite Balladur. Selon le sondage de la Sofres à la veille de l'élection, Chirac se situait à 24 points<sup>8</sup> et finalement il obtint un score décevant de 20,8 points au moment du scrutin, en deçà du candidat socialiste Jospin (Bréchon et Salomon dans Bréchon (dir.), 2002: 127). Aucun

---

<sup>7</sup> Nous renvoyons aux résultats des sondages du second tour de l'élection présidentielle de 1981 de l'Ifop (18 avril) et de la Sofres (17 avril) (Stoetzel, 1982 : 5).

<sup>8</sup> Dans les derniers sondages publiés à huit jours de l'élection, Chirac était estimé entre 26 et 27,5% (Bréchon et Salomon dans Bréchon (dir.), 2002 : 127).

sondage n'avait prévu l'arrivée de Jospin au premier rang; il était annoncé à 20,5 points dans le sondage de la Sofres à la veille du scrutin et, le jour du scrutin, il a obtenu 23,3% des suffrages exprimés (Bréchon et Salomon dans Bréchon (dir.), 2002: 127). Pour sa part, Balladur (deuxième candidat de droite) a fait mieux que ce que les prédictions annonçaient: il obtint 18,6 points comparativement à l'estimation de 16,5 points de la Sofres à la veille du scrutin (Bréchon et Salomon dans Bréchon (dir.), 2002: 127). L'inversion Chirac/Jospin à la tête du premier tour provoqua "une volée d'accusations et de critiques (...) sur les dirigeants des instituts qui, les uns après les autres, tentèrent de se défendre tout en s'interrogeant sur les faiblesses de l'outil qu'ils manient et en insistant sur les difficultés de l'indispensable opération de redressement" (Gazier et al., 1995: 144). L'inconvénient majeur dans le cas de cette élection ne se trouve pas dans les écarts de prévision, mais bien dans l'ordre des candidats. De plus, selon les sondages de la Sofres<sup>9</sup>, le candidat de l'extrême droite Le Pen qui a obtenu 15,0% des suffrages exprimés le jour du scrutin, demeure sous-estimé tout au long de la campagne.

En résumé, ce chapitre a permis de prendre connaissance des cas types d'erreurs de prédiction électorales de quelques pays et d'étudier la performance des sondeurs lors des élections présidentielles françaises de 1965 à 1995. Contrairement à d'autres cas, le biais pro-gauche n'est pas typique de l'histoire des sondages présidentiels en France. C'est davantage la surestimation de la droite et la sous-estimation de la gauche qui caractérisent les estimations des instituts français. Lors du premier tour de l'élection présidentielle en 2002, les pratiques "sondagières" des instituts français peuvent-elles être remises en cause?

---

<sup>9</sup> Voir le graphique des intentions de vote au premier tour de l'élection présidentielle 1995 (dans Bréchon (dir.), 2002: 134).

### CHAPITRE 3 - Problématique

Systematiquement, certaines erreurs de prédiction semblent se reproduire dans divers pays, et ce, malgré les différentes méthodes utilisées. Un premier type d'erreur de prédiction classique est la surestimation de la gauche et la sous-estimation de la droite. À ce propos, Durand, Blais et Vachon, indiquent cette erreur de prédiction en notant une surestimation du courant "*labor-democrat-left*" et une sous-estimation du courant "*conservative-republican-right*" dans les sondages (Durand et al., 2001: 109) . En France, Rivière (2001) propose une analyse des sondages d'intentions de vote par villes aux élections municipales de 2001 et il explique que ce sont quasi systématiquement, à l'exception d'Aurillac, les points attribués en trop à la gauche plurielle qui manquent à la droite républicaine, malgré les écarts plus ou moins importants des estimations de sondage et des résultats. Toutefois, cette situation va à l'encontre des élections présidentielles françaises mentionnées au chapitre précédent.

Une seconde erreur de prédiction largement acceptée est la sous-estimation par les sondages électoraux des partis politiques manifestant des opinions extrêmes (Breen, 2000: 375). Malgré l'attribution d'une "performance très honorable" aux sondages électoraux français, il demeure "une certaine difficulté à saisir correctement le poids électoral des formations extrêmes, qu'elles soient de droite ou de gauche" selon Brûlé (1988: 20). En France, la difficulté à estimer les formations se trouvant aux extrêmes de l'éventail politique est connue (Brûlé, 1988). Certaines études notent une sous-estimation de l'extrême droite et, particulièrement du vote "frontiste" (ceux qui votent pour le parti Front national) (Rivière, 2001).

Dans l'étude des résultats bruts des sondages d'intention de vote aux élections régionales françaises de 1998, Boy et Chiche (1999a) notent "une sous-estimation des listes regroupant des candidats "divers" ou "catégoriels" (il s'agit notamment des chasseurs)" (Boy et Chiche, 1999a: 247).

Ce troisième chapitre permet de faire le point et de poser les explications possibles des erreurs de prédiction au premier tour de l'élection présidentielle française de 2002. Tout d'abord, les explications d'ordre psychosociologique et méthodologique proposées par les auteurs permettront de tracer un portrait des explications possibles des erreurs de prédiction. D'une part, certains auteurs tels Durand et al. (2001, 2002), Curtice (1997) et Jowell et al. (1993) attribuent à certaines catégories de répondants une partie des erreurs de prédiction. De plus, certains auteurs tels Lau (1994), Curtice (1997), Curtice et Sparrow (1997), Durand et al. (2001, 2002) ainsi que Vachon et al. (1999) soutiennent que la méthodologie influence la qualité des estimations. Ensuite, ce qui s'est produit lors du premier tour de l'élection présidentielle française de 2002 sera présenté. La présentation de la méthodologie utilisée par les instituts français permettra de poser des hypothèses de recherche quant à l'explication possible des erreurs de prédiction du premier tour de l'élection présidentielle française en 2002. Enfin, la dernière partie présente les objectifs de cette recherche.

### **3.1 Explications des erreurs de prédiction**

Selon Durand et al. (2001, 2002), les erreurs de prédictions peuvent être attribuées à deux causes. La première explication est de type psychosociologique; elle relève des répondants eux-mêmes, de leurs comportements, et de leurs attitudes. La seconde explication est de type méthodologique; elle correspond à l'influence de l'échantillon et de la gestion de terrain sur la qualité des estimations.

Toutefois, les facteurs méthodologiques et les facteurs psychosociologiques causant les erreurs de prédiction dans les sondages se recoupent dans une large mesure. En fait, les méthodes utilisées par certains sondeurs et négligées par d'autres, ont pour objectif de corriger les biais que peuvent provoquer les facteurs psychosociologiques. Un exemple illustrant cette affirmation est le cas où un questionnaire porte sur un sujet délicat pouvant entraîner un effet de désirabilité chez le répondant face à l'enquêteur. Afin de contrer le biais que peut provoquer ce facteur psychosociologique et ainsi minimiser l'"effet de prestige", le mode de passation du questionnaire aura avantage à être auto-



administré plutôt qu’être face à face pour améliorer la qualité des estimations. Malgré le fait que les facteurs méthodologiques et psychologiques se recoupent, les facteurs explicatifs des erreurs de prédiction seront présentés séparément.

### 3.1.1 Explications psychosociologiques

L’étude de diverses catégories de l’électorat permet la compréhension de certaines erreurs de prédiction. Une première catégorie est celle des répondants qui changent d’idée au dernier moment. Ces répondants peuvent décider de s’abstenir ou voter pour un candidat ou parti différent de celui mentionné lors du sondage. Selon Jowell et al. (1993), cette hypothèse du “*late swing*”, changement de dernier moment, peut expliquer une partie de l’erreur de prédiction lors de l’élection générale britannique de 1992. L’hypothèse est toutefois invalidée dans le cas de l’élection québécoise de 1998, puisque 13% des gens changent d’idée au dernier moment mais que les mouvements s’annulent (Durand et al., 2001, 2002).

Une deuxième catégorie, est celle des répondants refusant de révéler leur intention de vote, ou disant ignorer pour qui ils voteront. Ces derniers se répartissent différemment de ceux qui déclarent leur intention. Plusieurs études (Durand et al., 2001, 2002; Curtice, 1997; Jowell et al., 1993) ont permis de confirmer que les gens qui refusent de dévoiler leurs intentions sont plus souvent à droite de l’échiquier. Donc, ce biais entraîne une sous-estimation des électeurs de droite dans les sondages.

Une troisième catégorie correspond aux gens plus difficiles à rejoindre. Ces répondants peuvent voter différemment des autres répondants. Tout comme les répondants changeant d’avis au dernier moment, et refusant de dévoiler leur intention, les gens plus difficiles à rejoindre sont généralement plus conservateurs (Curtice, 1997), mais pas selon Durand et al. (2002). Par conséquent, Lau (1994) et Vachon et coll. (1999) identifient une relation entre les efforts mis pour rejoindre les répondants (en augmentant ainsi le taux de réponse) et la qualité des estimations des intentions de vote. Comme Chevalier et Durand (1997), Triplett, Blair, Hamilton et Kang (1996) et Sebold (1988),

Vachon et coll. affirment qu' "il est en effet connu que les répondants moins coopératifs et ceux qui sont plus difficiles à rejoindre ont un profil socio-démographique, des attitudes et des comportements typés" (1999: 558).

La dernière explication de type psychosociologique est celle des fausses déclarations. Il semble que, dans la relation établie entre l'enquêteur et le répondant, ce dernier peut chercher à dissimuler son intention de vote réel afin de montrer un comportement "acceptable", valorisé socialement. La théorie de la "spirale du silence" de Noëlle-Neumann (1984) rend compte de ce problème de sincérité. Cette théorie se fonde sur l'idée que le public ressent certaines opinions comme étant plus désirables que d'autres et il tend à taire ses propres convictions pour se fondre dans l'opinion publique. De cette façon seront tués les opinions "illégitimes" et ce, par les fausses déclarations, les refus de répondre au sondage ou à la question d'intention de vote.

Boy et Chiche (1999a) posent le problème des fausses déclarations à l'aide de deux types d'explication de la "spirale du silence". Le premier type de "spirale du silence" est de type structurel. Dans la structure politique, les partis politiques sont parfois valorisés socialement tandis que d'autres souffrent d'une dévalorisation sociale. La sous-estimation des partis extrémistes est évidente: "il est difficile d'avouer un vote pour des partis qui se définissent comme extrémistes, c'est-à-dire comme se situant, de fait, aux marges du système républicain" (Boy et Chiche, 1999a: 242). En France, les valeurs républicaines furent d'abord remises en question par le parti communiste et ensuite par le Front national (Boy et Chiche, 1999a: 242-243). "Le Parti Communiste autrefois, le FN aujourd'hui n'ont pas bonne presse et l'on conçoit que ceux qui font ce choix tendent soit à refuser d'être interrogés soit à dissimuler leur intention à l'enquêteur" (Boy et Chiche, 1999b). La sous-estimation communiste entraînait une surestimation de la gauche non communiste puisque la réticence à déclarer une préférence pour le Parti communiste amenait "une proportion élevée de non-réponse, et parmi les réponses exprimées à une sous-estimation communiste entraînant d'ailleurs une surestimation de la gauche non communiste" (Stoetzel et Girard, 1982: 111).

Afin de rétablir ce sous-aveu, les instituts de sondage effectuent des redressements et “ces faits sont parfaitement connus et correctement traités la plupart du temps par les instituts de sondage” (Boy et Chiche, 1999a: 243). Comme l’indique Rivière:

Le cas du Front national ou du MNR illustre assez clairement les vertus de cette méthode. Il n’est pas rare, lorsqu’on demande aux sondés ce qu’ils ont voté lors des élections antérieures, que le score reconstitué de l’extrême droite soit trois à quatre fois moins élevé que ce qu’elle a réellement obtenu. Les sondeurs expliquent qu’il s’agit d’un vote “honteux”, les électeurs Frontistes ayant du mal à avouer leur choix, et/ou refusant tout simplement de répondre à l’enquête” (Rivière, 2001: 73).

D’après Boy et Chiche (1999a), un deuxième type de “spirale du silence” est de nature conjoncturelle. Il regroupe les dynamiques de domination et de légitimation à l’intérieur de la conjoncture politique au moment de la campagne. Certains candidats représenteront davantage les enjeux véhiculés lors de la campagne et donc ils seront valorisés et surestimés dans les sondages. Ce problème de sincérité peut également s’appliquer à la surestimation du parti dit gagnant, “le fait qu’un candidat ou un parti soit régulièrement présenté comme gagnant potentiel crée les conditions décrites par Noëlle-Neumann” (Boy et Chiche, 1999a: 243).

### 3.1.2 Explications méthodologiques

Les erreurs de prédiction peuvent être attribuées à des facteurs liés à l’échantillonnage: sous-représentation de certains secteurs ou de certains types de personnes dont les numéros de téléphone sont non-listés ou sur la liste rouge<sup>10</sup>, les personnes en déménagement fréquent ainsi que les personnes plus difficiles à rejoindre.

Tous les auteurs s’entendent pour dire que les répondants sur liste rouge ont des intentions de vote différentes de celles des autres répondants. Pour cette raison, la génération aléatoire de numéros de téléphone est nécessaire afin de rejoindre cette

---

<sup>10</sup> Numéro de téléphone volontairement confidentiel.

catégorie de gens et d'améliorer la qualité de l'échantillon.

Vachon et al. (1999) identifient trois indicateurs de la qualité de l'échantillon: l'utilisation des quotas, la substitution de numéros de téléphone et l'absence de récupération de refus. Selon eux, ces trois pratiques ont un effet sur le taux de réponse et influencent la représentativité de l'échantillon. Le premier indicateur étant l'utilisation des quotas "entraîne généralement une sous-représentation des personnes les moins coopératives en raison de la moins forte insistance sur le taux de réponse" (Vachon et al., 1999: 559). March et Scarborough (1990) considèrent également l'utilisation des quotas comme une pratique moins rigoureuse.

Selon plusieurs auteurs (Curtice, 1997; Curtice et Sparrow, 1997; Jowell et al., 1993; Lau, 1994 et Vachon et al., 1999), les sondages par quotas peuvent être moins fiables étant donné un taux de refus plus élevé, non contrôlé, une période de collecte trop courte et des quotas de classe sociale non contrôlables. Les études de Curtice (1997) et Curtice et Sparrow (1997) sur la comparaison des sondages par quotas effectués sur place et les sondages aléatoires téléphoniques, attribuent les moins bonnes estimations à la méthode des quotas.

De plus, la structure de l'échantillon par quotas ne permet pas d'obtenir constamment des échantillons représentatifs de la population. Par exemple, les échantillons incluent une plus grande proportion de gens diplômés qu'il y en a dans la population (Boy et Chiche, 1999a), mais c'est aussi le cas des échantillons aléatoires. Et puisque le Front national est beaucoup plus choisi par les personnes faiblement diplômées et que les échantillons sous-représentent cette catégorie, la sous-estimation du Front national tient aussi pour partie à la surévaluation culturelle (Boy et Chiche, 1999a).

Le second indicateur correspond à la substitution de numéros de téléphone. Cette pratique consiste à remplacer les personnes difficiles à rejoindre ou les gens qui refusent de répondre au sondage ou à la question d'intention de vote. Les répondants les plus difficiles à rejoindre sont sous-représentés et cette méthode nuit donc à la représentativité

de l'échantillon (Smith, 1983).

Le troisième indicateur est l'absence de récupération des refus. Ce procédé nuit à la qualité de l'échantillon puisqu'il ne permet pas d'augmenter la représentation des personnes non coopératives (Chevalier et Durand, 1997; Triplett et al., 1996).

Finalement, les sondages qui se déroulent sur une plus grande période de collecte produisent de meilleures estimations (Lau, 1994).

En bref, les facteurs psychosociologiques et méthodologiques entraînent des erreurs de prédiction. Le but des méthodes est de contrer les biais de sondages provoqués par les facteurs psychosociologiques comme nous venons de le voir. Avant de questionner les méthodes utilisées par les instituts français et afin de les mettre en relation avec la qualité des estimations, il convient de présenter de façon détaillée les événements qui sont survenus au premier tour de l'élection présidentielle française de 2002.

## **3.2 Premier tour de l'élection présidentielle française de 2002**

Les résultats du premier tour de l'élection présidentielle française de 2002 sont associés à un "séisme" marqué par la présence d'un candidat de l'extrême droite au second tour, au détriment de l'ancien Premier ministre et candidat de gauche Jospin. Un bref retour sur la campagne électorale et les résultats du vote semble essentiel pour saisir le contexte de cette élection.

### **3.2.1 Campagne électorale**

La campagne électorale présidentielle de 2002 se caractérise par la multiplicité des candidatures. De 1965 à 1995, le nombre de candidats fluctue entre six et douze lors des premiers tours des élections présidentielles. Les premières élections présidentielles au suffrage universel, celles de 1965, 1969 et 1974, ont donné lieu à des luttes entre respectivement six, sept et douze candidats. Pour devenir candidats, ces derniers

devaient obtenir 100 signatures, anonymes, d'élus provenant d'au moins dix départements. Suite à l'élection présidentielle de 1974 et au nombre grandissant de candidatures, une nouvelle mesure fut établie. À partir de 1976, les candidats doivent recueillir 500 signatures, rendues publiques, d'élus provenant d'au moins 30 départements et au maximum un dixième du même département. Suite à cette mesure, le nombre de candidats lors des élections présidentielles fut limité à dix candidats en 1981 et neuf candidats en 1988 et 1995. Alors qu'en 2002, le nombre de candidats fait un bond à seize candidats. Cette offre électorale diversifiée suppose de nouvelles stratégies pour l'électeur, le candidat, et le sondeur. Les électeurs doivent faire un choix parmi une multiplicité de candidatures jamais égalée. Pour leur part, les candidats tentent de se distinguer des autres candidats et d'attirer l'attention sur leurs thèmes de campagne. Les sondeurs peuvent éprouver devant cette dispersion des votes des difficultés à estimer la part relative des intentions de vote pour chacun des seize candidats.

Malgré un nombre élevé de candidats, un duel entre le président sortant et son Premier ministre est anticipé. Selon Gerstlé (dans Perrineau et Ysmal (dir.), 2003), "tous les commentaires de la presse depuis pratiquement les élections législatives de 1997, une grande partie des sondages, une grande partie de ce que disent et font les acteurs sont rapportés ou commentés en termes de promesse d'un rendez-vous ultime pour 2002" (Gerstlé dans Perrineau et Ysmal (dir.), 2003: 31).

Quant à la publication de sondages, l'interdiction de publication durant la dernière semaine a été modifiée le 19 février 2002 à une interdiction limitée uniquement à la veille et au jour même du scrutin. Les sondeurs peuvent donc estimer l'intention de vote des électeurs jusqu'à l'avant-veille du scrutin et en tracer un "portrait" plus juste. Selon le bilan de la Commission des sondages, 193 sondages ont été contrôlés lors de l'élection présidentielle de 2002 dont 115 sondages à l'échelon national, 19 sondages sur des populations spécifiques (les médecins, les agriculteurs, etc.), 51 sondages portant sur l'opinion (sans les questions sur les intentions de vote), 3 sondages réalisés à la sortie des urnes et 5 sondages sur le second tour de l'élection présidentielle.

### 3.2.2 Résultats du vote

Au premier tour de l'élection présidentielle de 2002, le niveau d'abstention de 28,4% correspond à un taux inégalé pour ce type d'élection en France. "La campagne a certainement joué un rôle dans le recul de la participation" (Perrineau et Ysmal dans Perrineau et Ysmal (dir.), 2003: 19). Les médias ont démontré peu d'attention à la campagne ou lorsqu'ils l'ont fait, l'ont décrite comme ennuyeuse. L'anticipation généralisée d'un duel Chirac-Jospin présentée par les médias ainsi que l'augmentation du scepticisme vis-à-vis les candidats n'ont certes pas favorisé la participation électorale (Gerstlé et Mercier dans Perrineau et Ysmal (dir.), 2003).

Lors du premier tour de l'élection présidentielle de 2002, l'extrême gauche (Laguiller de Lutte ouvrière, Besancenot de la Ligue communiste révolutionnaire et Gluckstein du Parti des travailleurs) recueille 10,5% des suffrages exprimés, l'extrême droite (Le Pen du Front national et Mégret du Mouvement national républicain) obtient 17,8% des suffrages exprimés et le mouvement catégoriel Chasse, pêche, nature et tradition avec son candidat Saint-Josse recueille 4,3% des suffrages exprimés. Au total, les partis extrêmes et marginaux (extrême gauche, extrême droite et mouvement catégoriel<sup>11</sup>) représentent 32,5% des suffrages exprimés.

