



## Économie publique/Public economics

22-23 | 2008/1-2  
Varia

---

# Économie, politique et résultats des élections présidentielles françaises

Antoine Auberger

---



### Édition électronique

URL : <http://journals.openedition.org/economiepublique/8467>  
ISSN : 1778-7440

### Éditeur

IDEP - Institut d'économie publique

### Édition imprimée

Date de publication : 20 août 2012  
Pagination : 247-271  
ISSN : 1373-8496

### Référence électronique

Antoine Auberger, « Économie, politique et résultats des élections présidentielles françaises », *Économie publique/Public economics* [En ligne], 22-23 | 2008/1-2, mis en ligne le 04 octobre 2012, consulté le 19 avril 2019. URL : <http://journals.openedition.org/economiepublique/8467>

---

# économie publique public economics

Revue de l'**Institut d'Économie Publique**

Deux numéros par an

**n° 22-23** – 2008/1-2



## Économie, politique et résultats des élections présidentielles françaises

Antoine Auberger \*

### Résumé

Cet article a pour objet la construction et l'estimation d'un modèle permettant d'expliquer et de prévoir le résultat du second tour des élections présidentielles françaises par département et au niveau national (1988-1995 et 1988-2007 ; les élections de 2002 sont laissées de côté car elles ne rentrent pas dans le cadre de notre article). Ce modèle met en évidence le rôle des facteurs politiques et économiques dans l'explication du vote pour la gauche au second tour des élections présidentielles françaises (en cas de duel gauche/droite). Les prévisions pour l'élection présidentielle de 2007 sont satisfaisantes.

### Summary

This article's aim consists in building and estimating a model which explains and forecasts the outcome of the second round of the french presidential elections by department and at the national level (1988-1995 and 1988-2007 without 2002). This model emphasises the role of the political and economic factors in the explanation of the presidential

---

\*. Université Paris 2 et associé au *CES-LAEP MATISSE*. Université Paris 1, Maison des sciences économiques, 106-112, bd de l'Hôpital, 75647 Paris cedex 13 ; Email : antoine.auberger@univ-paris1.fr. L'auteur remercie les participants du premier colloque mondial de *Public Choice* (29 mars-1<sup>er</sup> avril 2007), du séminaire d'économie publique du *LAEP* et *MATISSE*, université Paris 1 (30 janvier 2007), de l'atelier de l'*IRGEI* sur les prévisions électorales (3 mai 2007) et du séminaire du *CREM*, université de Caen (3 mars 2008) pour leurs commentaires sur des versions antérieures ainsi qu'un rapporteur anonyme pour ses remarques et suggestions qui ont permis d'améliorer cet article.

vote received by the Left at the second round with a duel Left/Right). Its forecasts for the 2007 French presidential election are satisfactory.

**Mots-clés :** Économie du vote, prévision électorale, élections présidentielles françaises.

**Keywords:** Economic Voting, Electoral Forecasting, French Presidential Elections.

J.E.L. : D72

---

## 1. Introduction

La construction et l'estimation de fonctions de vote (ou de modèles économiques/économétriques du vote) est un thème de recherches de l'économie publique. De nombreux travaux mettent notamment en évidence l'influence de l'économie sur le résultat des élections. Le premier article consacré à l'estimation d'une fonction de vote est écrit par Kramer (1971) sur les États-Unis (élections à la chambre des représentants, 1896-1964). Kramer (1971) montre que le revenu personnel réel est la variable économique qui a le plus d'influence sur le résultat de ces élections. Stigler (1973) critique l'article de Kramer (1971) et montre que le revenu personnel réel n'a pas une influence importante sur le résultat des élections à la chambre des représentants (1900-1932 et 1934-1970) en supposant que les électeurs ont un comportement plus sophistiqué et sont mieux informés que dans le modèle de Kramer (1971). Fair (1978) construit un modèle de vote plus général pour analyser l'influence de l'économie sur le vote aux élections présidentielles américaines et montre que le taux de croissance du PNB réel a une influence significative sur le résultat de ces élections (1916-1976). Depuis, de nombreuses fonctions de vote ont été construites et estimées pour différents pays et on trouve notamment dans Mueller (2003) et Auberge (2001) une synthèse de l'influence des différentes variables économiques sur le résultat des élections. Les modèles économétriques du vote permettent d'expliquer le résultat des élections passées (prévisions *ex post*) et de faire des prévisions avant une élection (prévisions *ex ante*). Aux États-Unis, les résultats de certaines élections présidentielles ont été bien prévus (l'élection présidentielle de 1996, par exemple) mais pour d'autres élections, les modèles économétriques du vote ont proposé des prévisions décevantes (l'élection présidentielle de 2000).