Par ailleurs, les deux principaux candidats de la droite et de la gauche traditionnelles, obtiennent conjointement 36,1% du vote: Chirac recueille 19,9% et Jospin, 16,2%. Le vote du premier tour de l'élection présidentielle se présente donc comme un refus des grandes forces politiques de gauche et de droite.

Le reste des voix est partagé entre huit candidats qui recueillent 30,1%. En ordre d'importance, ces candidats sont: le candidat de l'Union pour la démocratie française François Bayrou (6,8%), le candidat du Pôle républicain Jean-Pierre Chevènement (5,3%), le candidat des Verts Noël Mamère (5,3%), le candidat de Démocratie libérale

---

<sup>11</sup> Nous vous référons à Bussi qui étudie le vote de Saint-Josse, le candidat de Chasse, pêche, nature et tradition (dans Perrineau et Ysmal, 2003 : 311-336).

Alain Madelin (3,9%), le candidat du Parti communiste français Robert Hue (3,4%), la candidate du Parti des radicaux de gauche Christiane Taubira (2,3%), la candidate du parti Citoyens, action, participation pour le 21e siècle Corinne Lepage (1,9%) et la candidate du Forum des républicains sociaux Christine Boutin (1,2%).

Outre l'éclatement du système partisan, la nature de l'élection présidentielle change. "La plupart des candidats n'étaient pas là pour être élus mais pour affirmer l'existence de leur formation politique (souvent par rapport aux concurrents de leur camp) ou pour témoigner" (Perrineau et Ysmal dans Perrineau et Ysmal (dir.), 2003: 18). Dolez et Laurent (dans Perrineau et Ysmal (dir.), 2003) qualifient les résultats de la gauche au premier tour de l'élection présidentielle de "gauche en miettes", mais la droite se partage également entre huit candidats.

Le premier tour de l'élection présidentielle française de 2002 est marqué par le séisme créé par la qualification du candidat du Front national Le Pen au second tour. Ni les sondeurs, ni les journalistes, ni les hommes politiques, ni le public, ne croyaient à la présence du candidat d'extrême droite au second tour de l'élection présidentielle. Tous les sondages précédant le premier tour prédisaient un second tour opposant les candidats Chirac et Jospin. Les douze sondages publiés durant la dernière semaine précédant le premier tour prévoyaient en moyenne 20 points au candidat de droite, Chirac, 18 points au candidat de gauche, Jospin, et 12,7 points au candidat de l'extrême droite, Le Pen (Durand et al., 2003). Cependant le jour du scrutin, le candidat Le Pen devança le candidat Jospin par moins d'un point de pourcentage (16,9% comparativement à 16,2%). Pour sa part, le président sortant, Chirac, récolte 19,9 points.

Mercier fait référence aux propos tenus par le conseiller "sondage" de Jospin:

lors d'une séance de l'équipe de campagne, [il] a évoqué, sans y croire lui-même réellement, la possibilité d'un dépassement par J.-M. Le Pen. "Nous écoutâmes Gérard Le Gall nous alerter sur l'apparition du risque possible statistiquement, mais improbable politiquement, au vu de l'évolution des deux courbes, de voir J.-M. Le Pen passer devant L. Jospin au premier tour. Un frisson saisit l'assistance qui n'a pu le croire" (citation dans Perrineau et Ysmal (dir.), 2003: 66).



Évoquer la présence de Le Pen au second tour de l'élection présidentielle était coûteux pour les sondeurs, les médias et les hommes politiques. "Les journalistes (mais aussi les sondeurs) ne pouvaient visiblement franchir une ligne symbolique qui les aurait fait titrer sur la possible présence de J.-M. Le Pen au second tour" (Mercier dans Perrineau et Ysmal (dir.), 2003: 64). Selon cet auteur, évoquer la présence de Le Pen au second tour présidentiel à la place de Jospin aurait entraîné des accusations de parti pris, de pression pour un "vote utile" par les autres candidats de gauche et d'accorder du crédit à Le Pen.

Malgré la publication tardive des sondages et leur nombre élevé, les estimations sont contestées. Afin d'évaluer la qualité des estimations de sondages, la méthodologie des instituts français doit être analysée.

### 3.3 Méthodologie des instituts de sondage français

De façon à tracer un portrait exhaustif de la situation et poser éventuellement les hypothèses, il convient de présenter la méthodologie utilisée par les instituts<sup>12</sup>. La décision est prise de se limiter à la période du 20 mars au 18 avril 2002 inclusivement<sup>13</sup>. En France, les principaux instituts qui réalisent les sondages électoraux sont: BVA, CSA, Ifop, Ipsos, Louis Harris et Sofres. Le dernier mois de réalisation de sondages compte trente sondages: cinq de BVA, cinq de CSA, six d'Ifop, cinq d'Ipsos, deux de Louis Harris et sept de Sofres. Pour tous les sondages, la taille des échantillons est d'environ 1000 répondants (généralement définies comme personnes de 18 ans et plus ou inscrits sur les listes électorales).

Une première caractéristique retrouvée chez tous les sondeurs est l'utilisation de la méthode des quotas. En France, la pratique courante veut que l'échantillonnage par quotas soit basée sur de cinq caractéristiques: sexe, âge, profession du chef de ménage, région, et catégorie d'agglomération. La méthode des quotas vise la représentativité de

---

<sup>12</sup> Voir le tableau VI qui porte les sondages réalisés du 20 mars au 18 avril 2002 à l'annexe A.

<sup>13</sup> Ce choix sera justifié au quatrième chapitre.

la population selon ces caractéristiques. Ce type d'échantillon consiste à déterminer le nombre de personnes dans la population qui possèdent chaque caractéristique et à rejoindre les répondants dont les caractères permettent de remplir les cellules des quotas. La tableau I présente les différentes catégories pour chaque caractéristique.

<b>TABLEAU I</b>	
<b>LES CATÉGORIES SOCIO-DÉMOGRAPHIQUES DE BASE</b>	
<b>DE L'ÉCHANTILLONNAGE PAR QUOTAS</b>	
<b>Sexe</b>	Homme
	Femme
<b>Âge</b>	18-24 ans
	25-34 ans
	35-49 ans
	50-64 ans
	65 ans et +
<b>PCS chef de ménage</b>	Agriculteur
	Artisan, commerçant, chef d'entreprise
	Cadre, profession intellectuelle supérieure
	Profession intermédiaire
	Employé
	Ouvrier
	Inactif, retraité
<b>Catégorie d'agglomération</b>	Moins 2000
	2000 à 20 000 hab.
	20000 à 100000 hab.
	100000 hab. et +
	Agglomération parisienne
<b>Région</b>	Nord
	Ouest
	Sud-Ouest
	Sud-Est
	Centre
	Est
	Région parisienne

Malgré l'échantillon non-probabiliste de la méthode des quotas, les instituts considèrent que, dans la pratique la marge d'erreur est la même que pour les sondages à échantillon aléatoire. Dans le cas française, le tirage de numéros de téléphone est aléatoire dans l'échantillonnage par quotas.

Les sondages d'intentions de vote du mois précédant le premier tour ont été réalisés soit par entrevue téléphonique (23 sondages) soit en face à face (7 sondages). Quatre instituts (CSA, Ifop, Ipsos, Louis Harris) ont utilisé uniquement l'entrevue téléphonique. BVA a privilégié l'entrevue face à face alors que la Sofres a utilisé les deux modes de passation; cet institut a réalisé cinq sondages téléphoniques et deux sondages face à face. Une autre caractéristique méthodologique des sondages d'intentions de vote est la durée de la collecte. En moyenne, les sondeurs français réalisent leurs sondages sur 2,2 jours. En fait, la collecte de données s'est effectuée sur une période d'une à quatre journées. La majorité (80%) des sondages se déroulent sur deux jours. Tous les sondages des instituts CSA, Ifop, Ipsos et Louis Harris se sont effectués sur une période de deux jours. BVA a réalisé un sondage ayant eu lieu pendant une seule journée, trois sondages de trois jours de collecte et un sondage se déroulant sur quatre jours. L'institut Sofres a réalisé six sondages sur deux jours et un sur trois jours de collecte. Quatre des cinq sondages ayant une plus longue durée de collecte (3 ou 4 jours) ont été réalisés en face à face.

D'autres informations méthodologiques ont pu être recueillies à la Commission des sondages. Comme la loi du 19 février 2002 donne accès aux fiches techniques des sondages, nous avons fait quelques visites à la Commission des sondages. Ces visites ont permis de consulter entre autres les notices détaillées des derniers sondages de chaque institut qui doivent être présentées à la Commission des sondages dans les vingt-quatre heures qui suivent la publication d'un sondage. Même si on ne peut affirmer que les données recueillies sont totalement représentatives de tous les autres sondages d'un même institut, celles-ci demeurent des sources d'information très pertinentes sur leur travail. Les informations complémentaires sur les derniers sondages de chaque institut proviennent des sondages présentés au tableau II.

<b>TABLEAU II LES DATES DE RÉALISATION DES DERNIERS SONDAGES DE CHAQUE INSTITUT</b>	
<i>Instituts de sondage</i>	<i>Dates de réalisation</i>
BVA <sup>14</sup>	10 au 13 avril
CSA	17 au 18 avril
Ifop	17 au 18 avril
Ipsos	17 au 18 avril
Louis Harris	16 au 17 avril
Sofres	17 au 18 avril

Puisque les sondages indiquent l'état du rapport de forces politiques à un certain moment, les derniers sondages - ceux qui sont les plus proches du jour du scrutin - devraient mieux représenter l'état de l'opinion publique dans les derniers moments de la campagne électorale.

La proportion des répondants qui refusent de révéler leur intention de vote ou qui disent ignorer pour qui ils voteront est en moyenne de 21,6% des intentions de vote au cours du dernier mois avant le premier tour de l'élection présidentielle française. Cependant, l'institut Ifop présente des proportions généralement plus basses de "non-exprimés", variant de 9% à 12%<sup>15</sup>. Deux cas de faible taux de non-exprimés sont observés pour l'institut Ipsos<sup>16</sup>. Les autres sondages indiquent des taux de "non-exprimés" oscillant

---

<sup>14</sup> Le dernier sondage de l'institut BVA a été réalisé le 18 avril 2002, mais la notice de ce sondage disponible à la Commission des sondages étant le même document qu'il était possible de voir sur le site internet de BVA, le choix s'est donc porté sur le sondage réalisé du 10 au 13 avril 2002 qui fournissait plus d'informations.

<sup>15</sup> Voir le tableau VI à l'annexe A.

<sup>16</sup> Les deux sondages de l'institut Ipsos ayant une faible proportion de "non-exprimés" sont ceux du 29 et 30 mars et du 5 et 6 avril 2003 avec 10% et 9% de "non-exprimés" (Voir l'annexe A).

entre 18% et 36%<sup>17</sup>. Suite à ces observations, il apparaît nécessaire d'exclure les "non-exprimés" de l'institut Ifop du calcul de la moyenne des "non-exprimés" puisque les résultats sont vraiment inférieurs aux autres instituts. Un taux de "non-exprimés" moyen de 24,4% est obtenu.

La formulation de la question de l'intention de vote est similaire dans les cinq instituts sur six dont l'information a été rendue disponible<sup>18</sup>. La question d'intention de vote de l'institut CSA demeure manquante.

Le type d'enquête varie selon les instituts. Le sondage de BVA est un omnibus à sujets multiples, les sujets à caractère politique ayant été placés en début de questionnaire. L'enquête de CSA a été insérée en première partie dans un omnibus ne comprenant que des sujets politiques. Ipsos a réalisé une enquête spécifique pour le sondage. L'institut Sofres a réalisé une enquête à sujets multiples sans spécifier l'endroit où se retrouve la question sur l'intention de vote. Les instituts Ifop et Louis Harris n'ont pas spécifié, dans les notices, le type d'enquête utilisé.

Les notices détaillées indiquent que les résultats de l'enquête ont été établis à travers la combinaison de redressement sur plusieurs critères: la certitude d'aller voter, la fermeté ou la sûreté du choix, l'intention de participation, et les redressements politiques suivants: le vote à l'élection présidentielle de 1995, le vote à l'élection législative de 1997, le vote européen de 1999, et les sympathies ou les préférences partisans. Les redressements diffèrent selon les instituts. Les fiches techniques sont très inégales, certaines sont très complètes tandis que d'autres présentent peu d'informations ou les mêmes informations préalablement trouvées sur le site Internet de l'institut. De plus, ce que les instituts inscrivent dans les fiches techniques et ce qu'ils fournissent comme informations sur les redressements dans ces mêmes notices ne coïncident pas.

Dans la fiche technique du sondage réalisé du 10 au 13 avril 2002 par l'institut BVA, le

---

<sup>17</sup> Voir l'annexe A.

<sup>18</sup> Voir le tableau VII portant sur la méthodologie du dernier sondage de chaque institut à l'annexe B.

mode de redressement a été effectué selon: les critères socio-démographiques, le vote au premier tour de l'élection législative de 1997 ainsi que le vote au premier et au second tour de l'élection présidentielle de 1995. Les résultats sont également analysés en fonction de différentes questions d'opinion, l'intention d'aller voter, la fermeté du choix, etc. Concrètement, la notice fournit seulement des informations sur la certitude du choix des électeurs exprimant une intention de vote.

En ce qui concerne le sondage des 17 et 18 avril de CSA, les informations sur les résultats de l'enquête ont été établies à travers la combinaison de redressement sur les critères suivants: la structure socio-démographique de l'échantillon, la certitude d'aller voter, la sûreté du choix, le vote à l'élection présidentielle de 1995, et le vote à l'élection législative de 1997. Toutefois, la fiche technique comprend uniquement les informations sur la certitude d'aller voter selon l'électorat et la sûreté du choix selon l'électorat.

L'institut Ifop inscrit dans la fiche technique du sondage réalisé les 12 et 13 avril 2002 que l'enquête a fait l'objet d'un redressement socio-démographique à partir des critères ayant servi de quotas (sexe, âge, profession du chef de ménage, région, catégorie d'agglomération) et d'un redressement politique sur la base de la reconstitution du vote européen de 1999, du vote législatif 1997 (premier tour), du vote présidentiel 1995 (premier et second tour), de l'intention de participation, et de la certitude d'aller voter. Cependant, cette notice inclut uniquement les informations sur la structure socio-démographique.

La notice du sondage réalisé les 17 et 18 avril 2002 par l'institut Ipsos permet de constater les critères de redressement utilisés. Le redressement a été effectué à partir de la structure socio-démographique de la population française âgée de 18 ans et plus (sexe, âge, profession du chef de ménage, région, catégorie d'agglomération) et du souvenir du vote au premier tour des élections législatives de 1997. Effectivement, la notice inclut la répartition socio-démographique de la population (résultats bruts et pondérés), le vote reconstitué au premier tour des législatives de 1997, et en plus, la liste des points d'enquête. Les informations disponibles sur le site Internet de cet institut permettent de

connaître le niveau d'indécision, la fermeté du choix de vote au premier tour, etc. Ipsos est l'institut qui fournit le plus d'informations dans les fiches techniques qui peuvent être consultées à la Commission des sondages.

La notice du sondage réalisé les 16 et 17 avril par l'institut Louis Harris n'étant pas disponible auprès de la Commission des sondages, cette dernière a fait la demande auprès de l'institut afin qu'il soit possible de consulter la notice complète de ce sondage. Dans un premier temps, sera examinée la notice du sondage réalisé les 29 et 30 mars 2002 afin de voir ce que l'institut dit inclure dans la notice consultable, puis la notice du 16 et 17 avril 2002 viendra préciser les informations disponibles. La fiche technique du sondage réalisé les 29 et 30 mars 2002 indique les intentions de vote présidentiel selon le sexe, l'âge, la catégorie socio-professionnelle du chef de famille, la profession de l'interviewé, la proximité partisane et les intentions de vote présidentielles du premier et du second tour de 1995. Puis, la notice du sondage réalisé les 16 et 17 avril 2002 permet de constater que toutes les informations indiquées dans la fiche du sondage des 29 et 30 mars se retrouvent dans la fiche détaillée. De plus, la notice inclut le vote aux législatives de 1997, la certitude d'aller voter et la sûreté du choix.

Dans la notice du sondage réalisé les 17 et 18 avril 2002 par la Sofres, il est possible de lire que différentes pondérations ont été effectuées selon les critères socio-démographiques, le vote au premier tour des élections législatives de 1997, le vote au premier tour des élections présidentielles de 1995, le vote au second tour des élections présidentielles de 1995, et la préférence partisane. Toutes ces informations se retrouvent dans la notice, mais la Sofres ne présente pas les résultats bruts. Il demeure impossible d'évaluer le travail que cet institut a réalisé.

Les instituts BVA et Sofres inscrivent dans les notices qu'ils ont recours à "l'expérience professionnelle des sondeurs". BVA indique: "Les spécialistes de BVA établissent les résultats en tenant compte de ces différentes pondérations et en s'appuyant sur des années d'expérience en matière de sondage électoral et d'intention de vote". Il y a lieu de se questionner sur cette pratique de redressement selon "l'expérience professionnelle" des

sondeurs. Cette pratique est davantage “subjective” que scientifique. Les instituts refusent de rendre publiques les données brutes, et le redressement selon “l’expérience professionnelle” ne favorise aucunement la transparence de la démarche.

### 3.4 Objectifs de recherche

La présente recherche vise à décrire l’évolution probable des intentions de vote pour les seize candidats à l’élection présidentielle française de 2002 à la lumière des sondages réalisés et à analyser les raisons possibles des erreurs d’estimation et de prédiction. Il s’agit tout d’abord de savoir si les estimations des sondages s’écartent réellement des résultats du vote. Les grandes erreurs de prédiction classiques observées plus tôt, telles la sous-estimation de la gauche, la surestimation de la droite, la sous-estimation des candidats ou des partis extrêmes, et plus particulièrement, de l’extrême droite, se présentent-elles dans le cas français de 2002? Quels sont les facteurs psychosociologiques et méthodologiques qui ont pu influencer la qualité des estimations?

Divers moyens semblent pertinents afin d’expliquer les erreurs de prédiction. Une étude post-électorale pourrait éclairer les raisons de ces erreurs. L’analyse des données de la grande enquête électorale “Panel électoral français 2002” serait très intéressante dans la mesure où elle comprend trois vagues d’avril à juin 2002<sup>19</sup>. Les données post-électorales sur les abstentionnistes, les gens qui ont changé d’avis au dernier moment (*late swing*), les gens qui n’ont pas dit la vérité, et la structure de l’échantillon (la sous-représentation de certaines classes sociales), procureraient des éléments explicatifs de certaines erreurs de prédiction. Suite à l’élection, les répondants qui disaient s’abstenir, au moment du sondage, se sont-ils réellement privés de leur droit de vote? Les non-répondants ont-ils finalement voté, l’ont-ils fait proportionnellement à l’ensemble des répondants ou sont-ils plus à droite de l’échiquier? Plusieurs auteurs pensent ainsi. Les répondants votent-ils comme ils l’avaient déclaré au moment du sondage, ou ont-ils changé d’avis au dernier moment? Des événements qui auraient influencé et fait changer

---

<sup>19</sup> La première vague a été réalisée du 8 au 20 avril, la seconde vague a été effectuée du 15 au 31 mai et la troisième vague s’est déroulée du 20 au 28 juin.



d'idée les répondants se sont-ils produits entre le moment du sondage et du vote? L'enquête post-électorale permet-elle de constater un écart important entre les intentions de vote avant, et après le sondage?

L'échantillon représentait-il bien les classes sociales de la population? Cet aspect est important dans la mesure où les classes sociales peuvent être liées au vote. Les deux premières vagues, préélectorale et post-électorale, du Panel électoral français de 2002, permettraient de comprendre le comportement de certaines catégories de répondants.