De nombreux modèles économétriques du vote ont été développés en France depuis la fin des années 1970 pour expliquer et prévoir le résultat des élections nationales et locales comme le montre le *survey* de Dubois (2007). Les premiers modèles construits pour les élections présidentielles sont des modèles nationaux utilisant des données nationales<sup>1</sup>. Par exemple, les modèles de Courbis (1995) et de Lewis-Beck (1995, 1997) utilisent des données nationales mais comme le nombre d'élections présidentielles est faible (seulement sept élections de 1965 à 2002), l'utilisation de ces modèles pour expliquer et prévoir le vote aux élections présidentielles est assez délicate. Lewis-Beck (1991) utilise un modèle unique pour les élections présidentielles et législatives, ce qui augmente naturellement le nombre d'observations qui est alors égal à quatorze mais la variable économique (croissance du PNB) n'est pas statistiquement significative et les prévisions *ex post* sont moins bonnes que celles de Lewis-Beck (1995). Dans le modèle de Lewis-Beck (1995), la variable économique (croissance du PIB) et celle de popularité (popularité du président) ne sont pas simultanément statistiquement significatives et le coefficient de la variable popularité n'a pas le signe attendu. De plus, ce modèle prévoyait pour l'élection présidentielle de 1995 la victoire de la gauche et cette erreur s'explique car l'auteur suppose que la majorité présidentielle sortante était jugée responsable de la situation économique pendant une période de cohabitation alors que le président sortant ne se représentait pas. Le modèle de l'Iowa (Bélanger, Fauvelle-Aymar et Lewis-Beck, 2007) explique le vote pour la gauche au premier tour des élections présidentielles à l'aide du taux de chômage et de la popularité du président de la République et il prédisait la victoire de Ségolène Royal (52,7 %) au second tour de l'élection présidentielle de 2007 avec une marge d'erreur rendant possible la défaite de Ségolène Royal. On constate donc que ces modèles nationaux présentent des faiblesses sur le plan économétrique et qu'ils proposent des prévisions incertaines. Pour remédier au très faible nombre d'élections présidentielles, il est possible d'utiliser des données régionales ou départementales (modèles avec des données de panel<sup>2</sup>), ce qui permet d'employer éventuellement davantage de variables explicatives. Les premiers modèles pour la France ont été construits pour les élections législatives et sont plus adaptés que les modèles nationaux à l'établissement de prévisions *ex ante* précises en sièges : Jérôme, Lafay et Lewis-Beck (1993) pour les élections de 1993 et Jérôme, Jérôme-Speziari et Lewis-Beck (1999) pour celles de 1997. Pour les élections présidentielles, Dubois (2002) est le premier à construire un modèle de prévision par département, alors que Jérôme, Jérôme-Speziari et Lewis-Beck (2003), Jérôme et Jérôme-Speziari (2004b) et Fauvelle-Aymar, Lafay

---

1. On trouve dans Auberger (2004) une synthèse des principaux modèles nationaux pour les élections présidentielles américaines et françaises.