La deuxième approche projetée est l'étude de la méthodologie des instituts de sondage. À partir des fiches techniques des sondages, diffusées dans les médias écrits, ou sur les sites Internet, il est possible de comparer les méthodologies des instituts et de vérifier si la méthodologie de sondage influence la qualité des estimations. Le mode d'échantillonnage, la durée de la période de collecte, le mode de passation, le redressement, etc., sont des éléments explicatifs plausibles. Cette approche consiste à vérifier la présence de relations entre la méthodologie de sondage et la qualité des estimations. Selon les études précédentes, les sondages par quotas produisent de moins bons résultats que les sondages aléatoires. Un autre aspect méthodologique du sondage est la durée de la collecte. Certaines études soutiennent que plus la période de collecte est longue, plus les estimés seront bons. Puisque les sondages menés sur de plus longues périodes ont plus de possibilité d'inclure les répondants plus difficiles à rejoindre. Et comme les répondants plus difficiles à rejoindre votent différemment des autres répondants, il est important de connaître leur opinion. Afin de remédier à cette situation, la pratique de récupération de refus est essentielle. La question de la désirabilité est un aspect relatif au mode de passation du questionnaire qui peut influencer la qualité des estimations. Les questionnaires administrés en face à face impliquent une relation enquêteur/enquêté et un désir de bien paraître que les questionnaires téléphoniques ou auto-administrés ne présentent que dans une moindre mesure (Blais et Durand dans Gauthier (Ed.), 1997). Il sera intéressant de considérer le mode de passation comme un facteur influant la qualité des estimations.

Une troisième approche consiste en l'étude des redressements sur le vote antérieur. Les consultations auprès de la Commission des sondages permettent de recueillir les informations relatives aux redressements politiques exercés et ainsi comprendre un aspect supplémentaire de la méthodologie de sondage en France. Dans le cas présent, seulement quelques informations relatives aux redressements ont pu être recueillies.

Une quatrième approche consiste en la consultation des fiches détaillées et confidentielles des sondages à la Commission des sondages. Les instituts de sondage fournissent deux documents différents à la Commission des sondages. Un premier, auquel les gens qui en font la demande ont accès. Cependant, les données disponibles auprès de la Commission des sondages comprennent peu de renseignements supplémentaires comparativement aux données observées dans les fiches techniques diffusées dans les médias écrits ou sur les sites des instituts. Le second document est détaillé et confidentiel. Il inclut la méthodologie et notamment les redressements ainsi que les résultats bruts. Pourtant, la loi ne stipule à aucun moment l'existence de deux types de fiches différentes (une fiche consultable et une confidentielle). Afin de comprendre davantage la pratique des redressements, une demande a été effectuée auprès de la Commission des sondages afin d'avoir accès aux documents confidentiels. Cette requête a été refusée. Les informations fournies par cette requête auraient été d'une grande utilité dans le cadre de cette analyse. De plus, lors des visites il était impossible d'effectuer des photocopies des documents ce qui aurait permis de maximiser la consultation et de tracer un portrait plus détaillé de la méthodologie utilisée par les instituts de sondage.

Une cinquième et dernière approche consisterait en des consultations auprès des instituts de sondage. Des sources, sans doute inestimables, d'informations seraient disponibles et elles auraient permis de compléter l'analyse. Deux tentatives ont été faites auprès des instituts de sondage, mais faute de réponse, cette approche a dû être exclue du projet de recherche. De plus, comme la pratique de redressement semble être la "recette secrète" de chaque institut, la coopération des instituts demeurerait incertaine. Il est également envisageable que les sondeurs, fort occupés par le second tour de l'élection présidentielle

et les deux tours de l'élection législative, n'auraient pas été disponibles pour fournir les données.

Toutes ces perspectives de recherche sont très intéressantes. Toutefois, il demeure impossible d'étudier les sondages électoraux français de 2002 sous toutes ces approches dans le cadre d'un mémoire de maîtrise. Le volet descriptif du cas français consiste à analyser la justesse des estimations. Ce volet est indispensable à la compréhension du phénomène. Le volet explicatif, quant à lui, se restreindra à l'influence de la méthodologie de sondage sur la qualité des estimations. La formulation de la question de recherche se présente comme ceci:

*Quel portrait les sondages d'intentions de vote du premier tour de l'élection présidentielle française de 2002 tracent-ils l'évolution des intentions de vote? Les erreurs d'estimation ou de prédiction, le cas échéant, peuvent-elles être reliées à certaines pratiques méthodologiques?*

Les questions de recherche auxquelles nous devons répondre sont les suivantes:

- Au cours du dernier mois de la campagne électorale, quelle a été l'évolution des intentions de vote pour les seize candidats et les quatre regroupements de partis (extrême gauche, gauche modérée, droite modérée, extrême droite)?
- Les sondages s'écartent-ils de façon significative des estimations plausibles de l'évolution des intentions de vote?
- Certaines caractéristiques méthodologiques sont-elles liées à ces écarts?

À partir des questions mentionnées, l'analyse portera sur l'amont du "séisme" provoqué lors du premier tour de l'élection présidentielle de 2002. L'analyse de séries chronologiques permettra de comprendre ce qui s'est passé. Ensuite, l'étude de la méthodologie de sondage conduira à des pistes d'explications de certaines erreurs de

prédiction. Le prochain chapitre présente la méthodologie qui sera utilisée pour répondre aux questions de recherche.

## CHAPITRE 4. Méthodologie

Afin de répondre aux objectifs de recherche, la première section de ce chapitre présente les données utilisées dans le cadre du mémoire. La seconde section porte sur la justification du recours aux séries chronologiques pour l'analyse des données.

### 4.1 Données

Les données de ce mémoire proviennent des sondages d'intention de vote qui ont précédé le premier tour de l'élection présidentielle française de 2002. Les sondages mesurant les intentions de vote du premier tour ont été réalisés plusieurs mois avant le jour de l'élection. Dans le cadre de ce mémoire, il a paru pertinent d'étudier le dernier mois où il y a eu réalisation de sondages. Comme la loi du 19 février 2002 proscrit la publication de sondages la veille et le jour du scrutin, les 20 et 21 avril 2002, et qu'il n'y a pas eu de sondage réalisé le 19 avril, la période analysée s'étend du 20 mars au 18 avril 2002.

Plusieurs raisons justifient ce choix. Tout d'abord, les sondages qui précèdent cette période sont moins nombreux et il est préférable de détenir le plus de données quotidiennes possibles afin de réaliser les analyses de séries chronologiques. Comme il est nécessaire de calculer des moyennes d'intentions de vote quotidiennes afin d'effectuer les analyses de séries chronologiques, les intervalles sans sondage, demandent des imputations de données trop importantes.

De plus, les sondages réalisés plusieurs mois avant l'élection excluent certains candidats qui participent au premier tour, et ils incluent certains candidats qui ne figurent pas à la liste officielle. En effet, les intentions de vote pour certains candidats ne sont pas mesurées avant le début officiel de la campagne, le 4 avril, ce qui entraîne plusieurs données manquantes.

Les données de base ont été recueillies en partie dans les médias écrits tels les quotidiens - *Le Figaro*, *Le Monde*, *Aujourd'hui en France* et *Libération* - ainsi que les magazines ou les journaux hebdomadaires - *L'Express*, *Le Figaro Magazine*, *Le Journal du Dimanche*, *Le Nouvel Observateur*, *Le Point*, *Paris Match* et *Valeurs actuelles*. Il existe des ententes entre les instituts de sondage et les médias. Parfois quelques médias se réunissent et des baromètres naissent de ces ententes avec les instituts, par exemple *BFM*, *Libération* et *CSA*. *BVA* présente l'observatoire présidentiel et le tableau de bord politique dans le magazine *Paris Match* depuis 1981 (Bréchon et Salomon dans Bréchon (dir.), 2002). *CSA* produit un indicateur présidentiel des intentions de vote pour le quotidien *Libération*. *Ifop* réalise pour *L'Express* un baromètre mensuel depuis janvier 1999 et publie régulièrement des sondages dans le *Journal du Dimanche* (Bréchon et Salomon dans Bréchon (dir.), 2002). *Ipsos* produit un baromètre de l'action politique pour *Le Point*. *Louis Harris* réalise le baromètre *Louis Harris/Valeurs Actuelles* sur la mesure de la "cote d'action" du président de la République et du Premier Ministre. *Sofres* présente également un baromètre politique dans le *Figaro Magazine*. En plus de ces ententes, d'autres quotidiens et magazines commandent des sondages, ou publient les sondages commandités par d'autres médias.

Les données sont donc constituées des informations publiées par les médias écrits qui ont présenté l'ensemble des sondages du dernier mois avant le jour du scrutin. De plus, de nombreuses informations ont été dépouillées sur les sites Internet des instituts afin de compléter celles publiées par les médias écrits. Enfin, les fiches techniques des derniers sondages ont été consultées auprès de la Commission des sondages.

Ces sources d'information permettent non seulement de recueillir les renseignements sur les intentions de vote, mais également de documenter la méthodologie de sondage utilisée par les instituts. Les différentes caractéristiques méthodologiques qui seront étudiées sont: le redressement, la durée de la collecte, le moment de la collecte (sondage réalisé pendant au moins un jour du week-end ou la semaine seulement), le mode de passation, la taille de l'échantillon, le type d'échantillon, la proportion d'abstention, la formulation de la question et le type d'enquête.

## 4.2 Méthode d'analyse

Le premier objectif de recherche vise à tracer le portrait de l'évolution des intentions de vote des candidats et des regroupements de partis. Pour ce faire, il faut d'abord estimer l'évolution des intentions de vote, puis évaluer la qualité des estimations et enfin tenter d'analyser les explications possibles des écarts entre les estimations et le vote lorsqu'ils se produisent.

### 4.2.1 Procédure d'analyse choisie

Dans le cas présent, l'analyse de séries chronologiques servira à la fois à estimer l'évolution des intentions de vote et à prédire le vote des divers candidats et des regroupements de partis. Comme le souligne Ostrom (1990), un but de l'analyse peut être simplement la *description* de la série. Ce premier objectif consiste à mettre en évidence les diverses composantes des séquences d'observation dans le temps (tendance, composante saisonnière, etc.).

Pour ce faire, les estimations de sondages constituent les données de base. Initialement, la base de données inclut les estimations des trente sondages réalisés au cours du dernier mois précédant le premier tour de l'élection présidentielle. À partir de ces données, il s'agit de transformer la suite des estimations de sondages en une série quotidienne d'intentions de vote moyennes. La première étape est la transformation de la base de données par la répartition des intentions de vote de chaque sondage sur chacun des jours où il y a eu réalisation<sup>20</sup>. La deuxième étape est le recours à l'agrégation des résultats et le calcul de la moyenne des intentions de vote pour chaque journée où plusieurs sondages sont réalisés. Si aucun sondage n'a été effectué une journée, les séries chronologiques permettent de remplacer les cas manquants à partir des valeurs connues de la tendance afin d'imputer les valeurs manquantes. Il s'agit d'un processus d'intrapolation linéaire.

---

<sup>20</sup> Si un sondage a lieu les 17 et 18 avril, les résultats seront attribués aux deux journées.

L'analyse de série chronologique permet l'estimation la plus plausible de l'évolution à partir de la série des moyennes quotidiennes des données.

De façon à bien évaluer l'évolution des intentions de vote, la création d'une première variable indépendante de temps est nécessaire. Les valeurs de la variable de temps correspondent à la période d'observation, le mois précédant l'élection, et plus précisément, du 20 mars au 21 avril 2002. Cette période correspond aux valeurs de 0 à 32. L'analyse de la contribution de cette variable permettra de savoir si les intentions de vote moyennes varient dans le temps de façon linéaire.

Une deuxième variable indépendante est le temps au carré. Dans un premier temps, il s'agit de centrer la variable de temps en lui attribuant les valeurs de -16 à +16 et ensuite de la mettre au carré. L'analyse de cette variable permet de vérifier l'existence d'une évolution quadratique ( $\cup$  ou  $\cap$ ) des intentions de vote.

Les variables indépendantes peuvent également représenter des événements qui ont influencé le déroulement de la série chronologique. Dans le cas présent, aucun événement marquant n'a été noté qui aurait pu entraîner des variations dans l'évolution des intentions de vote des candidats pendant le dernier mois de la campagne.

Le second objectif de l'analyse de série chronologique est la *prévision*. La série chronologique permet de prédire des valeurs futures à court terme. En général, il est plutôt difficile de réaliser des prévisions à long terme. Cela dépend du comportement de la série dans le passé. Dans le cas français 2002, seulement quelques observations séparent la dernière observation et la fin de la série. Les derniers sondages réalisés ont eu lieu le 18 avril 2002, il y aura donc des extrapolations sur trois jours, les 19, 20 et 21 avril 2002. À partir de la prédiction finale de l'évolution des intentions de vote moyennes, le jour du vote, il sera possible de calculer l'écart entre la fin de la série chronologique et le résultat du vote.



L'ensemble de ces analyses seront effectuées pour chacun des candidats ainsi que pour quatre regroupements de partis. L'espace politique français est divisé en catégories<sup>21</sup> (Grunberg et Schweisguth dans Perrineau et Ysmal, 2003). L'extrême gauche rassemble les candidats Arlette Laguiller (Lutte ouvrière - LO), Olivier Besancenot (Ligue communiste révolutionnaire - LCR) et Daniel Glückstein (Parti des travailleurs) - PT). La gauche modérée regroupe Robert Hue (Parti communiste français - PCF), Lionel Jospin (Parti socialiste - PS), Jean-Pierre Chevènement (Pôle Républicain - Pôle Rep), Noël Mamère (Verts) et Christiane Taubira (Parti radical de gauche - PRG). La droite modérée se compose de Jacques Chirac (Rassemblement pour la République - RPR), François Bayrou (Union pour la démocratie française - UDF), Corinne Lepage (Citoyens, action, participation pour le 21e siècle - CAP 21), Christine Boutin (Forum des républicains sociaux - FRS), Jean Saint-Josse (Chasse, Pêche, Nature et Traditions - CPNT) et Alain Madelin (Démocratie libérale - DL). L'extrême droite rassemble Jean-Marie Le Pen (Front national - FN) et Bruno Mégret (Mouvement national républicain - MNR). À partir des séries des regroupements de partis, une analyse du travail individuel des sondeurs sera effectuée. Il sera possible de vérifier si certains sondeurs s'écartent des estimations des intentions de vote plus régulièrement que d'autres.

#### 4.2.2 Justification de la procédure d'analyse

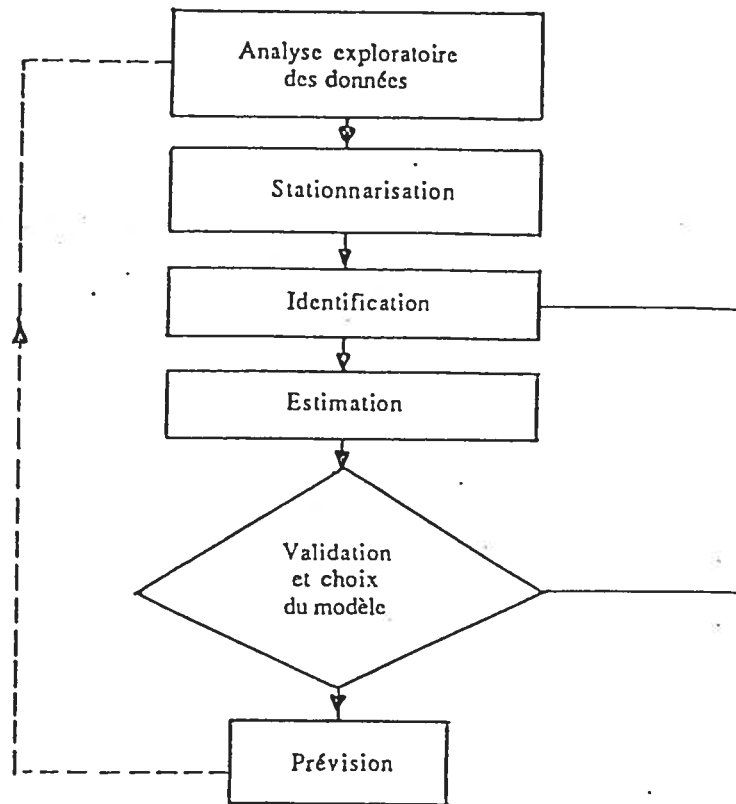
Pour mesurer et modéliser l'effet des variables indépendantes sur l'évolution des intentions de vote moyennes des divers candidats et des regroupements de candidats ou partis, l'analyse de régression pourrait sembler appropriée. Toutefois comme les observations quotidiennes d'intentions de vote moyennes ne sont pas indépendantes entre elles, c'est l'analyse de série chronologique ARIMA (AutoRegressive Integrated Moving Average) qui est adéquate. En effet, une première composante de la série chronologique est l'auto-corrélation de ses termes. Les observations dépendent statistiquement des observations précédentes. Donc, la régression linéaire ou multiple, qui suppose le

---

<sup>21</sup> Le classement des divers candidats dans les catégories ne fait pas l'unanimité chez les auteurs. Nous justifierons nos regroupements au moment de l'analyse de l'évolution des intentions de vote des regroupements de partis ou candidats.

traitement d'observations indépendantes, ne peut être utilisée. L'analyse de série chronologique ARIMA permet, quant à elle, le contrôle de l'auto-corrélation des termes des séries chronologiques.

En 1976, Box et Jenkins propose la modélisation ARIMA, une analyse de séries chronologiques pouvant être résumée par le schéma suivant:



### Graphique 1

#### La modélisation ARIMA

Source : Droesbeke, Fichet et Tassi, 1989: 198.

*L'analyse exploratoire des données* consiste à représenter graphiquement les observations. Cette étape permet de visualiser les effets répétitifs, comme les composantes tendancielle, cyclique, saisonnière et irrégulière. La tendance représente en fait l'évolution moyenne à long terme de la variable étudiée. La composante saisonnière ou cyclique est une composante périodique dont la période est généralement dépendante de l'environnement, par exemple, la saison ou le mode d'organisation de la

société (vacances, fêtes de fin d'année, etc.). Elle présente théoriquement un comportement périodique, dont la période est suffisamment importante pour être qualifiée de mouvement à moyen ou à long terme. La composante irrégulière regroupe tout ce qui n'est pas pris en compte par les autres composantes (saisonniers, cyclique et tendancielle); ce sont les résidus.

La *stationnarisation*, tout comme l'analyse exploratoire, est un préalable au traitement des observations d'une série chronologique. Le processus de stationnarisation, particulièrement utilisé dans le cas d'une série saisonnière, se réalise par des transformations généralement linéaires. Si la composante saisonnière est présente, il convient de l'isoler à l'aide d'un premier paramètre qui est l'ordre différencié ( $d$ ). Il consiste à faire de la série un processus immobile, une série chronologique stable. Comme l'évolution des intentions de vote moyennes ne présente pas de composante saisonnière, le modèle est immobile et le paramètre  $d$  prend la valeur de 0. En effet, les intentions de vote ne sont pas liées à un cycle comme il est possible d'en apercevoir dans les dépenses d'une entreprise, qui sont toujours plus élevées au quatrième trimestre. Dans le cas français de 2002, l'analyse se déroule sur une durée d'un mois. La période est trop courte pour observer une composante saisonnière. L'évolution des intentions de vote ne se prête pas à cet aspect de l'analyse.

L'*identification* est l'étape permettant de déterminer le degré d'auto-régression ( $p$ ) et le degré de la moyenne mobile ( $q$ ) du modèle. Le terme d'auto-régression ( $p$ ) représente la mémoire du processus de prédiction des observations. Le terme d'ordre auto-régressif ( $p$ ) correspond au nombre de termes dans le modèle qui décrit la dépendance par les observations successives. Par exemple, si la valeur de  $p$  est 0, il n'y a pas de relation entre les observations adjacentes. Comme les intentions de vote dans le cas français de 2002 dépendent de l'observation précédente, le modèle comprendra un terme d'auto-régression ( $p$ ) de valeur 1.

La dernière composante à identifier est le degré de la moyenne mobile ( $q$ ). Elle représente la mémoire du processus pour le choc aléatoire précédent. Les moyennes

mobiles ( $q$ ) permettent de remplacer les observations par la moyenne des valeurs qui l'entourent. Cette technique, utilisée pour étudier le problème de tendance dans la mesure, permet d'éliminer les variations accidentelles ou périodiques. Les moyennes mobiles sont particulièrement utilisées dans les méthodes de désaisonnalisation, c'est-à-dire dans la séparation des fluctuations saisonnières des autres composantes. Les moyennes mobiles s'expriment sous forme d'un filtre linéaire. Dans le cas français de 2002, le  $q$  prend la valeur 0 puisqu'il n'y a pas de composante saisonnière, accidentelle ou périodique à corriger.