2. Sevestre (2003) fait une présentation en français de l'économétrie des données de panel. Pour une présentation en anglais, on peut notamment consulter Baltagi (2003), Hsiao (2003) et Wooldridge (2002).

et Servais (2000) utilisent des données régionales<sup>3</sup>. Jérôme, Jérôme-Speziari et Lewis-Beck (2003) construisent un modèle pour expliquer le vote obtenu par la majorité parlementaire sortante (au sens large) au premier tour des élections présidentielles (1974-1995) à l'aide de variables explicatives régionales et nationales. Malheureusement, leur prévision *ex ante* pour l'élection présidentielle de 2002 est très inexacte : leur modèle prévoyait un total de 52,36 % des voix pour les partis de gauche au premier tour alors que les partis de gauche ont obtenu au total 42,68 % des voix au premier tour (métropole) en classant à droite les voix (4,33 %) obtenues par Jean Saint-Josse (CPNT). Cette erreur de prévision s'explique en partie par le score très élevé du Front national (extrême droite) au premier tour. En effet, une partie des électeurs du Front national est constituée d'anciens électeurs des partis de gauche. Jérôme et Jérôme-Speziari (2004b) construisent également un modèle pour expliquer le vote obtenu par la majorité parlementaire sortante (au sens large) au premier tour des élections présidentielles (1974-2002) à l'aide de variables explicatives régionales et nationales. Ils expliquent ensuite le vote obtenu au second tour par le candidat de la droite et construisent aussi une fonction de vote pour le Front national (premier tour). Les prévisions *ex post* pour l'élection présidentielle de 2002 sont bonnes mais les auteurs ne présentent pas de prévisions *ex ante* pour 2002. Dubois (2002) construit un modèle expliquant le vote obtenu par le candidat de la majorité présidentielle sortante au second tour des élections présidentielles (1981-1995) à l'aide de variables explicatives départementales et nationales. Pour l'élection présidentielle de 2002, l'auteur présente des prévisions *ex ante* pour le second tour (victoire de Jacques Chirac).

Dans cet article, nous construisons une fonction de vote pour les élections présidentielles françaises permettant d'expliquer le vote pour la gauche au second tour (en cas de duel classique gauche/droite) sur la période 1988-1995 puis sur la période 1988-2007 (sans 2002). Ce modèle, qui utilise des données au plan départemental, permet d'établir des prévisions départementales et nationales *ex post* du vote pour la gauche au second tour sur la période 1988-1995, des prévisions départementales et nationales *ex ante* du vote pour la gauche au second tour pour l'élection présidentielle de 2007 puis des prévisions départementales et nationales *ex post* du vote pour la gauche au second tour, sur la période 1988-2007 (à l'exception de 2002). Nous étudions notamment l'influence de la situation économique départementale sur le vote pour la gauche au second tour des élections présidentielles. Pour les variables politiques, est mise en évidence l'influence positive de la popularité du parti socialiste sur le vote pour la gauche.

Après avoir présenté la fonction de vote (second tour pour la gauche) et les

---

3. Le modèle de Fauvelle-Aymar, Lafay et Servais (2000) est construit pour différents types d'élections dont l'élection présidentielle et il comporte une autre équation permettant d'expliquer le taux de participation électorale simultanément au vote pour la gauche.

variables utilisées, nous analysons les estimations obtenues. Nous faisons état des prévisions *ex post* sur la période 1988-1995 et des prévisions *ex ante* pour l'élection présidentielle de 2007. Enfin, nous présentons une mise à jour des estimations en intégrant l'élection présidentielle de 2007 et de nouvelles prévisions *ex post* (période 1988-2007, à l'exception de 2002).