La prochaine étape dans la modélisation de séries est *l'estimation*. La modélisation ARIMA permet d'estimer les coefficients identifiés du modèle. En remplaçant les paramètres  $(p, d, q)$  par les valeurs préalablement identifiées  $(1,0,0)$ , ARIMA utilise les itérations nécessaires afin de déterminer par maximum de vraisemblance l'évolution plausible de la série. De nouvelles variables sont ajoutées au fichier: les valeurs prédites, les résidus de la prédiction (erreurs) et les intervalles de confiance.

La *validation* consiste principalement en l'analyse des résidus de l'estimation. Les résultats résiduels sont les différences entre les résultats prédits par le modèle et les résultats réels pour les séries. Ce diagnostic est l'étape consistant à valider la phase d'identification, les paramètres  $(p, d, q)$  choisis. La validation porte sur l'analyse de la qualité statistique du modèle estimé ainsi que sur l'analyse des résidus d'estimation. Les tests d'autocorrélation (Autocorrelation Function - ACF) et d'autocorrélation partielle (Partial autocorrelation function - PACF) des erreurs permettent de vérifier si le processus d'autocorrélation est bien contrôlé par l'analyse de série chronologique ARIMA.

Lorsque le modèle est correct, l'autocorrélation sera contrôlée. Dans plusieurs applications des séries chronologiques, identifier et modéliser les " patterns " des données est suffisant pour produire une équation avec laquelle il est possible de prédire le futur de la série. C'est seulement si le modèle est correct que la *prévision* est possible, et cette phase est souvent la finalité du processus.

## CHAPITRE 5. Résultats

Ce cinquième chapitre se divise en trois sections. *Primo*, les résultats des séries chronologiques pour chacun des candidats et des regroupements seront présentés: les tendances à la hausse ou à la baisse de l'évolution des intentions de vote, et l'écart entre la prédiction de la série et le résultat de l'élection. *Secundo*, un retour sera fait sur les prédictions des sondages afin de résumer les erreurs qui se sont produites au premier tour de l'élection présidentielle française de 2002. *Tertio*, à partir des séries des regroupements de candidats ou de partis, une analyse du travail individuel des sondeurs sera effectuée. Il sera possible de vérifier si certains sondeurs s'écartent des estimations des intentions de vote plus régulièrement que d'autres.

Les analyses de séries chronologiques ARIMA permettront donc 1) de tracer un portrait de l'évolution des intentions de vote des candidats et des regroupements de partis, 2) de calculer le nombre de points gagnés ou perdus entre le début et la fin de la série chronologique, et 3) de calculer l'écart entre la prédiction finale de la série chronologique et le résultat de l'élection.

### 5.1 Évolution des intentions de vote

Les graphiques suivants illustrent l'estimation de l'évolution des intentions de vote moyennes<sup>22</sup> du dernier mois de sondages précédant le jour du scrutin pour les candidats et les regroupements. Les échelles de graphiques varient selon l'évolution des intentions de vote moyennes de chaque candidat et de chaque regroupement. La lecture du graphique devra donc être faite en tenant compte de cette information.

---

<sup>22</sup> Sur le graphique, l'estimation de l'évolution des intentions de vote est représentée par une ligne pleine. Les lignes pointillées représentent l'intervalle de confiance des séries.

De plus, le graphique illustre les marges d'erreur. "La marge d'erreur, c'est la précision du résultat obtenu étant donné le seuil de confiance que l'on désire. C'est une distance à l'intérieur de laquelle on s'attend à ce que soit situé le paramètre estimé (dans la population)" (Durand, 1999:12). Pour ce faire, il s'agit de déterminer le seuil de confiance recherché, c'est-à-dire "la certitude que l'on a quant à la justesse des résultats" (Durand, 1999:12). Le critère généralement admis est 95%, ce qui signifie dans le cas présent, la distance à l'intérieur de laquelle les résultats se retrouvent dans 95% des fois. Ainsi, il est possible de voir si le résultat du vote des candidats<sup>23</sup> se retrouve à l'intérieur de l'intervalle de confiance de l'estimation moyenne de l'évolution des intentions de vote. À partir de ce constat, il sera possible de confirmer la présence ou l'absence d'erreur de prédiction des sondages.

### 5.1.1 Évolution des candidats

À gauche, huit candidats se partagent 43,0% des suffrages exprimés au premier tour de l'élection présidentielle française. Les instituts ont-ils réussi à bien estimer les intentions de vote de cette "gauche en miettes" comme la qualifie Dolez et Laurent (dans Perrineau et Ysmal (dir.), 2003)? Ont-ils bien saisi les tendances à la hausse ou à la baisse des candidats? Quels sont les écarts qui séparent les prévisions des séries et les résultats de l'élection? Finalement, les sondages se sont-ils trompés?

L'ordre de présentation des candidats est établi selon le résultat du vote le plus élevé, donc Jospin sera le premier candidat de la gauche à être présenté.

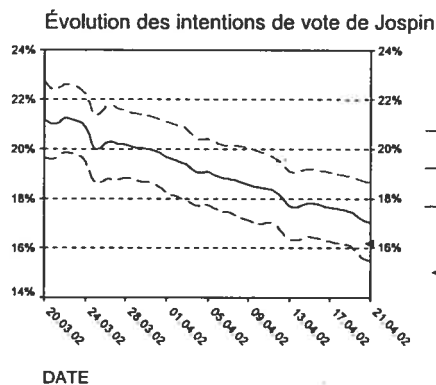
Le graphique 2 présente l'évolution des intentions de vote du candidat du Parti socialiste *Lionel Jospin*. Il a connu une baisse linéaire significative de 4,1<sup>24</sup> points de pourcentage

<sup>23</sup> Le résultat du vote est représenté par une flèche.

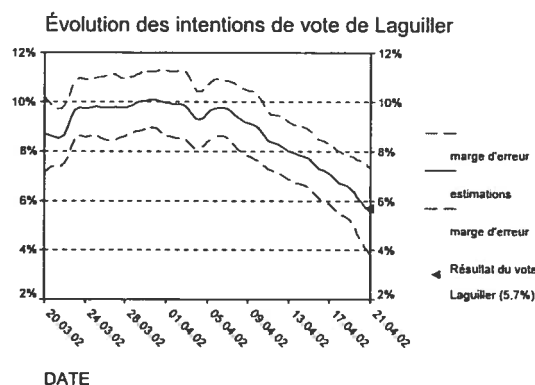
<sup>24</sup> L'écart est calculé à partir des résultats de l'analyse de série chronologique ARIMA, plus précisément la différence obtenue entre le début et la fin de la série. Il est également possible de mesurer l'écart à partir du coefficient de régression (B) de la variable de "temps" (Voir les trois tableaux de l'annexe C). Il s'agit de multiplier ce coefficient de régression par 32, puisqu'il y a 32 intervalles entre le 20 mars et le 21 avril 2002, et de multiplier par 100 pour obtenir le résultat en pourcentage. Par exemple, la série chronologique ARIMA de Jospin nous informe d'un B= -0,00130 et un p= 0,00000. Le coefficient est significatif si  $p < 0,05$ . Le signe du coefficient de régression indique une relation positive ou négative de la tendance. Selon

du 20 mars au 21 avril 2002. Selon la série chronologique, il passe de 21,2 à 17,1 points. La prédiction de l'analyse montre une surestimation de 0,9 point, car Jospin a obtenu un résultat de 16,2 points au premier tour présidentiel. D'après le modèle, le résultat du vote se situe à l'intérieur des marges d'erreur des estimations prédites à partir des sondages réalisés.

Graphique 2



Graphique 3

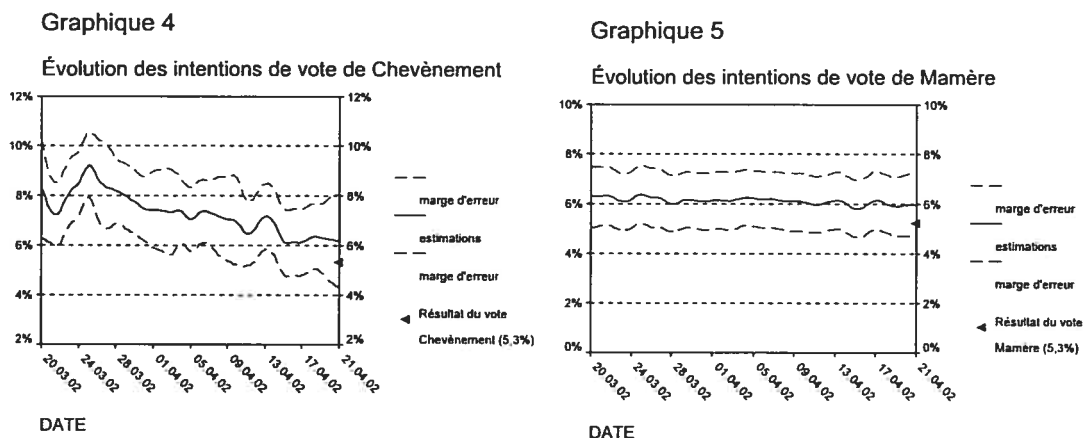


Le graphique 3 présente l'évolution des intentions de vote de la candidate *Arlette Laguiller* du parti Lutte ouvrière. Le graphique montre une évolution selon des fonctions quadratique ( $\cap$ ) et linéaire significatives. Le 20 mars, elle obtient 8,7 points et elle connaît une évolution positive jusqu'à environ la moitié de la période, moment où elle recueille environ 10,1 points, sa plus haute intention de vote. Par la suite, elle perd des points jusqu'au jour du scrutin où elle atteint 5,6 points. Du début à la fin de la période, elle connaît une baisse linéaire significative de 3,1 points. Ce modèle de prédiction prévoit très bien son résultat à l'élection de 5,7% des suffrages exprimés, soit une sous-estimation de 0,1 point. Le résultat du vote se trouve à l'intérieur de l'intervalle de confiance.

Le graphique 4 illustre la baisse linéaire significative de 2,1 points entre le 20 mars et le 21 avril 2002 pour le candidat du Pôle républicain *Jean-Pierre Chevènement*. Au

les résultats de l'analyse, Jospin a donc connu une baisse linéaire significative de 4,2 points ( $(32 \times -0,0013) \times 100$ ) du 20 mars au 21 avril 2002. La baisse est linéaire puisque seule la variable indépendante de "temps" est significative ( $p=0,0000$ ). La variable de "temps au carré" n'est pas significative. Les résultats des deux méthodes présentées diffèrent parfois de quelques décimales puisqu'il y a arrondissement dans les deux cas. Les résultats de cette analyse correspondront à la première méthode, c'est-à-dire la différence entre le début et la fin de la série chronologique ARIMA (Voir les trois tableaux de l'annexe D).

début de la période étudiée, il obtenait environ 8,3 points et il termine avec 6,2 points. La prédiction de la série surestime de 0,9 point le résultat du vote de Chevènement qui est de 5,3% des suffrages exprimés. L'estimation de l'évolution des intentions de vote prédit bien le vote de ce candidat puisque le résultat de l'élection se trouve à l'intérieur des marges d'erreur des estimations de sondages.

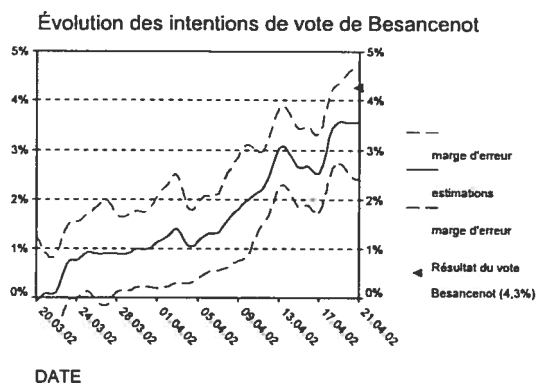


Le graphique 5 présente l'évolution des intentions de vote du candidat des Verts *Noël Mamère* et montre une perte de 0,3 point entre le début et la fin de la série chronologique. Cette baisse est non-significative. Pour cette raison, l'estimation de vote du candidat Mamère est considérée stable. En fait, il est passé de 6,3 à 6,0 points entre le 20 mars et le 21 avril 2002. L'écart qui sépare la prédiction de la série et le résultat de l'élection est de 0,7 point au-dessus du vote prédit et le résultat du vote est à l'intérieur de l'intervalle de confiance. Les sondages ont bien prédit le résultat du vote de Mamère.

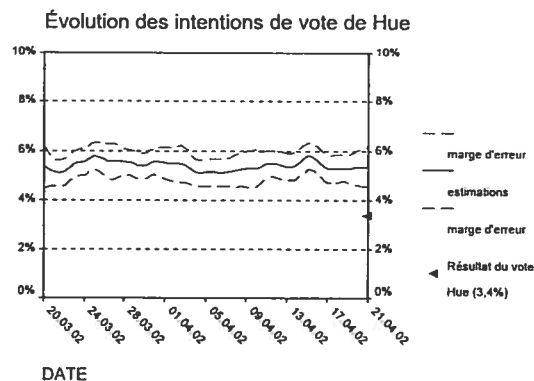
Le graphique 6 montre l'évolution des intentions de vote du candidat de la Ligue communiste révolutionnaire *Olivier Besancenot*. Ce candidat a obtenu la plus forte hausse linéaire comparativement à tous les autres candidats au cours de la campagne. En effet, la série chronologique indique une augmentation significative de 3,6 points entre le 20 mars et le 21 avril. Selon la série, Besancenot est passé d'une intention de vote nulle le 20 mars à 3,6 points le jour du scrutin. De plus, la prédiction finale des estimations de vote a sous-estimé de 0,7 point son résultat à l'élection, qui est de 4,3 points. Ce candidat semble bien estimé par les sondages, puisque le résultat du vote se trouve à l'intérieur des marges d'erreur.



Graphique 6



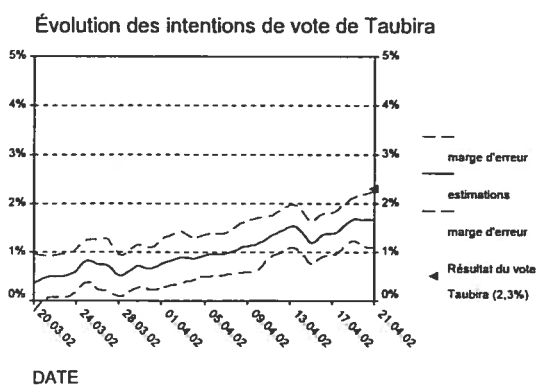
Graphique 7



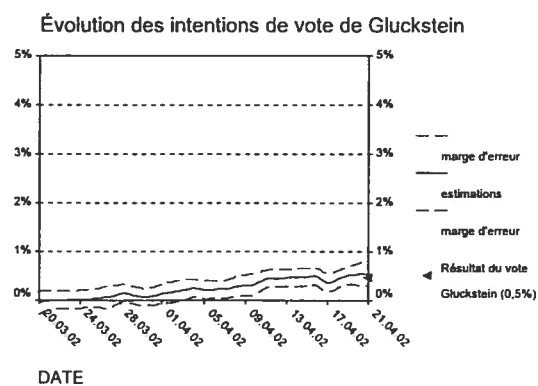
Le graphique 7 présente une stabilité de l'intention de vote pour le communiste *Robert Hue* du 20 mars au 21 avril 2002. L'écart calculé sur la période étudiée est de 0,1 point et cette très faible baisse linéaire est non-significative. Toutefois, le candidat du parti communiste est le plus surestimé de tous les candidats de gauche (1,9 point). Il a obtenu 3,4% des suffrages exprimés comparativement aux 5,3 points prédits par la série chronologique. La prédiction de la série est à l'extérieur de l'intervalle de confiance.

Le graphique 8 illustre la hausse significative de l'intention de vote pour la candidate du Parti des radicaux de gauche *Christiane Taubira* (1,3 point entre le 20 mars et le 21 avril). Elle est passée de 0,4 à 1,7 point selon une fonction linéaire. La prévision de la série a sous-estimé la candidate de 0,6 point comparativement au vote puisqu'elle a obtenu 2,3% des suffrages exprimés. Le résultat du vote se situe à l'extérieur de l'intervalle de confiance.

Graphique 8



Graphique 9

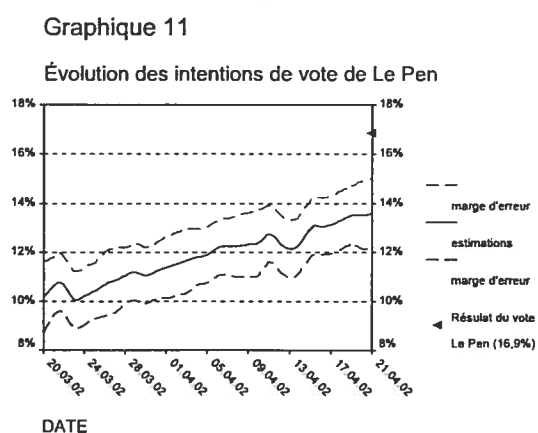
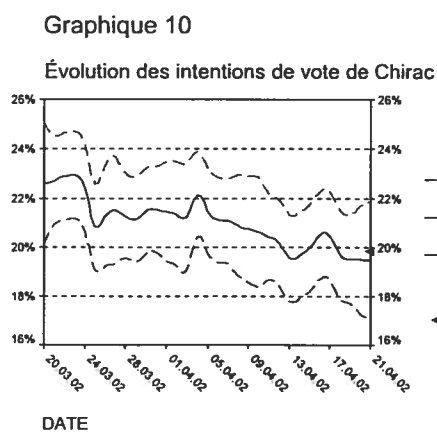


Le graphique 9 montre une hausse significative de l'intention de vote pour le candidat *Daniel Gluckstein* du Parti des travailleurs. Selon la série chronologique, il a débuté le 20 mars avec une intention de vote nulle, et il est passé à 0,6 point, soit une augmentation de 0,6 point. La prédiction de la série indique une légère surestimation de 0,1 point par rapport au résultat du vote de 0,5 point.

Du côté droit de l'échiquier politique, huit candidats se partagent 57,1% des suffrages exprimés. Quels sont les portraits que tracent l'évolution des candidats? Les estimations d'intentions de vote s'écartent-elles des résultats du vote?

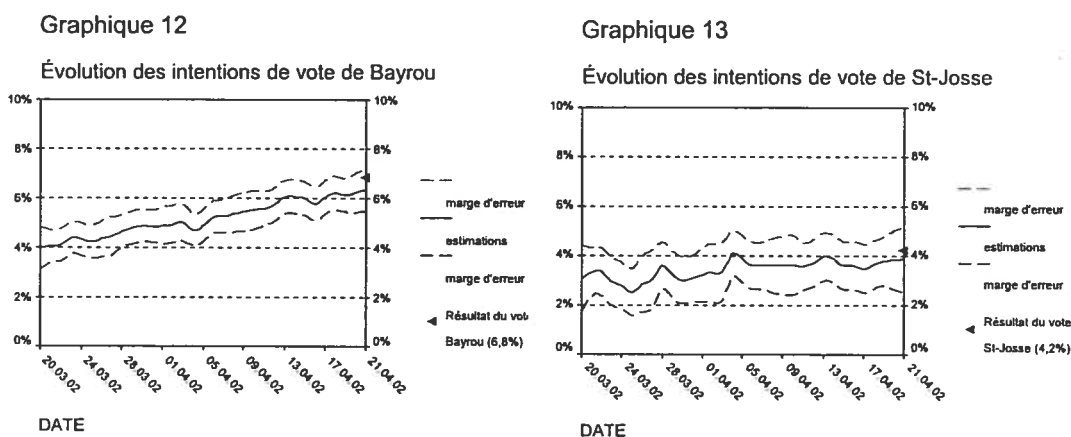
Les candidats de droite seront également présentés en ordre d'importance. Le premier candidat, étant celui qui récolte le plus de suffrages exprimés, sera Chirac.

Le graphique 10 illustre l'évolution des intentions de vote du candidat du Rassemblement pour la République *Jacques Chirac* et montre une baisse significative de 3,1 points du 20 mars au 21 avril 2002. Selon la série chronologique, il a débuté cette période avec 22,6% des intentions de vote et il termine avec 19,5%. Les estimations de sondage du président sortant ont sous-estimé le vote réel de 0,4 point, puisqu'il obtient 19,9% des suffrages exprimés lors du premier tour. Le résultat de l'élection est à l'intérieur de l'intervalle de confiance. Il est donc possible d'affirmer que les sondages ont bien prédit le score du président sortant Chirac.