## 2. Présentation de la fonction de vote

Nous cherchons à construire une fonction expliquant le vote pour la gauche au second tour des élections présidentielles (en cas de duel classique gauche/droite). Nous supposons que les électeurs sont rationnels au sens de Downs (1957). Ils devraient donc avoir un comportement prospectif, mais, en réalité, ils sont mal informés des programmes des candidats et de leurs éventuelles conséquences sur l'économie car l'acquisition d'information est coûteuse par rapport au bénéfice escompté. Ils se comportent donc suivant l'hypothèse de récompense-punition de Key (1966) ou de responsabilité de Paldam (1981), c'est-à-dire qu'ils récompensent (resp. sanctionnent) le gouvernement et les partis de la majorité parlementaire pour de bonnes (resp. mauvaises) performances économiques pendant une période normale et éventuellement pendant une période de cohabitation. Pendant une période de cohabitation, le président sortant s'il se représente, peut être récompensé (resp. sanctionné) pour de bonnes (resp. mauvaises) performances économiques car il apparaît comme étant responsable de la situation économique (en partie responsable ou entièrement responsable suivant le modèle).

Le pourcentage de voix obtenu par la gauche au second tour des élections présidentielles s'explique par la situation économique et des facteurs politiques.

Notre fonction de vote est de la forme :

$$\text{Vote gauche} = f(\text{facteurs économiques, facteurs politiques})$$

Soit la fonction de vote définie de la manière suivante :

$$\text{VOTG2} = f(\text{ECO}) + \alpha_0 + \alpha_1 \text{POPPS} + \alpha_2 \text{LOC} + \omega$$

où VOTG2 est le vecteur  $NT \times 1$  ( $N = 96$  : taille de l'échantillon et  $T$  : nombre d'élections présidentielles) représentant la variable à expliquer : le pourcentage de voix obtenu par la gauche au second tour des élections présidentielles dans chaque département de la France métropolitaine à l'élection présidentielle<sup>4, 5</sup>.

---

4. Les données utilisées sont celles du CIDSP (Grenoble) pour les élections présidentielles de 1988 et 1995 et celles du Conseil constitutionnel pour l'élection présidentielle de 2007.

5. Nous n'avons pas pris en compte l'élection présidentielle de 1981 car, avant cette élection, le taux de croissance sur un an des demandes d'emploi au niveau départemental a beaucoup augmenté dans

Nous commençons la présentation des variables explicatives pour  $T = 2$  (élections présidentielles de 1988 et de 1995).

$f(\text{ECO})$  représente la prise en compte de l'influence de l'économie dans la fonction de vote. En général, les facteurs économiques peuvent dépendre de la situation économique nationale et/ou départementale (passée et/ou future). Dans cette étude,  $f(\text{ECO})$  dépend de la situation économique départementale. Cinq formes de  $f(\text{ECO})$  sont testées représentant cinq hypothèses.

$\alpha_0$  est le vecteur  $NT \times 1$  représentant la constante.

POPSS est le vecteur  $NT \times 1$  représentant la popularité du parti socialiste<sup>6</sup>. C'est une variable qui permet de prendre en compte des facteurs politiques. Nous mesurons la popularité du parti socialiste et non la popularité des partis de gauche car pour les élections présidentielles, à l'exception de 2002 pour la période étudiée, c'est un candidat du parti socialiste qui est au second tour. Nous choisissons de retenir la moyenne du pourcentage de personnes ayant une bonne opinion du parti socialiste dans les trois dernières enquêtes de popularité de la *SOFRES* avant le second tour des élections présidentielles<sup>7</sup>. Nous attendons donc un coefficient positif ( $\alpha_1 > 0$ ) pour la variable POPSS car si la popularité du parti socialiste est élevée (resp. faible), la situation est favorable (resp. défavorable) au candidat de la gauche.

LOC est le vecteur  $NT \times 1$  prenant en compte l'influence du localisme (variable politique), c'est-à-dire l'avantage électoral dont bénéficie le candidat dans le département où il est un élu important ou a été un élu important (cas notamment d'un président sortant qui n'est plus député). Cette variable est égale à 1 dans les départements où le candidat socialiste présent au second tour est ou a été un élu important, à -1 dans les départements où le candidat de droite présent au second tour est un élu important et à 0 pour les autres départements. Sur la période 1988-1995, nous avons donc  $LOC = 1$  pour la Nièvre,  $LOC = -1$  pour la Corrèze et Paris (Jacques Chirac était alors député de la Corrèze et maire de Paris) et  $LOC = 0$  dans les autres départements pour l'élection présidentielle de 1988 ;  $LOC = -1$  pour la Corrèze et Paris car Jacques Chirac était alors député de la Corrèze et maire de Paris<sup>8</sup> et  $LOC = 0$  dans les autres départements pour l'élection présidentielle

---

certain départements (maximum = 45,40 %, moyenne = 16,84 %) et les demandes d'emploi étaient encore assez faibles en 1980 ; le taux de croissance n'a varié qu'entre -15,92 % et 11,75 % suivant les départements avant les élections présidentielles de 1988 (moyenne = -2,86 %) et de 1995 (moyenne = -1,14 %).

6. Lafay, Facchini et Auberger (2007) montrent l'intérêt d'utiliser cette variable pour prévoir le résultat du second tour de l'élection présidentielle de 2007 en cas de duel classique gauche/droite (dans l'article cité, le pourcentage de personnes ayant une bonne opinion du parti socialiste au mois de mars de l'année de l'élection présidentielle).

7. Données disponibles en ligne sur le site internet de la *SOFRES* (<http://www.tns-sofres.com/>) et publiées chaque mois dans le *Figaro magazine* (sauf celui de mai 1988).

8. Dubois (2002) construit une variable localisme mais il prend en compte les origines affectives



Tableau 1 : Variables pour le chômage (1988-1995)

	CHODG	CHODCG	CHODCGA	CHODCGB
1988	-CHOD	-CHOD	CHOD	CHOD
1995	-CHOD	0	-CHOD	0

Tableau 2 : Situation politique (1988-1995)

	Majorité présidentielle	Président sortant se représentant ?	Majorité parlementaire
1988	Gauche	Oui	Droite
1995	Gauche	Non	Droite

de 1995 (Lionel Jospin avait été battu aux élections législatives de 1993 et son implantation électorale en Haute-Garonne était récente et non-comparable à celle des autres candidats présents au second tour). Un coefficient positif ( $\alpha_2 > 0$ ) est attendu pour la variable LOC. Sur la période 1988-1995, on peut difficilement prendre en compte l'influence du localisme car Jacques Chirac était le candidat de la droite au second tour des élections présidentielles de 1988 et de 1995 et, de plus, nous supposons que Lionel Jospin n'a pas bénéficié de cette influence du localisme au second tour de l'élection présidentielle de 1995.

$\omega$  est le vecteur  $NT \times 1$  représentant le terme d'erreurs (ou de perturbations).

Soit CHOD le vecteur  $NT \times 1$  représentant le taux de croissance sur un an des demandes d'emploi au niveau départemental (au mois de février)<sup>9</sup>.

Nous construisons les variables suivantes pour le chômage (tableau 1) et le tableau 2 synthétise la situation politique.

Modèle 1 :  $f(\text{ECO}) = \beta_1 \text{CHODG}$ . La majorité parlementaire est jugée responsable de la situation économique nationale et, par conséquent, de la situation économique départementale.  $\beta_1 < 0$  est attendu car une amélioration (resp. dégradation) de la situation économique départementale est favorable (resp. défavorable) à la droite avant les élections présidentielles de 1988 et de 1995 et défavorable (resp. favorable) à la gauche. Il n'y a pas d'effet présidentiel.