Le graphique 11 montre une augmentation linéaire significative durant le dernier mois de l'évolution des intentions de vote du candidat du Front national *Jean-Marie Le Pen*. Le candidat de l'extrême droite a gagné 3,4 points, selon la série chronologique. L'analyse montre le candidat Le Pen à 10,2 points au début du mois précédant le scrutin, et la prédiction finale de la série le place à 13,6 points. Il est certainement le candidat le plus sous-estimé de tous puisqu'il obtient 16,9% des intentions de vote, soit un écart de 3,3 points avec la prédiction. Les sondages ont saisi une tendance à la hausse du candidat, mais la hausse s'avère plus importante que celle prédite par les sondages. Comme le résultat du vote se situe à l'extérieur des marges d'erreurs, il est possible de conclure que les sondages ont mal prédit le vote du candidat Le Pen.

Le graphique 12 présente l'évolution de l'intention de vote du candidat de l'Union pour la démocratie française *François Bayrou*. Ce graphique montre une augmentation linéaire significative de 2,3 points au cours de la période étudiée. Il passe de 4,0 à 6,3% d'intentions de vote. Son résultat lors de l'élection qui est 6,8 points, soit une sous-estimation de 0,5 point. La prédiction de la série est bonne puisque le résultat de l'élection se situe à l'intérieur de l'intervalle de confiance des estimations.

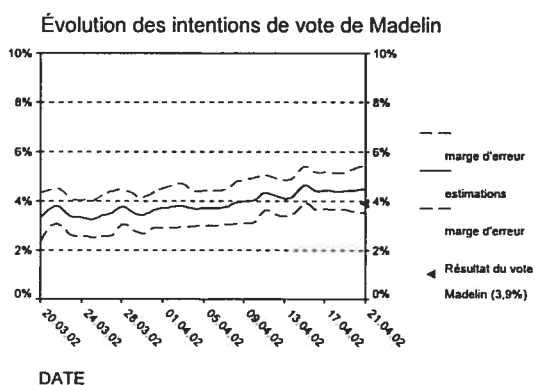


Le graphique 13 illustre une tendance plutôt stable de l'évolution des intentions de vote du candidat du parti Chasse, pêche, nature et traditions *Jean Saint-Josse* du 20 mars au 21 avril 2002. Selon la série chronologique, il bénéficie d'une hausse de 0,9 point, mais

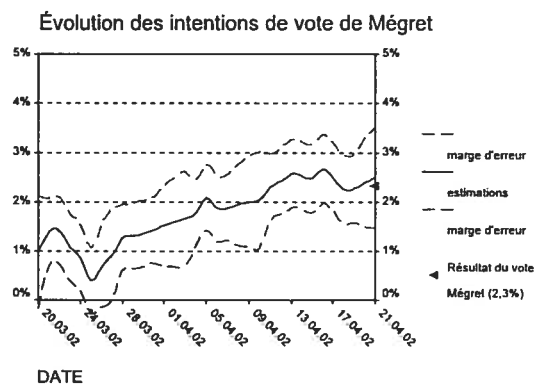
cette augmentation linéaire est non-significative. Ce candidat est donné à 3,0 points le 20 mars et à 3,9 points le 21 avril. Les sondages ont bien prédit le vote de Saint-Josse puisqu'il obtient 4,2 points, le jour du vote. La sous-estimation est donc de 0,3 point entre la prédiction de la série et le résultat du vote, et le résultat se situe à l'intérieur de l'intervalle de confiance.

Le graphique 14 présente une hausse de l'évolution des intentions de vote de 1,2 point du candidat de Démocratie libérale *Alain Madelin*. Cette tendance est significative et linéaire. Le candidat passe de 3,3 à 4,5 points dans la série chronologique. Le jour de l'élection, Madelin obtient 3,9% des suffrages exprimés. La prédiction de la série présente donc une surestimation de 0,6 point. Comme le résultat du vote se situe à l'intérieur des marges d'erreur, il est possible de conclure que les sondages ont bien prédit le résultat de Madelin.

Graphique 14



Graphique 15

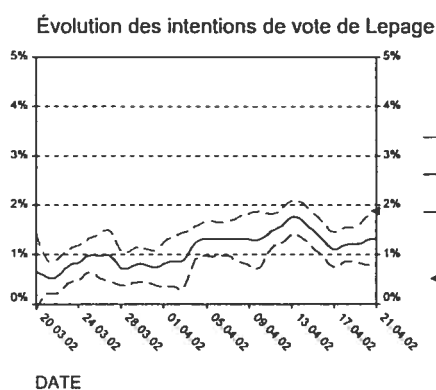


Le graphique 15 montre l'évolution des intentions de vote du candidat du Mouvement national républicain *Bruno Mégret*. Ce graphique présente une hausse linéaire significative de 1,5 point. Du 20 mars au 21 avril, il passe de 1,0 à 2,5 points. L'évolution des intentions de vote prédit bien le résultat du vote en surestimant de 0,2 point puisque le jour du scrutin Mégret obtient 2,3% des suffrages exprimés.

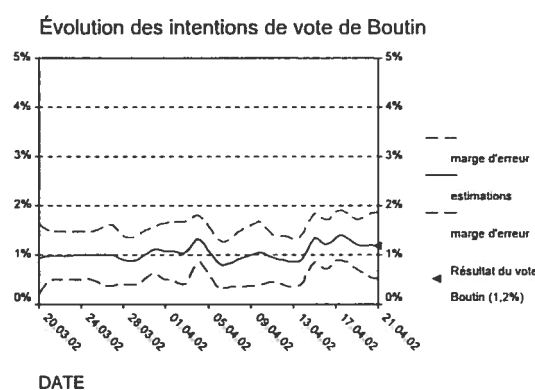
Le graphique 16 montre une stabilité de l'évolution des intentions de vote pendant le mois d'observation de la candidate du parti Citoyens, action, participation pour le 21e

siècle *Corinne Lepage*. Elle passe d'environ 0,7 à 1,3% d'intentions de vote dans la série chronologique. Selon la prédiction de la série, la candidate gagne 0,6 point, mais cette hausse est non-significative. Avec un résultat de vote de 1,9% des suffrages exprimés, la candidate est sous-estimée de 0,6 point. Cet écart est relativement important étant donné la faible intention de vote de cette candidate. Le résultat du vote se situe exactement sur la ligne extérieure (du haut) de la marge d'erreur.

Graphique 16



Graphique 17



Le graphique 17 de la candidate du Forum des républicains sociaux *Christine Boutin* montre également une stabilité de l'évolution des intentions de vote du 20 mars au 21 avril 2002. Elle connaît une hausse linéaire non-significative de 0,3 point, passant approximativement de 0,9 à 1,2 point. La prédiction de la série est exacte, étant donné le résultat de 1,2% des suffrages exprimés.

### 5.1.2 Évolution des regroupements de partis

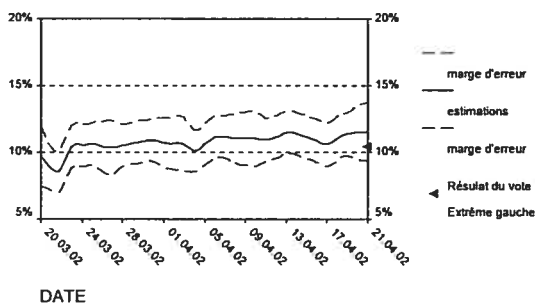
Afin de vérifier si les sondages du premier tour de l'élection présidentielle de 2002 présentent des erreurs de prédiction classiques telles la surestimation de la gauche, la sous-estimation de la droite et la sous-estimation des extrêmes, il convient d'effectuer quatre regroupements de candidats ou de partis: extrême gauche, gauche modérée, droite

modérée et extrême droite<sup>25</sup>. Les candidats de gauche sont divisés en deux groupes: la gauche modérée représentée par Jospin, Chevènement<sup>26</sup>, Mamère, Hue, et Taubira et l'extrême gauche incluant Laguiller, Besancenot et Gluckstein.

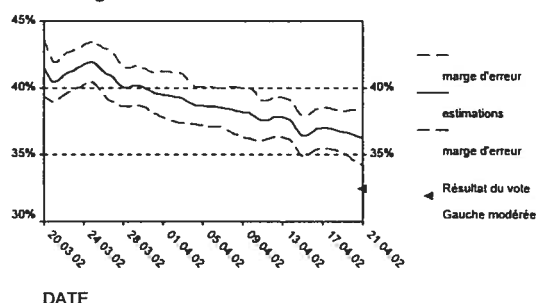
Les regroupements de partis seront présentés selon l'échiquier politique de la gauche vers la droite: l'extrême gauche, la gauche modérée, la droite modérée et l'extrême droite.

Tout d'abord, l'évolution des intentions de vote de l'extrême gauche indique une hausse non significative de 1,9 point. Le graphique 18 montre que du 20 mars au 21 avril 2002, les intentions de vote de l'extrême gauche passent de 9,7 à 11,6 points. L'extrême gauche est considérée comme étant stable pendant la période étudiée. La prédiction finale est supérieure de 1,1 point par rapport au résultat du vote qui atteint 10,5% des suffrages exprimés. Les sondages ont un peu surestimé les résultats du vote des candidats de l'extrême gauche, mais le résultat de l'élection se situe à l'intérieur des marges d'erreur.

Graphique 18  
Évolution des intentions de vote  
de l'extrême gauche



Graphique 19  
Évolution des intentions de vote  
de la gauche modérée

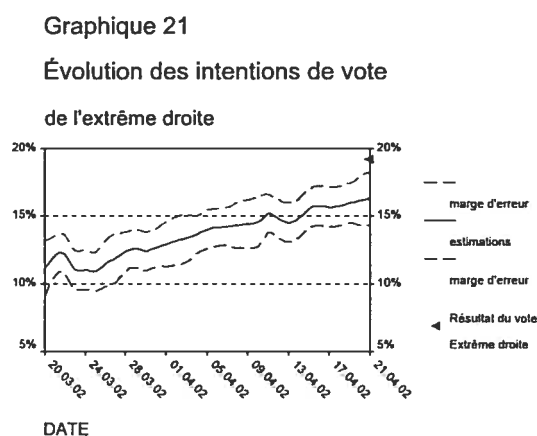
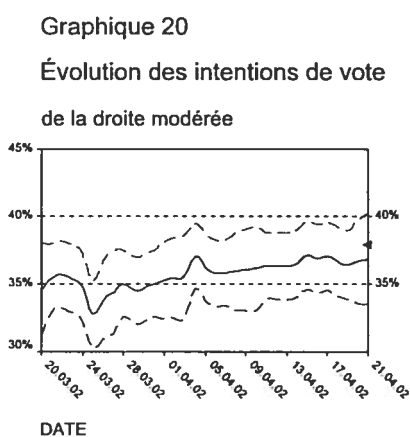


<sup>25</sup> Ces regroupements de partis ou de candidats ont été choisis en fonction de ce qui présenté dans les médias. Malgré que certains refusent d'accorder une idéologie spécifique à l'extrême gauche (Grunberg et Schweisguth dans Perrineau et Ysmal (dir.), 2003), il convient de diviser les candidats ou les partis en quatre regroupements, dans le cas présent, afin de vérifier la présence des grandes erreurs de prédiction classiques dans le cas du premier tour français de 2002.

<sup>26</sup> Grunberg et Schweisguth (dans Perrineau et Ysmal (dir.), 2003 : 349) excluent Chevènement de la gauche modérée et ils le posent comme étant intermédiaire entre la gauche et la droite modérée. En tenant compte de ce propos, une analyse de série chronologique de la gauche modérée sans le candidat Chevènement a été effectuée et les résultats demeurent semblables. Le maintien du candidat Chevènement dans le regroupement de la gauche modérée est donc justifié.

Le graphique 19 de l'évolution des intentions de vote de la gauche modérée présente une baisse linéaire significative de 5,2 points. Au début du mois, la gauche modérée obtient 41,5 points et elle termine le jour du vote avec 36,3 points. Les résultats de l'élection indiquent 32,5% des suffrages exprimés pour les candidats de la gauche modérée. Ce qui signifie que la prédiction de la série surestime de 3,8 points la gauche modérée. Les sondages avaient saisi une tendance à la baisse de la gauche modérée, mais leur constat est insuffisant. Le résultat de vote se situe à l'extérieur de l'intervalle de confiance, ce qui permet d'affirmer que les sondages ont surestimé le vote des candidats de la gauche modérée.

D'autres séries chronologiques ARIMA ont été effectuées en regroupant les candidats de la droite modérée<sup>27</sup> (Chirac, Bayrou, Saint-Josse, Madelin, Lepage et Boutin) et de l'extrême droite (Le Pen et Mégrét). Le graphique 20 illustre l'évolution des intentions de vote de la droite modérée: une hausse non-significative de 2,2 points où l'intention de vote est passée de 34,6 à 36,8 points du 20 mars au 21 avril 2002. Puisque le résultat du vote atteint 37,9% des suffrages exprimés, les sondages ont un peu sous-estimé les résultats du vote des candidats de la droite modérée, mais la prévision est valide puisque le résultat de l'élection se trouve à l'intérieur des marges d'erreur.



<sup>27</sup> Grunberg et Schweisguth (dans Perrineau et Ysmal (dir.), 2003 : 349) excluent les candidats Lepage et Saint-Josse de la droite modérée. Selon eux, Lepage se trouve entre la gauche et la droite modérée et Saint-Josse entre la droite modérée et l'extrême droite. Une série chronologique de la droite modérée a été réalisée en excluant ces candidats. L'analyse mène aux mêmes conclusions que celle incluant ces candidats. Les candidats ont donc été maintenus dans l'analyse.

Finalement, le graphique 21 présente l'évolution des intentions de vote de l'extrême droite et montre une hausse significative de 5,2 points. Les candidats Le Pen et Mégrét cumulaient 11,1 points le 20 mars et ils ont atteint 16,3 points le 21 avril. Comparativement au résultat de l'élection qui est de 19,2% des suffrages exprimés, le modèle a sous-estimé de 2,9 points le résultat du vote. La série chronologique montre que le résultat du vote s'écarte de l'intervalle de confiance et donc que les sondages ont sous-estimé le résultat du vote des candidats de l'extrême droite.



## 5.2 Prédictions des séries chronologiques

Le tableau III présente de façon synthétique les écarts entre le résultat de l'élection et la prédiction de la série chronologique pour les candidats et les regroupements<sup>28</sup>.

<b>TABLEAU III</b>			
<b>LES ÉCARTS ENTRE LES RÉSULTATS DU VOTE ET LES PRÉDICTIONS DES SÉRIES POUR LES CANDIDATS ET POUR LES REGROUPEMENTS DE PARTIS</b>			
	<i>Élections</i>	<i>Prédictions séries</i>	<i>Écarts</i>
Extrême gauche	10,5	11,6	+1,1
Laguiller	5,7	5,6	-0,1
Besancenot	4,3	3,6	-0,7
Gluckstein	0,5	0,6	+0,1
Gauche modérée	32,5	36,3	+3,8
Jospin	16,2	17,1	+0,9
Chevènement	5,3	6,2	+0,9
Mamère	5,3	6	+0,7
Hue	3,4	5,3	+1,9
Taubira	2,3	1,7	-0,6
Droite modérée	37,9	36,8	-1,1
Chirac	19,9	19,5	-0,4
Bayrou	6,8	6,3	-0,5
Saint-Josse	4,2	3,9	-0,3
Madelin	3,9	4,5	+0,6
Lepage	1,9	1,3	-0,6
Boutin	1,2	1,2	0
Extrême droite	19,2	16,3	-2,9
Le Pen	16,9	13,6	-3,3
Mégret	2,3	2,5	0,2

<sup>28</sup> Les résultats des regroupements proviennent des écarts entre la prédiction de la série et le résultat du vote, et non de la somme des écarts calculés pour les partis que les regroupements incluent.

Les prédictions des séries chronologiques permettent d'affirmer que les sondages ont bien prédit les résultats des candidats l'extrême gauche et de la droite modérée puisque les résultats du vote se situent à l'intérieur de l'intervalle de confiance. Les sondages ont surestimé de 1,1 point l'extrême gauche et sous-estimé la droite modérée de 1,1 point. Par contre, les sondeurs ont surestimé la gauche modérée de 3,8 points et sous-estimé l'extrême-droite de 2,9 points.

À gauche, l'évolution des intentions de vote des candidats est fiable, excepté pour Hue et Taubira. Le candidat communiste Hue semble le plus surestimé des candidats de gauche avec 1,9 point. La candidate Taubira est sous-estimée de 0,6 point.

À droite, la part la plus importante de la sous-estimation est attribuée au 3,3 points de Le Pen. La candidate Lepage est estimée de justesse par les instituts de sondage qui la sous-estime de 0,6 point. Les candidats Chirac, Bayrou, Saint-Josse, Madelin, Mégret et Boutin sont bien estimés par les sondages.

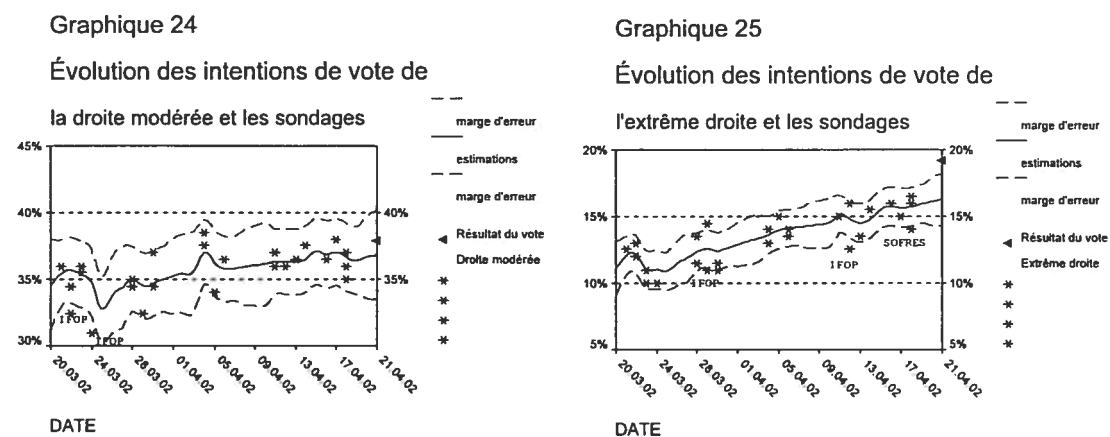
D'autres mesures ont été utilisées afin de vérifier les écarts entre les estimations de sondages et les résultats du vote. Durand et al. (2003) présentent une comparaison entre les douze sondages publiés lors de la dernière semaine de la campagne et les résultats de l'élection. L'analyse des sondages de la dernière semaine indique que l'ensemble des candidats de l'extrême gauche sont surestimés de 0,5 point, les candidats de la gauche modérée sont surestimés de 5,1 points, les candidats de la droite modérée sont sous-estimés de 1,4 point et les candidats de l'extrême droite sont sous-estimés de 4,1 points.

Une moyenne des derniers sondages de chaque institut a également été calculée par Durand et al. (2003). Les estimations moyennes produites par les derniers sondages de chaque institut accordent une surestimation de 0,7 point à l'extrême gauche, une surestimation de 4,8 points de la gauche modérée, une sous-estimation de 1,5 point pour la droite modérée ainsi qu'une sous-estimation de 4,0 points pour l'extrême droite.



Dans le cas de la gauche modérée, les sondages surestiment les candidats. Donc si les sondages présentent une estimation au-dessus de la marge d'erreur supérieure, il sera possible d'affirmer que ces sondages prédisent mal le résultat du vote, ou, s'éloignent de l'estimation moyenne des sondages. Le graphique 23 présente deux sondages qui s'écartent de la marge d'erreur supérieure de la série chronologique de la gauche modérée, et ces deux sondages ont été réalisés par l'institut Ifop. À l'extérieur de la marge d'erreur inférieure, Ipsos a réalisé un sondage, mais à une date hâtive. Ce qui signifie que cet institut se rapproche davantage du résultat de l'élection, et que ce sondage est plus précis.

Le graphique 24 présente l'évolution des intentions de vote de la droite modérée. Les sondages ont bien estimé le résultat du vote puisqu'il se trouve à l'intérieur des marges d'erreur. Les séries indiquent deux sondages qui s'écartent des marges d'erreur, et ce sont deux sondages de l'institut Ifop.