---

(département où la famille est installée) en plus des origines politiques (département électoral de chaque candidat au second tour), ce qui nous paraît moins bien adapté. Il est vrai que la famille de Jacques Chirac est corrézienne mais si ce dernier a bénéficié d'un avantage électoral en Corrèze lors du second tour des élections présidentielles de 1988 (49,14 % seulement mais à l'élection présidentielle de 2007, Nicolas Sarkozy a obtenu 47,01 %) et de 1995 (61,37 %), c'est en grande partie lié à ses mandats de député (élu pour la première fois en 1967) et il a été, également, président du conseil général de la Corrèze de 1970 à 1979.

9. On le calcule avant chaque élection présidentielle en utilisant les données du mois de février qui sont connues pendant la première quinzaine du mois d'avril. Les données proviennent de la DARES – bulletin mensuel des Statistiques du travail (données mensuelles).

Modèle 2 :  $f(\text{ECO}) = \beta_1 \text{CHODG} + \beta_2 \text{CHODCG}$ . Pendant une période de cohabitation, le président sortant, s'il se représente (élection présidentielle de 1988), est jugé en partie responsable de la situation économique car les électeurs ne distinguent pas nettement la responsabilité du président et celle du gouvernement : c'est-à-dire, que le président peut être récompensé (resp. sanctionné) si la situation économique s'améliore (resp. se dégrade). Nous attendons  $\beta_1 < 0$  et  $\beta_2 > 0$  sans que l'on puisse connaître le signe de  $\beta_1 + \beta_2$ . Si  $\beta_1 + \beta_2 > 0$ , nous devrions avoir :  $\beta_1 + \beta_2 < |\beta_1|$  car pendant une période de cohabitation, le président sortant, s'il se représente, ne peut pas être plus récompensé (resp. sanctionné) si la situation économique départementale s'améliore (resp. se dégrade) que la majorité parlementaire pendant une période de cohabitation avec un président sortant qui ne se représente pas (élection présidentielle de 1995). Cette hypothèse est différente de celle de Lewis-Beck (1997) qui suppose que, pendant une période de cohabitation, le président sortant, s'il se représente est pleinement responsable de la situation économique. Elle est également différente de celle de Dubois (2002)<sup>10</sup> qui suppose que la majorité présidentielle sortante (pendant une période normale ou pendant une période de cohabitation) est jugée entièrement responsable de la situation économique.

Modèle 3 :  $f(\text{ECO}) = \beta_1 \text{CHODCGA}$ . Pendant une période de cohabitation, le président sortant, s'il se représente (élection présidentielle de 1988), est jugé responsable de la situation économique car les électeurs mesurent mal la responsabilité du président et celle du gouvernement : le président est récompensé (resp. sanctionné) si la situation économique s'améliore (resp. se dégrade). Nous supposons également que pendant une période normale ou une période de cohabitation, si le président sortant ne se représente pas, la majorité parlementaire est jugée responsable de la situation économique (la droite lors de l'élection présidentielle de 1995). On attend :  $\beta_1 < 0$  car une amélioration (resp. dégradation) de la situation économique départementale est favorable (resp. défavorable) à la gauche avant l'élection présidentielle de 1988 et défavorable (resp. favorable) à la gauche avant l'élection présidentielle de 1995. Cette hypothèse est également différente de celles de Lewis-Beck (1997) et de Dubois (2002).

Modèle 4 :  $f(\text{ECO}) = \beta_1 \text{CHODCGB}$ . Nous supposons que, pendant une période de cohabitation, le président sortant, s'il se représente (élection présidentielle de 1988), est jugé responsable de la situation économique comme dans le modèle 3 et l'on suppose également que pendant une période de cohabitation, si le président sortant ne se représente pas, la situation économique n'a pas d'influence sur le vote (élection présidentielle de 1995). Ce modèle convient pour l'élection présidentielle de 1995 car Jacques Chirac était le candidat de la droite au second tour et il

---

10. Dubois (2002) utilise également le taux de croissance sur un an des demandes d'emploi au niveau départemental.

avait critiqué le bilan du gouvernement d'Édouard Balladur (premier ministre et candidat au premier tour). Dans ce cas, nous attendons  $\beta_1 < 0$  car une amélioration (resp. dégradation) de la situation économique départementale est favorable (resp. défavorable) à la gauche avant l'élection présidentielle de 1988. Cette hypothèse est celle du modèle présidentiel de Lewis-Beck (1997).

Dans ces quatre modèles, les électeurs ont un comportement rétrospectif par rapport à la situation économique départementale.

Modèle 5 :  $f(\text{ECO}) = 0$ . La situation économique n'a pas d'influence sur le vote pour la gauche.

### 3. Estimations et prévisions (1988-1995)

Pour la fonction de vote, notre étude est d'abord réalisée sur la période 1988-1995 (deux élections présidentielles : 1988 et 1995) et sur l'ensemble des départements de la France métropolitaine ( $i$  varie de 1 à 96). Nous avons donc :  $NT = 192$  observations. Le tableau A de l'annexe présente des statistiques descriptives sur les variables utilisées.

Nous commençons par préciser le choix du modèle. Nous avons le choix entre :

- un modèle sans effets :  $\alpha_{0it} = \alpha_0$   
avec  $\omega_{it} = \varepsilon_{it}$  terme d'erreurs aléatoires ;
- un modèle à effets fixes individuels (n'intégrant pas d'effets fixes temporels) :

$$\alpha_{0it} = \gamma_i = \sum_{j=1}^{96} \gamma_j D_j$$

avec  $\omega_{it} = \varepsilon_{it}$  terme d'erreurs aléatoires ;  $D_j$  est une variable muette par département qui prend la valeur 1 pour le département  $j$  et 0, dans le cas contraire (ces variables muettes permettent de tenir compte des caractéristiques personnelles de chaque département, c'est-à-dire des facteurs qui ne sont pas pris en compte par les autres variables explicatives) ;

- un modèle à effets aléatoires individuels ou à erreurs composées (n'intégrant pas d'effets aléatoires temporels) :

$$\alpha_{0it} = \alpha_0 \text{ et } \omega_{it} = u_i + \varepsilon_{it}$$

avec  $\varepsilon_{it}$  terme d'erreurs aléatoires ;  $u_i$  est un élément aléatoire spécifique au département  $i$ .

Dans un modèle à effets fixes, il peut y avoir une corrélation entre les effets individuels ( $\gamma_i$ ) et les variables explicatives du modèle ; en revanche, dans un modèle à effets aléatoires, nous supposons qu'il y a absence de corrélation entre les effets individuels ( $u_i$ ) et les variables explicatives du modèle.

Nous procédons à un test de Fisher afin de choisir entre un modèle à effets fixes (EF) et un modèle sans effets ; pour tous les modèles, les résultats sont favorables

au modèle à effets fixes au seuil statistique de 5 % et, par exemple, pour les modèles (2') à (4') avec effets fixes, nous avons respectivement :  $F(95,93) = 55,18$ ,  $F(95,94) = 54,37$  et  $F(95,94) = 48,29$  (la valeur critique est 1,42). Pour choisir entre un modèle à effets aléatoires (EA) et un modèle sans effets, nous effectuons le test de Breusch et Pagan ; pour tous les modèles, les résultats sont favorables au modèle à effets aléatoires au seuil statistique de 5 % et, par exemple, pour les modèles (2') à (4') avec effets aléatoires, on a respectivement :  $LM = 87,89$ ,  $LM = 87,78$  et  $LM = 86,84$  (la valeur critique est :  $\chi^2_{0,05}(1) = 3,84$ ). Pour choisir entre un modèle à effets fixes (EF) et un modèle à effets aléatoires (EA), nous effectuons le test d'Hausman (1978)<sup>11</sup> et pour tous les modèles, choisissons le modèle à effets aléatoires au seuil statistique de 5 % (il apparaît que l'hypothèse d'absence de corrélation entre les effets aléatoires et les variables explicatives est acceptée au seuil de 5 %) ; par exemple, pour les estimations (2') à (4'), nous avons respectivement :  $H = 1,45$  (la valeur critique est :  $\chi^2_{0,05}(3) = 7,81$ ),  $H = 2,33$  et  $H = 2$  (la valeur critique est :  $\chi^2_{0,05}(2) = 5,99$ ). Le choix du modèle à effets aléatoires est confirmé par le critère d'information de Schwartz ou bayésien (BIC) que Hsiao et Sun (2000)<sup>12</sup> recommandent d'utiliser ; par exemple, pour l'estimation (2') avec des effets aléatoires,  $\ln BIC = -0,77$  alors qu'avec des effets fixes,  $\ln BIC = 1,84$ .