Le graphique 25 montre que la série chronologique de l'extrême droite a sous-estimé le résultat du vote. Donc, les sondages à l'extérieur de la marge d'erreur au niveau inférieur font erreur, et ceux qui s'écartent de la marge d'erreur supérieure se rapprochent davantage de la réalité du vote. Dans le cas présent, BVA a bien fait. Cet institut a présenté des résultats supérieurs à l'estimation moyenne qui sous-estimaient l'extrême droite. Trois sondages se sont écartés de la marge d'erreur inférieure, c'est-à-dire qu'ils

ont sous-estimé encore plus que l'estimation moyenne le faisait. Un des ces sondages est de la Sofres et deux sondages ont été réalisés par l'Ifop. Au total, six sondages sur huit s'éloignent de l'estimation moyenne dans la mauvaise direction, sont de l'institut Ifop. Ce qui signifie que ces sondages sont moins bons que les autres.

Le prochain chapitre aborde les explications possibles des mauvaises prédictions notées dans les séries de même que des écarts plus importants de certains sondeurs.

## CHAPITRE 6. Discussion des résultats

L'analyse de séries chronologiques présentant l'estimation de l'évolution des intentions de vote des candidats et des regroupements de partis a permis d'identifier les principales erreurs de prédiction du premier tour de l'élection présidentielle française de 2002: la surestimation de la gauche modérée, la sous-estimation de l'extrême droite et plus particulièrement la sous-estimation de Le Pen, le candidat du Front national. De plus, tous les instituts se sont trompés sur le duel du second tour. Pourtant, d'importantes tendances à la baisse du candidat Jospin et à la hausse du candidat Le Pen étaient enregistrées par tous les sondeurs. Aucun n'a osé prédire ce qui est de fait arrivé au premier tour, la victoire de Le Pen sur Jospin. Enfin, les analyses de l'évolution des intentions de vote des regroupements de partis montrent que les estimations de l'institut Ifop s'écartent des marges d'erreur plus souvent que celles des autres instituts.

Les caractéristiques méthodologiques sont-elles liées aux erreurs d'estimation et de prédiction de sondages? Ce dernier chapitre porte sur trois points principaux: les hypothèses au sujet des biais de sondage qui demeurent impossibles à vérifier faute d'informations adéquates, les principales découvertes de ce mémoire et les pratiques "sondagières" pouvant être remises en cause, associées aux redressements.

### 6.1 Hypothèses possibles

Une grande homogénéité caractérise la méthodologie de sondage des instituts français au premier tour de l'élection présidentielle française de 2002. Une caractéristique méthodologique utilisée par tous les sondeurs est la méthode des quotas<sup>30</sup>.

---

<sup>30</sup> Selon la fiche technique du sondage réalisé les 17 et 18 avril par la Sofres, les catégories socio-démographiques de l'échantillon respectent la structure de la population. Voir le tableau à l'annexe E.

L'utilisation des quotas soulève beaucoup de débats depuis le début des sondages. Au niveau théorique, la méthode des quotas utilisant l'échantillon non-probabiliste est non justifiée "puisque'on ne donne pas à chaque unité de l'échantillon la même chance d'être retenue" (Javeau, 1982: 50). Cet échantillon ne permet pas théoriquement l'inférence statistique, ce qui entrave la représentativité de la population visée par l'enquête. "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). Le statisticien Kish (1965) affirme que "l'échantillonnage par quota n'est pas une méthode scientifique définie. Au contraire (chaque sondage par quota) ressemble à une production artistique, difficile à définir ou à décrire" (citation dans Blondiaux, 1991: 761). Avec les quotas, les ménages peuvent être sélectionnés au hasard, mais le choix du répondant ne sera pas fait de manière aléatoire à l'intérieur du ménage puisque l'enquêteur doit remplir ses quotas. Cependant, au plan pratique, l'utilisation de la méthode des quotas suscite des avis partagés. Pour sa part, Bon soutient:

"Ces méthodes [choix au hasard] donnent des résultats satisfaisants dans des pays, comme les États-Unis, où l'habitat est suffisamment groupé et homogène. Elles restent inapplicables en France: les villes et les régions sont si différentes, les communes ont des tailles si variables qu'il faudrait trop de soins, de contrôles, et donc de temps et d'argent, pour s'assurer que le tirage au sort porte bien sur la totalité de la population étudiée. Voilà pourquoi les instituts français ont recours à une seconde méthode où l'échantillon - tout en conservant globalement les propriétés de représentativité définies plus haut - se compose au départ de toute autre façon" (Bon, 1974: 29).

Cette définition des sondages aléatoires semble erronée. L'échantillonnage aléatoire permet au contraire d'obtenir une meilleure représentativité de la population.

Dans un échantillon aléatoire, les sondeurs ont avantage à rejoindre le plus de répondants sélectionnés par le mode d'échantillonnage et à développer des méthodes pour améliorer leur taux de réponse. De cette façon, les personnes difficiles à rejoindre ayant des

“comportements typés” auront plus de chances de faire partie de l'échantillon (Vachon et al., 1999: 558). Le fait de connaître leur opinion permettra d'améliorer la qualité de l'échantillon. Toutefois, tous les sondages de l'élection présidentielle française de 2002 utilisent l'échantillonnage par quotas. L'objectif des quotas est de créer une micro-population où certaines caractéristiques représentent bien celles de la population. L'échantillon est complet lorsque le nombre de personne possédant chaque caractéristique de base dans l'échantillon est atteint. Comme le taux de non-réponse n'est pas un critère tenu en compte dans cette logique, aucun effort n'est déployé pour interroger les personnes difficiles à rejoindre afin d'améliorer la qualité de l'échantillon.

La méthode des quotas suppose que “le phénomène à mesurer (en l'occurrence les opinions ou les attitudes) ait une corrélation très forte avec les caractères de contrôle retenus” (dans Blondiaux, 1991: 762). Dans le cas des sondages électoraux français, les catégories sont l'âge, le sexe, la profession du chef de ménage, la taille d'agglomération et la région. L'enquêteur devra rejoindre des gens selon des caractéristiques spécifiques très fines, par exemple: une femme âgée de 50 à 64 ans, ouvrière habitant dans une agglomération de 100 000 habitants et plus dans le sud-ouest de la France. Considérant que la taille de l'échantillon est d'environ 1000 répondants, très peu de répondants se retrouvent dans chaque catégorie socio-démographique. L'échantillon peut être biaisé lorsqu'un petit nombre de répondants dans une catégorie représente l'opinion de l'ensemble des gens appartenant à cette catégorie. Enfin, la relation est moins forte entre ces caractéristiques et le vote comparativement à autrefois où les idéologies politiques étaient fortement liées à ces caractéristiques socio-démographiques.

De plus, la méthode des quotas semble entraîner systématiquement l'inclusion d'une plus grande proportion de gens diplômés<sup>31</sup> qu'il n'y en a dans la population (Boy et Chiche, 1999a). Comme le candidat du Front national recueille des intentions de vote élevées auprès des chômeurs, des ouvriers, des personnes de faible niveau d'études (Perrineau dans Perrineau et Ysmal, 2003: 206), il est possible de penser qu'une partie de la sous-estimation du candidat du Front national Le Pen par la sous-estimation des gens

---

<sup>31</sup> Le niveau de scolarité ne fait pas partie des caractéristiques de base des quotas.



faiblement diplômés (Boy et Chiche, 1999a). Toutefois, la sous-estimation des gens diplômés se retrouve également dans les échantillons aléatoires. Il s'agit donc plus d'une question relative à la capacité de joindre les personnes moins scolarisées que de la méthode des quotas.

Les erreurs de prédiction peuvent également être attribuées à la sous-représentation de certains secteurs ou de certains types de personnes, particulièrement dans les sondages téléphoniques. Une première catégorie est celle des personnes non-listés vivant en institution, en résidence étudiante, etc., qui n'ont pas accès à un numéro de téléphone privé. Selon Durand et al., (2001, 2002), les gens vivant en institutions votent davantage à droite de l'échiquier politique. Or, ces personnes, ne pouvant être rejointes à un numéro de téléphone privé ne sont pas sondées, ce qui peut entraîner un biais défavorable à la droite.

Une deuxième catégorie est celle des numéros de téléphone sur la liste rouge (numéros confidentiels). Les gens sur la liste rouge ont des intentions de vote différentes des autres répondants (Vachon et al., 1999), donc leur inclusion dans l'échantillon est importante. Dans le cas français, les échantillons sont constitués de numéros de téléphone choisis aléatoirement, mais les gens non-listés y sont exclus. Pour cette raison, les échantillons entraînent probablement un biais.

En outre, il semble que les efforts mis pour rejoindre les gens permet d'obtenir une meilleure qualité des estimations (Vachon et al., 1999). Un indicateur possible peut être la durée de la période de la collecte (Lau, 1994). Traugott (1987) notait "pour plusieurs sondages commerciaux des périodes de collecte tellement courtes qu'il était impossible au cours de l'enquête de tenter à nouveau de rejoindre les répondants n'étant pas disponibles au début de l'enquête" (citation de Traugott, 1987: 51 (traduction) dans Vachon, 1998: 16). Lors du premier tour de l'élection présidentielle française de 2002, la période de collecte des sondages a été de 2,2 jours en moyenne. Cette période de collecte est plutôt courte comparativement à celle de certains pays, notamment au Québec avec 6,5 jours et au Canada avec 7,1 jours lors de l'élection fédérale de 2000 (Durand et

al., 2003). En France, puisque la période de collecte des sondages est semblable pour chaque sondeur, il est impossible de comparer leur travail respectif. Toutefois, cette caractéristique méthodologique peut entraîner un biais de sondage pour l'ensemble des instituts français.

En résumé, il est impossible de vérifier l'influence de l'utilisation des quotas sur la qualité des estimations puisque cette méthode est utilisée par tous les instituts. Cependant, il est possible que cette pratique entraîne des biais pour l'ensemble des sondages comme la surestimation de la gauche modérée, la sous-estimation de l'extrême droite et plus particulièrement la sous-estimation du candidat du Front National Le Pen.

## 6.2 Principales découvertes

Dans le cas des sondages électoraux, le comportement des répondants est vérifiable et il semble qu'une "proportion significative ne dit pas la vérité" (Hyman, 1944: 557-559 dans Blondiaux, 1991: 777). Blondiaux note:

"La validité de leurs réponses est toujours fragile, surtout lorsqu'il s'agit de mesurer des opinions et des comportements socialement normés. Les répondants auront tendance à choisir la réponse qui convient, celle qui leur apparaît la plus légitime. Les chercheurs trouveront un nom pour décrire cet effet: l'"effet de prestige". L'expérience prouve que les individus se comportent différemment face à un enquêteur ou face à une urne, qu'ils dissimulent au premier ce qu'ils acceptent de confier à la seconde" (Blondiaux, 1991: 777).

En effet, "l'entrevue face à face est la plus susceptible d'entraîner des biais de "désirabilité sociale", c'est-à-dire que l'interviewer étant sur les lieux, la tendance pour le répondant à vouloir donner une bonne image de lui-même sera d'autant plus forte" (Blais et Durand, 1997: 377). Dans le cas français, les sondages d'intentions de vote réalisés en face à face ne s'écartent pas plus des marges d'erreur que les sondages téléphoniques. En fait, les graphiques 22 à 25 présentant les séries chronologiques des

quatre regroupements de partis montrent que sept mauvaises estimations sur huit<sup>32</sup> ont été réalisées par des sondages téléphoniques. Le tableau IV permet d'observer les estimations de sondage qui s'écartent des marges d'erreur de l'évolution moyenne des intentions de vote pour les quatre regroupements de partis.

<b>TABLEAU IV</b> <b>LA COMPARAISON DES SONDAGES PRÉSENTANT DES ÉCARTS SIGNIFICATIFS PAR RAPPORT À L'ESTIMATION MOYENNE DES INTENTIONS DE VOTE DES REGROUPEMENTS DE PARTIS</b>							
Nom de l'institut	Date de l'estimation dans la série	Mode de passation	Au moins un jour du week-end	Estimation moyenne de la série	Estimation du sondage	Valeur écart	Marge d'erreur
<i>Selon l'évolution des intentions de vote de l'extrême gauche</i>							
BVA	81	Face à face	oui	8,8	11	2,2	1,6
<i>Selon l'évolution des intentions de vote de la gauche modérée</i>							
IFOP	88	Téléphone	non	40,1	42	1,9	1,5
IFOP	102	Téléphone	non	37,8	39,5	1,7	1,5
<i>Selon l'évolution des intentions de vote de la droite modérée</i>							
IFOP	81	Téléphone	non	35,7	32,5	-3,2	-2,5
IFOP	83	Téléphone	oui	34,7	31	-3,7	-2,5
<i>Selon l'évolution des intentions de vote de l'extrême droite</i>							
IFOP	88	Téléphone	non	12,6	11	-1,6	-1,4
IFOP	102	Téléphone	non	14,8	12,5	-2,3	-1,4
SOFRÉS	108	Téléphone	non	15,7	14	-1,7	-1,5

Dans le tableau IV, la valeur de l'écart correspond à la différence entre l'estimation du sondage et l'évolution moyenne des intentions de vote à une date donnée. Comme chaque valeur dépasse celle de la marge d'erreur, ces sondages s'écartent significativement de l'évolution moyenne des intentions de vote des regroupements.

<sup>32</sup> Les huit estimations proviennent de six sondages. Les sondages réalisés les 28 et 29 mars et 11 et 12 avril de l'Ifop présentent deux estimations qui s'éloignent des marges d'erreur dans des graphiques différents et les quatre autres sondages s'écartent des marges d'erreur une seule fois dans les graphiques des quatre regroupements de partis.

Si l'«effet de prestige» expliquait de façon importante les biais de sondages, les sondages face à face devrait produire de moins bonnes estimations. Or, c'est l'inverse qui s'est produit dans le cas français. Toutefois, il est difficile de départager l'influence du mode de passation du questionnaire et celle de la durée de la collecte puisque les sondages de plus longue durée (3 ou 4 jours) sont plus souvent des sondages face à face que des sondages téléphoniques.

De plus, il est possible de croire que les sondages réalisés pendant la fin de semaine permettent d'interroger les gens les plus difficiles à rejoindre. Sur l'ensemble des trente sondages, treize ont au moins un jour de terrain le week-end et dix-sept ont eu lieu uniquement les jours de la semaine<sup>33</sup>. Est-il possible d'associer cette pratique à une meilleure qualité des estimations? Parmi les huit estimations qui s'éloignent des marges d'erreur des estimations prédites des regroupements de partis, six estimations ont eu lieu uniquement du lundi au vendredi<sup>34</sup>. Le fait de réaliser les sondages au moins un jour pendant la fin de semaine pourrait améliorer la qualité des sondages. Mais des vérifications ultérieures devront être faites puisque peu d'observations sont disponibles dans le cas présent. Et puis, l'institut CSA ne produit aucun sondage ayant au moins un jour de terrain le week-end, mais aucun de ces sondages ne s'écarte des marges d'erreur.

Le tableau IV montre également que l'institut Ifop produit de moins bonnes estimations que les autres instituts (BVA, CSA, Ipsos, Louis Harris et Sofres). Pourtant, la formulation de la question, la population de référence, la taille de l'échantillon, les redressements et le type d'échantillon sont très semblables<sup>35</sup>. Le tableau V présente la comparaison de caractéristiques méthodologiques de l'institut Ifop et des autres sondeurs.

---

<sup>33</sup> Voir le tableau VI à l'annexe A.

<sup>34</sup> Voir le tableau IV.

<sup>35</sup> Voir le tableau VII qui présente la méthodologie du dernier sondage de chaque institut à l'annexe B.

TABLEAU V LA COMPARAISON DES CARACTÉRISTIQUES MÉTHODOLOGIQUES DE L'INSTITUT IFOP ET DES AUTRES INSTITUTS				
	Mode de passation	Nombre de jours de terrain	Au moins un jour de week-end	Taux d'abstention
Ifop	6 téléphoniques 0 face à face	2	154	10,5%
Autres instituts	17 téléphoniques 7 face à face	2,2	10/24	24,4%

Le tableau V montre que l'institut Ifop réalise seulement des sondages téléphoniques comme trois autres sondeurs. En tout, les autres sondeurs produisent dix-sept sondages téléphoniques et sept sondages face à face. Le nombre de jours de terrain est similaire. En moyenne, Ifop réalise ses sondages sur 2 jours et les autres instituts sur 2,2 jours. L'institut Ifop produit proportionnellement plus que les autres sondeurs de sondages réalisés durant au moins un jour de week-end et pourtant la qualité de ses estimations est moindre. Finalement, le taux d'abstention de l'institut Ifop est en moyenne beaucoup plus faible que celui des autres instituts. Malgré toutes ces informations, il est impossible d'expliquer pourquoi les estimations de l'institut Ifop, comparativement à celles des autres, s'écartent plus souvent de l'évolution des intentions de vote moyennes.

Enfin, les instituts ont pratiquement tous enregistré la forte tendance à la baisse du vote Jospin et la forte tendance à la hausse du vote Le Pen<sup>36</sup>. Pourtant, aucun sondeur n'a osé prédire la qualification de Le Pen pour le second tour. La théorie de la "spirale du silence" de Noëlle-Neumann (1984) pourrait expliquer cette situation. En considérant certaines opinions comme étant plus désirables que d'autres, les sondeurs ont préféré se taire sur la possibilité d'un second tour Chirac/Le Pen. Dans les sondages, cette théorie s'applique habituellement aux répondants. Les opinions "illégitimes" peuvent être tuées soit par les fausses déclarations, soit par les refus de répondre au sondage ou à la question d'intention de vote. Mais dans ce cas-ci, la théorie pourrait s'appliquer également aux

<sup>36</sup> Voir les graphiques 2 et 11 des candidats Jospin et Le Pen.

sondeurs puisque ceux-ci utilisent un mode de redressement où leurs propres perceptions jouent un rôle.

### 6.3 Redressements

Les redressements résultent d'une combinaison de corrections: le vote d'une élection antérieure, la certitude d'aller voter, l' "expérience professionnelle des sondeurs", etc. En fait, les instituts modifient les résultats bruts des sondages par diverses corrections et publient les résultats redressés. D'après les informations recueillies dans les fiches techniques disponibles à la Commission des sondages, les sondeurs effectuent des redressements afin de corriger entre autres la sous-estimation du Front national. À titre d'exemple, dans la fiche fournie par Ipsos, le redressement effectué pour le premier tour de l'élection présidentielle de 2002<sup>37</sup>. À partir de la base non-redressée, le vote de 1997 du Front national est de 5,7% des intentions de vote et le résultat redressé de 15,4%<sup>38</sup>. Le redressement peut également corriger la sous-estimation de la gauche modérée<sup>39</sup>.

Une partie de la sous-estimation de l'extrême droite, du candidat Le Pen et la surestimation de la gauche modérée pourrait s'expliquer par les fausses déclarations ou par les non-déclarations à la question sur l'intention de vote. Les électeurs avouent plus facilement un vote pour la gauche que pour l'extrême droite (Boy et Chiche, 1999a), ce qui entraîne une surestimation de la gauche et une sous-estimation de l'extrême droite. Comme les sondeurs connaissent ce phénomène, ils tentent de le corriger en redressant les résultats de sondage. Cependant, il est possible d'expliquer ces biais de sondages par d'autres raisons.

La non-réponse à la question des intentions de vote moyenne des sondages étant de 24,4%, il y a lieu de se questionner sur cette catégorie de répondants. Certains auteurs

---

<sup>37</sup> L'institut Ipsos est celui qui fournit le plus d'informations à la disposition du public. Il est possible de croire que les autres sondeurs pratiquent des redressements semblables.

<sup>38</sup> Voir les redressements effectués par Ipsos pour le sondage réalisé les 17 et 18 avril 2002 à l'annexe F.

<sup>39</sup> Voir l'annexe F.

(Durand et al., 2001, 2002; Curtice, 1997; Jowell et al., 1993) expliquent que les gens qui refusent de dévoiler leurs intentions de vote sont plus souvent à droite de l'échiquier politique. Sachant que ces derniers se répartissent différemment de ceux qui déclarent leurs intentions de vote, quelles répartitions les sondeurs devraient-ils effectuer de cette catégorie de répondants? La répartition proportionnelle est-elle appropriée dans les circonstances?

Les erreurs de prédiction peuvent être attribuées aux non-répondants aux sondages. Comme, "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), imputer les résultats des répondants aux non-répondants n'est pas sans risque de biais" (Vachon, 1998: 22).

La pratique des redressements de résultats amène différentes interrogations. Tout d'abord, les réalités électorales peuvent être différentes d'une élection à l'autre et les redressements politiques sur les votes antérieurs peuvent fausser les calculs. De plus, selon les fiches techniques de la Commission des sondages, les gens qui révèlent leur vote antérieur sont peu nombreux. À titre d'exemple, la fiche technique du sondage réalisé les 17 et 18 avril 2002 par l'institut Ipsos rend compte d'un redressement sur le vote aux élections législatives de 1997. Seulement 598 répondants sur 1103 (54,2%) ont révélé leur vote à ces élections<sup>40</sup>. Le sondage de la Sofres réalisé les 17 et 18 avril 2002 présente un taux de non-réponse de 47% dans les reconstitutions des votes antérieurs<sup>41</sup>. La reconstitution du vote au premier tour de l'élection présidentielle de 1995 montre une "abstention, sans réponse<sup>42</sup>" de 32%<sup>43</sup>. Et puis, la reconstitution du vote au deuxième

---

<sup>40</sup> Voir l'annexe F.

<sup>41</sup> Voir le tableau XVI sur la reconstitution du vote aux élections législatives de 1997 de l'institut Sofres pour le sondage réalisé les 17 et 18 avril à l'annexe G.

<sup>42</sup> "Abstention, sans réponse" regroupe les non-voteurs par choix, les non-voteurs qui n'avaient pas le droit de vote selon leur âge ou leur statut, les discrets, les voteurs qui refusent de dévoiler leur vote antérieur et les voteurs qui ne se souviennent pas de leur vote antérieur.

<sup>43</sup> Voir le tableau XVII sur la reconstitution du vote au premier tour de l'élection présidentielle de 1995 de l'institut Sofres pour le sondage réalisé les 17 et 18 avril à l'annexe G.

tour de l'élection présidentielle de 1995 est de 29%<sup>44</sup>. Cette pratique, en plus de ne pas toujours tenir compte de la nouvelle réalité politique, semble se baser sur de faibles souvenirs de vote. Pour cette raison, la pratique du redressement selon le vote antérieur avec de faibles proportions de rappels semble peu rigoureuse.

La pratique de redressements demeure dans une certaine mesure un "mystère" et elle semble très subjective. Il est impossible à l'heure actuelle de savoir dans quelle proportion les sondeurs se basent sur les redressements et à partir de quels redressements ils arrivent aux résultats finaux, aux intentions de vote publiées. La Commission des sondages aurait avantage à rendre accessibles les informations sur les redressements afin de rendre la démarche des instituts "transparente".

---

<sup>44</sup> Voir le tableau XVIII sur le reconstitution du vote au deuxième tour de l'élection présidentielle de 1995 de l'institut Sofres pour le sondage réalisé les 17 et 18 avril 2002 à l'annexe G.



## CONCLUSION

En conclusion, un bref retour sur les six chapitres de ce mémoire permettra de résumer la démarche de cette recherche portant sur les erreurs de prédiction des sondages au premier tour de l'élection présidentielle française de 2002. De plus, les nombreuses limites, les principales découvertes et quelques pistes de recherche seront exposées.

Le premier chapitre a permis de comprendre l'avènement des sondages d'opinion. En France, c'est à partir de 1965 que les hommes politiques, les scientifiques et les journalistes ont accordé une place de choix aux sondages. La situation politique française a favorisé le recours aux sondages afin de mesurer les rapports entre les coalitions et à l'intérieur de celles-ci, les cotes de popularité des gouvernants, les intérêts des citoyens, etc.

Au second chapitre, afin d'analyser les sondages du premier tour de l'élection présidentielle de 2002 il s'est avéré nécessaire d'examiner de plus près les cas types d'erreurs de prédiction électorales connues dans différents pays et plus particulièrement en France. Un premier type d'erreur de prédiction électorale qui s'est reproduit dans le temps et dans différents pays est la surestimation de la gauche et la sous-estimation de la droite. Un second type d'erreur de prédiction est la sous-estimation des partis politiques manifestant des opinions extrêmes. Au cours de la V<sup>e</sup> République (1965 à 1995), les élections présidentielles françaises présentent effectivement une sous-estimation de l'extrême droite et du parti du Front national. Toutefois, c'est plutôt la surestimation de la droite et la sous-estimation de la gauche qui sont observées lors de ces élections. Comment expliquer ces erreurs de prédiction? La méthodologie de sondage peut-elle être reliée à ces erreurs?

Au troisième chapitre, les explications d'ordres psychosociologique et méthodologique proposées par plusieurs auteurs ont permis de tracer un portrait des explications possibles

des erreurs de prédiction. L'étude de la campagne électorale, des résultats du vote et de la méthodologie utilisée par les instituts français au premier tour de l'élection présidentielle de 2002 a permis de poser des hypothèses de recherche quant aux explications possibles des erreurs de prédiction.

Au quatrième chapitre, le recours à l'analyse de série chronologique ARIMA s'est avéré nécessaire pour traiter le problème des erreurs de prédiction. Cette analyse nous a permis d'estimer l'évolution des intentions de vote pour les candidats et pour les regroupements de partis (extrême gauche, gauche modérée, droite modérée et extrême droite), puis d'évaluer la qualité des prédictions. De plus, l'analyse de l'estimation moyenne des intentions de vote des regroupements de partis a permis d'évaluer la qualité du travail de chacun des sondeurs.

Ensuite, le cinquième chapitre a présenté les résultats des analyses. Comme dans plusieurs pays et contrairement à l'histoire des prédictions de sondages en France, les sondages du premier tour présentent une surestimation de la gauche modérée. L'erreur de prédiction la plus importante est sans aucun doute la sous-estimation de l'extrême droite et tout particulièrement la sous-estimation de Le Pen, le candidat du Front national. Ces écarts entre la prédiction des sondages et le résultat du vote étaient-ils reliés à certaines caractéristiques méthodologiques?

Les explications possibles ont été évoquées au sixième chapitre. Malheureusement, plusieurs hypothèses intéressantes au sujet des biais de sondages n'ont pu être vérifiées faute de données adéquates en provenance des instituts et de la Commission des sondages.

Tout d'abord, les données disponibles dans le cadre de cette recherche n'ont pas permis de savoir si les répondants changent d'idée au dernier moment (late swing). Les résultats des deux premières vagues du "Panel électoral français 2002" (une avant et une après le premier tour) permettraient peut-être de vérifier si les gens ont changé d'idée entre le moment du sondage et le jour du scrutin.

La méthode des quotas pratiquée par l'ensemble des sondeurs a rendu difficile la vérification de l'impact de ce mode d'échantillonnage. D'une part, l'homogénéité du travail des instituts ne permet pas de comparer leur mode d'échantillonnage. D'autre part, les quotas peuvent présenter des biais, et ce, pour l'ensemble des instituts.

Les instituts français en utilisant la méthode des quotas ont pu entraîner une "sous-représentation des personnes les moins coopératives en raison de la moins forte insistance sur le taux de réponse" (Vachon et al., 1999: 559). En fait, il est possible de croire que les instituts français ne se soient pas référés au taux de réponse puisque ce critère n'est pas tenu en compte dans la logique des quotas.

Les erreurs de prédictions peuvent être associées à la sous-représentation de certaines catégories de personnes. Par exemple, les gens non-listés vivant en institution qui n'ont pas accès à un numéro de téléphone privé ne peuvent être sondés. Ces gens sont généralement plus à droite, ce qui peut entraîner une sous-estimation de la droite (Durand et al., 2001, 2002).

De plus, la méthode des quotas produit des échantillons ayant un niveau de diplôme plus élevé que dans la population. Considérant que le Front national recueille des intentions de vote élevées auprès des chômeurs, des ouvriers et des personnes de faible niveau d'étude, il est possible que la structure de l'échantillon ait entraîné une sous-estimation des électeurs du Front national. Comme la sous-estimation des gens faiblement diplômés se reproduit également dans les sondages aléatoires, il serait intéressant de développer une technique pour arriver à rejoindre ces gens.

Également, la non-réponse à la question sur l'intention de vote moyenne dans les sondages est de 24,4%<sup>45</sup>. Les gens qui refusent de dévoiler leurs intentions de vote sont plus souvent à droite de l'échiquier politique (Durand et al., 2001, 2002; Curtice, 1997; Jowell et al., 1993). Comme ces derniers se répartissent différemment de ceux qui

---

<sup>22</sup>Ce taux d'abstention moyen exclut les résultats de l'institut Ifop.

déclarent leurs intentions de vote, il est possible que les sondeurs n'aient pas effectué une répartition appropriée de ces répondants.

La pratique des redressements demeure un grand "mystère" et ce malgré les efforts pour accéder à ces informations. Des études supplémentaires et l'accès aux données brutes des instituts de sondages permettraient de connaître les détails de cette pratique. Un questionnement demeure quant à la pratique des redressements. Les spécialistes des instituts de sondages se réfèrent à des redressements sur les votes aux élections présidentielles de 1995, aux élections législatives de 1997, aux élections européennes de 1999; puis certains redressements politiques portent sur la certitude d'aller voter, la sûreté du choix de vote ou la fermeté du choix, l'intention de participation et la préférence partisane afin de déterminer les intentions de vote du premier tour de l'élection présidentielle française de 2002. Mais de quelle façon les instituts combinent-ils les différents redressements? Les sondeurs tiennent-ils compte des "non-exprimés" sur les questions des votes antérieurs? Les instituts de sondages font-ils des répartitions des "non-exprimés" selon les candidats ou les partis? De plus, le redressement ou l'ajustement selon "l'expérience professionnelle" du sondeur est une pratique qui demeure plus ou moins scientifique.

Malgré les nombreuses limites de ce mémoire, certains points ont pu être vérifiés. La première découverte est que le mode de passation face à face, comparativement au questionnaire téléphonique, ne peut être associé à un plus grand besoin de bien paraître dans la relation enquêteur/répondant dans le cas français 2002. Dans ce cas-ci, les sondages réalisés en face à face connaissent de meilleures estimations que les sondages téléphoniques. Cependant, il faut tenir compte que les sondages face à face ont lieu plus souvent sur de plus longues périodes de collecte que les sondages téléphoniques, ce qui permet d'améliorer la qualité des estimations.

La seconde découverte concerne la différence entre le terrain réalisé au moins pendant un jour de la fin de semaine ou seulement la semaine. Dans le cas présent, il a été possible d'observer que les sondages menés pendant au moins un jour du week-end sont

de meilleure qualité. Ce résultat vient renforcer l'hypothèse que les instituts qui déploient des efforts pour sonder les gens plus difficiles à rejoindre améliorent la qualité de leur échantillon. Mais à quel moment de la journée sont réalisés les sondages? Un indicateur de la qualité des estimations pourrait être les efforts effectués pour rejoindre cette catégorie de gens, par exemple, en soirée.

La troisième découverte est que les estimations de l'institut Ifop s'écartent davantage des marges d'erreur que celles des autres instituts. La proportion des gens qui ne se prononcent pas est faible comparativement à celle des autres instituts, ce qui amène un questionnement sur la gestion de terrain. Toutefois, les informations méthodologiques (le redressement, la durée de la collecte, le moment de la collecte (sondage réalisé pendant au moins un jour du week-end ou la semaine seulement), le mode de passation, la taille de l'échantillon, le type d'échantillon, la proportion d'abstention, la formulation de la question et le type d'enquête) n'ont pas permis d'expliquer en quoi la méthodologie de l'institut Ifop se distingue de celles des autres sondeurs. La consultation des fiches techniques détaillées à la Commission des sondages aurait permis de comparer les procédures spécifiques de chaque institut et d'y attribuer, le cas échéant, les explications des erreurs d'estimation et de prédiction.

Finalement, les sondeurs avaient bien enregistré les lourdes tendances à la baisse de Jospin et à la hausse pour Le Pen, sans toutefois oser prédire ce qui est arrivé au premier tour, c'est-à-dire la qualification de Le Pen au second tour. Dans le cas français, la théorie de la "spirale de silence" pourrait s'appliquer aux sondeurs. Les redressements utilisés par les instituts, où leurs propres perceptions jouent un rôle, ont vraisemblablement conduit à leur erreur de prédiction. Tout laisse croire que la plus grande erreur de prédiction du premier tour de l'élection présidentielle française de 2002 serait la "spirale du silence des sondeurs".

## Bibliographie

- Beaud, Jean-Pierre (1992), "L'échantillonnage" dans Gauthier, Benoit (Ed.) : *Recherche sociale; de la problématique à la collecte de données*. Presses de l'Université du Québec, Sillery, Québec, p. 185-216.
- Belli, Robert F., Traugott, Michael W., Young, Margaret et Katherine A. McGonagle (1999), "Reducing Vote Overreporting in surveys, Social Desirability, Memory Failure, and Source Monitoring", *Public Opinion Quarterly*, Vol. 63, p. 90-108.
- Blais, André, Durand, Claire et Marc-André Bodet (2003), "The Dynamics of Vote Intentions in the First Round of the 2002 French Presidential Election", *French Politics*, 1, p. 369-375.
- Blais, André et Claire Durand (1997), "Le sondage" dans Gauthier, Benoit (Ed.) : *Recherche sociale; de la problématique à la collecte de données*. Presses de l'Université du Québec, Sillery, Québec, p. 357-400.
- Blondiaux, Loïc (1998), *La fabrique de l'opinion, une histoire sociale des sondages*, Paris, Seuil.
- Blondiaux, Loïc (1991), "L'invention des sondages d'opinion", *Revue française de science politique*, 41 (6), p. 756-780.
- Blondiaux, Loïc (1990), "Paul F. Lazarsfeld (1901-1976) et Jean Stoetzel (1910-1987) et les sondages d'opinion : genèse d'un discours scientifique", *Mots*, no 23, p. 5-22.
- Bon, Frédéric (1974), *Les sondages se trompent-ils?*, Paris, Calmann-Lévy.
- Bourdieu, Pierre (1985), "Remarques à propos de la valeur scientifique et des effets politiques des enquêtes d'opinion", *Pouvoirs*, no 33, p. 131-139.
- Bourdieu, Pierre (1973), "L'opinion publique n'existe pas", *Les Temps modernes*, no 318, 29<sup>e</sup> année, p. 1292-1309.
- Boy, Daniel et Jean Chiche (1999a), "La qualité des enquêtes d'intention de vote : le cas des régionales de 1998" dans Sofres, *L'état de l'opinion 1999*, Paris, Editions du Seuil.
- Boy, Daniel et Jean Chiche (1999b), "Pourquoi les sondeurs se trompent", *Libération*, 22 juin 1999.
- Bréchon, Pierre (dir.) (2002), *Les élections présidentielles en France, quarante ans d'histoire politique*, Paris, Les Études de La Documentation française.

Breen, Richard (2000), "Why Is Support for Extreme Parties Underestimated by Surveys? A Latent Class Analysis", *British Journal of Political Science*, 30, 2, p. 375-382.

Brûlé, Michel (1988), *L'empire des sondages, transparence ou manipulation?*, Paris, Éditions Robert Laffont.

Cabin, Philippe (1999), "La construction de l'opinion publique", *Sciences humaines*, no 96, p. 12-15.

Cantrell, Paul (1992), "Opinion polling and American democratic culture", *International Journal of Politics, Culture and Society*, vol. 5, no 3, p. 415-416.

Cayrol, Roland (2000), *Sondages, mode d'emploi*, Paris, Presses de Sciences Po.

Champagne, Patrick (1990), *Faire l'opinion, le nouveau jeu politique*, Paris, Les Éditions de Minuit.

Chevalier, Serge et Claire Durand (1997), *Le taux de réponse dans les sondages téléphoniques, déterminants et conséquence*, 65e Congrès de l'ACFAS (Montréal : Université du Québec à Trois-Rivières).

Chiche, Jean et Dominique Reynié (2002), "La France en dépression électorale" dans Sofres, *L'état de l'opinion 2002*, Paris, Editions du Seuil.

Courtial, Jean-Pierre (1977), "Les sondages peuvent-ils se tromper", *Psychologie*, no 91, p. 43-47.

Coutrot, Bernard et Fernand Droesbeke (1984), *Les méthodes de prévision*, Paris, Presses universitaires de France.

Cowans, Jon (2002), "Fear and Loathing in Paris, The Reception of Opinion Polling in France, 1938-1977", *Social Science History*, no 26, 1, p. 71-104.

Curtice, John (1997), "So how well did they do? The polls in the 1997 Election", *Journal of the Market Research Society*, 39, no 3, p. 449-461.

Curtice, John et Nick Sparrow (1997), "How accurate are traditional quota opinion polls?", *Journal of the Market Research Society*, 39 (3), p. 433-448.

Droesbeke, Jean-Jacques, Fichet, Bernard et Philippe Tassi (1989), *Séries chronologiques, Théorie et pratique des modèles ARIMA*, Paris, Economica.

Durand, Claire, Blais André et Mylène Larochelle (2003a) (*à paraître*), "The Polls in the 2002 French Presidential Election: An Autopsy", *Public Opinion Quarterly*.

Durand, Claire, Blais André et Mylène Larochelle (2002), "Sondages et estimation du vote, une comparaison France-Canada", Colloque francophone sur les sondages, Grenoble-Autrans 2002.

Durand, Claire (2002), "Are the Polls Biased Against Quebec Liberals?", *Policy Options*, p. 51-56.

Durand, Claire, Blais, André et Sébastien Vachon (2002), "Accounting For Biases in Election Surveys : The Case of the 1998 Quebec Election », *Journal of Official Statistics*, 18, no 1, p. 25-44.

Durand, Claire, Blais, André et Sébastien Vachon (2001). « A Late Campaign Swing or a Failure of the Polls? The Case of the 1998 Quebec Election », *Public Opinion Quarterly*, 108-123.

Durand, Claire (1999), *L'échantillonnage et la gestion du terrain*, notes de cours.

Gazier, François et Ronny Abraham (1995), "La Commission des sondages face à l'élection présidentielle et aux élections municipales de 1995", *Pouvoirs*, no 75, p. 139-147.

Jadot, Anne (2001), "Sondages et votes : les leçons méthodologiques des élections générales britanniques", dans *Enquêtes, modèles et applications*, sous la direction de DROESBEKE, J.-J. et L. LEBART, Paris, Dunod.

Jaffré, Jérôme (1985), "Réflexions sur la "sondomanie"", *Pouvoirs*, no 33, p. 15-23.

Jaffré, Jérôme (1988), "Incertitudes et avenir de la démocratie des sondages", *Le Débat*, no 52, p. 41-54.

Javeau, Claude (1982), *L'enquête par questionnaire, manuel à l'usage du praticien*, Les Editions d'Organisation, Paris.

Jowell, Roger, Hedges, Barry, Lynn, Peter, Farrant, Graham et Anthony Heath (1993), "The 1992 British Election : the Failure of the Polls", *Public Opinion Quarterly*, 57, p. 238-263.

Lau, Richard R. (1994), "An Analysis of the Accuracy of "Trial Heat" Polls During The 1992 Presidential Election", *Public Opinion Quarterly*, 57, p. 238-263.

Lech, Jean- Marc (2001), *Sondages privés, Les secrets de l'opinion*, Paris, Stock.

Marcel, Jean-Christophe (1998), "Jean Stoetzel élève de Maurice Halbwachs : les origines françaises de la théorie des opinions", *L'Année sociologique*, no 48, 2, p. 319-351.



March, Catherine et Elinor Scarborough (1990), "Testing nine hypotheses about quota sampling", *Journal of the Market Research Society*, 32 (4), p. 485-506.

Max, Alfred (1981), *La république des sondages*, Paris, Gallimard.

McElroy, Gail et Michael Marsh, "Why the Opinion Polls Got it Wrong in 2002" in Gallagher, Michael, Marsh, Michael et Paul Mitchell (eds.), *How Ireland Voted 2002*, London, Palgrave, p. 159-176.

Meynaud, Hélène Y. et Denis Duclos (1985), *Les sondages d'opinion*, Paris, Éditions La Découverte.

Mitofsky, Warren (1998), "Was 1996 a Worse Year for Polls than 1948?", *Public Opinion Quarterly*, 62, p. 230-249.

Noëlle-Neumann, Elisabeth (1984), *The Spiral of Silence*, Chicago, The University of Chicago Press.

Ostrom, Charles W. (1990). *Time Series Analysis, Regression Techniques*, Newbury Park : SAGE Publications, 95 p.

Palm, Rodolphe (1992), *Comment interpréter les résultats d'une série chronologique*, ITCF.

Passelecq, Olivier (dir.) (1995), *Modes de scrutin et systèmes électoraux*, Documents d'études, droit constitutionnel et institutions politiques, no 1.05.

Panagakis, Nick (1999), "Response to "Was 1996 a Worse Year for Polls than 1948?", *Public Opinion Quarterly*, 63, p. 278-281.

Perrineau, Pascal et Colette Ysmal (dir.) (2003), *Le vote de tous les refus, les élections présidentielles et législatives 2002*, Paris, Presses de Sciences Po.

Quivy, Raymond et Luc Van Campenhoudt (1995), *Manuel de recherche en sciences sociales*, Paris, Dunod.

Rivière, Emmanuel (2001), "L'intention de vote existe-t-elle? Les sondages électoraux à nouveau en question", *Revue politique et parlementaire*, no 1011, p. 64-88.

Smith, Tom W. (1983), "The hidden 25 percent: An analysis of nonresponse on the 1980 general social survey", *Public Opinion Quarterly*, 47 (3), p. 50-68.

*SPSS for Windows : Trends, Release 6.0*, Copyright 1993 by SPSS Inc.

Stoetzel, Jean (1982), "Les élections françaises de 1981 et les sondages", *Revue française de sociologie*, 23, 1, p. 3-14.

Stoetzel, Jean et Alain Girard (1973), *Les sondages d'opinion publique*, Paris, Presses Universitaires de France.

Tabachnik, Barbara G. et Linda S. Fidell (1996), *Using multivariate statistics*. New York, Harper Collins. Chapitre 16, p. 837-900.

Triplett Timothy, Blair, Jhonny, Hamilton, Teresa 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.

Traugott, Michael W. (2001), "Assessing Poll Performance in the 2000 Campaign", *Public Opinion Quarterly*, 65, p. 389-419.

Vachon, Sébastien (1998), *L'effet des facteurs méthodologiques, empiriques et sociologiques sur la qualité de l'estimation des intentions de vote*, Mémoire de maîtrise, Université de Montréal.

Vachon, Sébastien, Durand, Claire et André Blais (1999), "Les sondages moins rigoureux sont-ils moins fiables?", *Analyse de politiques*, vol. 25, no 4, p. 557-561.

Sites Internet:

[www.bva.fr](http://www.bva.fr)

[www.csa-tmo.fr](http://www.csa-tmo.fr)

[www.ifop.com](http://www.ifop.com)

[www.canalipsos.com](http://www.canalipsos.com)

[www.louis-harris.fr](http://www.louis-harris.fr)

[www.sofres.com](http://www.sofres.com)

**ANNEXE A**

TABLEAU VI									
LES SONDAGES RÉALISÉS DU 20 MARS AU 18 AVRIL 2002									
Nom de l'institut	Début de la collecte	Fin de la collecte	Au moins un jour de week-end	Nombre de jours de terrain	Taille de l'échantillon	Taux d'abstention (%)	Mode de passation		
CSA	20.03	21.03	NON	2	1000	34	Téléphone		
BVA	21.03	23.03	OUI	3	981	24	Face à face		
IFOP	21.03	22.03	NON	2	947	11	Téléphone		
IPSOS	22.03	23.03	OUI	2	938	24	Téléphone		
SOFRES	22.03	23.03	OUI	2	1000	24	Téléphone		
IFOP	23.03	24.03	OUI	2	1005	10	Téléphone		
SOFRES	27.03	28.03	NON	2	1000	18	Téléphone		
CSA	27.03	28.03	NON	2	1003	35	Téléphone		
SOFRES	27.03	28.03	NON	2	1000	19	Face à face		
BVA	28.03	30.03	OUI	3	928	20	Face à face		
IFOP	28.03	29.03	NON	2	936	12	Téléphone		
IPSOS	29.03	30.03	OUI	2	938	10	Téléphone		
LOUIS HARRIS	29.03	30.03	OUI	2	1001	28	Téléphone		
SOFRES	03.04	04.04	NON	2	1000	18	Téléphone		
CSA	03.04	04.04	NON	2	1000	36	Téléphone		
BVA	04.04	06.04	OUI	3	963	24	Face à face		
IFOP	05.04	06.04	OUI	2	1004	10	Téléphone		

Nom de l'institut	Début de la collecte	Fin de la collecte	Au moins un jour de week-end	Nombre de jours de terrain	Taille de l'échantillon	Taux d'abstention	Mode de passation
IPSOS	05.04	06.04	OUI	2	952	9	Téléphone
SOFRES	10.04	11.04	NON	2	1000	26	Face à face
BVA	10.04	13.04	OUI	4	1082	22	Face à face
CSA	10.04	11.04	NON	2	1000	35	Téléphone
IFOP	11.04	12.04	NON	2	951	11	Téléphone
IFOP	12.04	13.04	OUI	2	1006	9	Téléphone
SOFRES	13.04	15.04	OUI	3	1000	22	Téléphone
IPSOS	15.04	16.04	NON	2	1005	22	Téléphone
LOUIS HARRIS	16.04	17.04	NON	2	1005	31	Téléphone
CSA	17.04	18.04	NON	2	1002	31	Téléphone
IPSOS	17.04	18.04	NON	2	989	19	Téléphone
SOFRES	17.04	18.04	NON	2	1000	30	Téléphone
BVA	18.04	18.04	NON	1	918	25	Face à face

**ANNEXE B**

TABLEAU VII LA MÉTHODOLOGIE DU DERNIER SONDAGE DE CHAQUE INSTITUT						
	BVA (18 avril)	CSA (17 et 18 avril)	IFOP (12 et 13 avril)	IPSOS (17-18 avril)	LOUIS HARRIS (16 et 17 avril)	SOFRES (17 et 18 avril)
Formulation de la question	Si le premier tour de l'élection présidentielle avait lieu dimanche prochain et que vous ayez le choix entre les candidats suivants, pour lequel y aurait-il plus de chances que vous votiez?	N/D	Si dimanche prochain devait se dérouler le premier tour de l'élection présidentielle pour lequel des candidats suivants y aurait-il le plus de chances que vous votiez?	Quel est, parmi les suivants, le candidat pour lequel il y a le plus de chances que vous votiez dimanche prochain au premier tour de l'élection présidentielle?	Si le premier tour de l'élection présidentielle avait lieu dimanche prochain et que vous ayez le choix entre les candidats suivants, pour lequel y aurait-il le plus de chances que vous votiez?	Supposez que le premier tour de l'élection présidentielle ait lieu dimanche prochain et que vous ayez le choix entre les candidats suivants, pour lequel y aurait-il le plus de chances que vous votiez?
Population de référence	Population française âgée de 18 ans et plus	Personnes âgées de 18 ans et plus inscrites sur les listes électorales	Population française âgée de 18 ans et plus, inscrite sur les listes électorales	Population française inscrite sur les listes électorales	Population française âgée de 18 ans et plus	Population française âgée de 18 ans et plus
Mode de passation	Face à face	Téléphone	Téléphone	Téléphone	Téléphone	Téléphone
Taille de l'échantillon	918	1000, (1002 dans <i>Libération</i> )	1006	989	1005	1000

	BVA (18 avril)	CSA (17 et 18 avril)	IFOP (12 et 13 avril)	IPSOS (17-18 avril)	LOUIS HARRIS (16 et 17 avril)	SOFRES (17 et 18 avril)
Nombre de jours de collecte	1	2	2	2	2	2
Redressements	(Sondage 10-13/04) - vote au 1er tour des élections législatives de 1997 - vote au premier tour de l'élection présidentielle de 1995 - vote au second tour de l'élection présidentielle de 1995 -fermé du vote	- certitude d'aller voter - sûreté du choix - vote à l'élection présidentielle de 1995 - vote à l'élection législative de 1997	- vote européen 1999 - vote législatif 1997 1er tour - vote présidentiel 1995 1er et 2e tours - intention de participation - certitude d'aller voter	Le souvenir de vote au 1er tour des élections législatives de 1997.	(16 et 17 avril) - intentions de vote présidentielles du 1 <sup>er</sup> et de 2 <sup>e</sup> tour de 1995 - vote législatif 1997 - certitude d'aller voter - sûreté du vote (29 et 30 mars) - profession de l'interviewé - proximité partisane	- vote au 1er tour des élections législatives de 1997 - vote au 1er et 2e tours des élections présidentielles de 1995 - préférence partisane - fermé du choix (pour l'intention de vote)
Type d'enquête	(Sondage 10-13/04) Omnibus à sujets multiples, les sujets à caractère politique ayant été placé en début de questionnaire.	Enquête insérée en première partie dans un omnibus ne comprenant que des sujets politiques	N/D	Enquête spécifique	N/D	Enquête à sujets multiples
Type d'échantillon	Quotas : sexe, âge, profession du chef de famille, région et catégorie d'agglomération	Quotas (sexe, âge, catégorie socioprofessionnelle du chef de ménage, stratification par région et par agglomération	Quotas (sexe, âge, profession du chef de ménage, région et catégorie d'agglomération	Quotas (sexe, âge, catégorie socio-professionnelle du chef de famille, catégorie d'agglomération et région)	Quotas (sexe, âge, profession du chef de ménage) après stratification par région et catégorie d'agglomération	Quotas (sexe, âge, profession du chef de ménage PCS) et stratification par région et catégorie d'agglomération



**ANNEXE C**

**TABLEAU VIII**  
**LES COEFFICIENTS DE RÉGRESSION DES SÉRIES CHRONOLOGIQUES DES**  
**CANDIDATS DE GAUCHE**

<i>Fonctions linéaires</i>	<i>B</i>	<i>Erreur-type</i>	<i>Valeur de T</i>	<i>P</i>
Jospin	-0,0013	0,00019	-6,72019	0,00000*
Laguiller	-0,00053	0,00045	-1,17964	0,2527
Chevènement	-0,00066	0,00031	-2,14859	0,04477*
Mamère	-0,0001	0,00014	-0,67405	0,5084
Besancenot	0,00107	0,00021	5,14455	0,00006*
Hue	0,00001	0,00013	0,04597	0,96381
Taubira	0,00039	0,00009	4,64756	0,00018*
Gluckstein	0,00019	0,00004	5,44904	0,00003*

\* si  $p \leq 0,05$  la fonction est significative.

**TABLEAU IX**  
**LES COEFFICIENTS DE RÉGRESSION DES SÉRIES CHRONOLOGIQUES DES**  
**CANDIDATS DE DROITE**

<i>Fonctions linéaires</i>	<i>B</i>	<i>Erreur-type</i>	<i>Valeur de T</i>	<i>P</i>
Chirac	-0,00096	0,00036	-2,67838	0,01487*
Le Pen	0,00106	0,00019	5,52618	0,00003*
Bayrou	0,00074	0,00012	6,39583	0,00000*
Saint-Josse	0,00025	0,00019	1,29201	0,21185
Madelin	0,00036	0,00014	2,57491	0,01855*
Mégret	0,00051	0,00018	2,80049	0,01141*
Lepage	0,00026	0,00013	2,0911	0,05019
Boutin	0,00007	0,00011	0,69879	0,49315

\* si  $p \leq 0,05$  la fonction est significative.

**TABLEAU X**  
**LES COEFFICIENTS DE RÉGRESSION DES SÉRIES CHRONOLOGIQUES DES**  
**REGROUPEMENTS DE PARTIS**

<i>Fonctions linéaires</i>	<i>B</i>	<i>Erreur-type</i>	<i>Valeur de T</i>	<i>P</i>
Extrême gauche	0,00059	0,00034	1,74877	0,09647
Gauche modérée	-0,00166	0,00031	-5,37011	0,00004*
Droite modérée	0,00051	0,00037	1,37916	0,18386
Extrême droite	0,00161	0,0003	5,39569	0,00003*

\* si  $p \leq 0,05$  la fonction est significative.

## **ANNEXE D**

**TABLEAU XI**  
**L'ÉVOLUTION DES INTENTIONS DE VOTE POUR LES CANDIDATS DE**  
**GAUCHE DU 20 MARS AU 21 AVRIL 2002**

Candidats	20/03/02	21/04/02	Évolution	Tendance
Jospin	21,2	17,1	- 4,1	↓
Laguiller	8,7	5,6	- 3,1	∩
Chevènement	8,3	6,2	- 2,1	↓
Mamère	6,3	6,0	- 0,3	–
Besancenot	0,0	3,6	3,6	↑
Hue	5,4	5,3	- 0,1	–
Taubira	0,4	1,7	1,3	↑
Gluckstein	0,0	0,6	0,6	↑

**TABLEAU XII**  
**L'ÉVOLUTION DES INTENTIONS DE VOTE POUR LES CANDIDATS DE**  
**DROITE DU 20 MARS AU 21 AVRIL 2002**

Candidats	20/03/02	21/04/02	Évolution de la série	Tendance significative
Chirac	22,6	19,5	-3,1	↓
Le Pen	10,2	13,6	3,4	↑
Bayrou	4	6,3	2,3	↑
Saint-Josse	3	3,9	0,9	–
Madelin	3,3	4,5	1,2	↑
Mégret	1	2,5	1,5	↑
Lepage	0,7	1,3	0,6	–
Boutin	0,9	1,2	0,3	–

**TABLEAU XIII**  
**L'ÉVOLUTION DES INTENTIONS DE VOTE POUR LES REGROUPEMENTS**  
**DE PARTIS DU 20 MARS AU 21 AVRIL 2002**

Regroupements de partis	20/03/02	21/04/02	Évolution	Tendance
Extrême gauche	9,7	11,6	1,9	–
Gauche modérée	41,5	36,3	-5,2	↓
Droite modérée	34,6	36,8	2,2	–
Extrême droite	11,1	16,3	5,2	↑

## **ANNEXE E**

<b>TABLEAU XIV</b>			
<b>LES REDRESSEMENTS EFFECTUÉS PAR L'INSTITUT SOFRES</b>			
<b>POUR LE SONDAGE RÉALISÉ LES 17 ET 18 AVRIL 2002</b>			
	<b>Structure recherchée</b>	<b>Résultats observés</b>	<b>Résultats redressés</b>
<b>Sexe</b>			
Homme	47,7	46,5	47,7
Femme	52,3	53,5	52,3
<b>Âge</b>			
18-24 ans	11,7	9,9	11,7
25-34 ans	18,2	18,5	18,2
35-49 ans	27,7	28,6	27,7
50-64 ans	21,0	22,5	21,0
65 ans et +	21,4	20,5	21,4
<b>PCS chef de ménage</b>			
Agriculteur	2,0	2,0	2,0
Artisan, commerçant, chef d'entreprise	5,4	5,4	5,4
Cadre, profession intellectuelle	11,0	11,9	11,0
Profession intermédiaire	13,9	14,4	13,9
Employé	10,4	10,8	10,4
Ouvrier	21,6	21,0	21,6
Inactif, retraité	35,7	34,5	35,7
<b>Catégorie d'agglomération</b>			
Moins 2000	26,3	25,9	26,3
2000 à 20 000 hab.	17,5	17,1	17,5
20000 à 100000 hab.	13,1	13,2	13,1
100000 hab. et +	29,4	29,7	29,4
Agglomération parisienne	13,7	14,1	13,7
<b>Région</b>			
Nord	10,3	10,3	10,3
Ouest	17,2	17,7	17,2
Sud-Ouest	13,2	13,2	13,2
Sud-Est	21,3	20,8	21,3
Centre	11,2	10,8	11,2
Est	11,4	11,5	11,2
Région parisienne	15,4	15,8	15,4

**ANNEXE F**



<b>TABLEAU XV</b>					
<b>LES REDRESSEMENTS EFFECTUÉS PAR L'INSTITUT IPSOS POUR LE</b>					
<b>SONDAGE RÉALISÉ LES 17 ET 18 AVRIL 2002</b>					
	<b>Base non pondérée</b>	<b>%</b>	<b>Coeff. de pondération</b>	<b>Base pondérée</b>	<b>%</b>
<b>Ensemble</b>	<b>1103</b>	<b>100</b>	<b>1</b>	<b>1103</b>	<b>100</b>
<b>Sexe</b>					
Hommes	529	48	0,997	527	47,8
Femmes	574	52	1,003	576	52,2
<b>Âge</b>					
18-24 ans	119	10,8	1,066	127	11,5
25-34 ans	207	18,8	1,012	210	19,0
35-44 ans	216	19,6	0,986	213	19,3
45-59 ans	273	24,8	0,974	26,6	24,1
60-69 ans	146	13,2	0,997	146	13,2
70 ans et +	142	12,9	1,001	142	12,9
S/t - de 35 ans	326	29,6	1,032	336	30,5
S/t 35 ans et +	777	70,4	0,987	767	69,5
<b>PCS chef de famille</b>					
Agriculteur	24	2,2	0,919	22	2,0
Artisan, commerçant, chef	51	4,6	1,254	64	5,8
Cadre supérieur	120	10,9	0,937	112	10,2
Profession intermédiaire	169	15,3	0,94	159	14,4
Employé	120	10,9	0,947	114	10,3
Ouvrier	240	21,8	1,007	242	21,9
Retraité	321	29,1	0,983	316	28,6
Inactif	58	5,3	1,293	75	6,8
<b>Région</b>					
Ile-de-France	205	18,6	0,997	204	18,5
Nord-Ouest	251	22,8	1,011	254	23,0
Nord-Est	258	23,4	1,01	261	23,6
Sud-Ouest	122	11,1	0,995	121	11,0
Sud-Est	267	24,2	0,984	263	23,8
S/t Province	898	81,4	1,001	899	81,5
<b>Catégorie d'agglomération</b>					
Rural	268	24,3	1,012	271	24,6
- de 20 000 hab.	186	16,9	0,996	185	16,8
20000 à 100000 hab.	148	13,4	0,99	147	13,3
+ de 100000 hab.	319	28,9	0,995	317	28,8
Agglomération de Paris	182	16,5	1,004	183	16,6
Exprimés Leg 97	<b>598</b>	<b>100</b>	<b>1</b>	<b>598</b>	<b>100</b>

<b>Législatives 97 1er tour</b>					
EXG	13	2,2	1,013	13	2,2
PC	43	7,2	1,403	60	10,1
PS-DVG	207	<b>34,6</b>	0,846	175	<b>29,3</b>
ECO	54	9	0,72	39	6,5
UDF-RPR-DVD	247	41,3	0,884	218	36,5
FN	34	<b>5,7</b>	2,706	92	<b>15,4</b>

# **ANNEXE G**

**TABLEAU XVI**  
**LA RECONSTITUTION DU VOTE AUX ÉLECTIONS**  
**LÉGISLATIVES DE 1997 PAR L'INSTITUT SOFRES**  
**POUR LE SONDAGE RÉALISÉ LES 17 ET 18 AVRIL 2002**

	<b>% des suffrages exprimés</b>
Parti communiste	94
Extrême gauche	4
Parti socialiste	351
Divers gauche	32
Verts	49
Autres écologistes	16
UDF	170
RPR	190
Droite indépendante	19
Divers droite	23
FN	48
Divers	3
Total	100
Abstention, sans réponse	47

**TABLEAU XVII**  
**LA RECONSTITUTION DU VOTE AU PREMIER TOUR**  
**DE L'ÉLECTION PRÉSIDENTIELLE DE 1995 PAR**  
**L'INSTITUT SOFRES POUR LE SONDAGE RÉALISÉ**  
**LES 17 ET 18 AVRIL 2002**

	<b>% des suffrages exprimés</b>
R. Hue	47
A. Laguiller	45
L. Jospin	307
D. Voynet	33
J. Chirac	394
E. Balladur	72
P. De Villiers	15
J-M. Le Pen	65
J. Cheminage	1
Total	100
Abstention, sans réponse	32

**TABLEAU XVIII**  
**LA RECONSTITUTION DU VOTE AU DEUXIÈME**  
**TOUR DE L'ÉLECTION PRÉSIDENTIELLE DE 1995**  
**PAR L'INSTITUT SOFRES POUR LE SONDAGE**  
**RÉALISÉ LES 17 ET 18 AVRIL 2002**

	<b>% des suffrages exprimés</b>
L. Jospin	436
J. Chirac	564
Total	1000
Abstention, sans réponse	29

8300 10000 1303