Pour le modèle économétrique, nous choisissons donc un modèle à effets aléatoires (calcul des effets aléatoires par la méthode de Swamy-Arora).

Le tableau B de l'annexe présente les coefficients de corrélation simples entre les variables explicatives : comme le  $R^2$  des cinq estimations (1') à (5') est supérieur aux carrés des coefficients de corrélation simples, d'après le test de Klein (voir Bourbonnais, 2002, p. 108-109), il n'y a pas de présomption de multicolinéarité. Le tableau C de l'annexe présente une détection de la multicolinéarité : les  $R^2$  obtenus à partir des régressions auxiliaires sont inférieurs au  $R^2$  de chaque estimation (1') à (5') donc d'après Gujarati (2004, p. 365) et la règle empirique de Klein, « la multicolinéarité n'est pas un problème fâcheux ».

Pour expliquer le vote obtenu par la gauche au second tour des élections présidentielles (1988-1995), nous estimons les modèles à effets aléatoires suivants :

$$VOTEG2_{it} = \alpha_0 + \beta_1 CHODG_{it} + \alpha_1 POPPS_t + \alpha_2 LOC_{it} + u_i + \varepsilon_{it} \quad (1)$$

$$VOTEG2_{it} = \alpha_0 + \beta_1 CHODG_{it} + \beta_2 CHODCG_{it} + \alpha_1 POPPS_t + \alpha_2 LOC_{it} + u_i + \varepsilon_{it} \quad (2)$$

$$VOTEG2_{it} = \alpha_0 + \beta_1 CHODGA_{it} + \alpha_1 POPPS_t + \alpha_2 LOC_{it} + u_i + \varepsilon_{it} \quad (3)$$

$$VOTEG2_{it} = \alpha_0 + \beta_1 CHODGB_{it} + \alpha_1 POPPS_t + \alpha_2 LOC_{it} + u_i + \varepsilon_{it} \quad (4)$$

$$VOTEG2_{it} = \alpha_0 + \alpha_1 POPPS_t + \alpha_2 LOC_{it} + u_i + \varepsilon_{it} \quad (5)$$

11. Une autre possibilité consisterait à privilégier le modèle à effets fixes car notre échantillon est formé par les 96 départements de la France métropolitaine. C'est ce que semble suggérer Hsiao (2003, p. 43).

12. Cet article est cité dans Baltagi (2003, p. 27).

Pour chaque modèle, l'hypothèse  $H_0 : \alpha_2 = 0$  est testée.

Nous nous intéressons également aux modèles contraints :

$$\text{VOTEG2}_{it} = \alpha_0 + \beta_1 \text{CHODG}_{it} + \alpha_1 \text{POPPTS}_t + u_i + \varepsilon_{it} \quad (1')$$

$$\text{VOTEG2}_{it} = \alpha_0 + \beta_1 \text{CHODG}_{it} + \beta_2 \text{CHODCG}_{it} + \alpha_1 \text{POPPTS}_t + u_i + \varepsilon_{it} \quad (2')$$

$$\text{VOTEG2}_{it} = \alpha_0 + \beta_1 \text{CHODGA}_{it} + \alpha_1 \text{POPPTS}_t + u_i + \varepsilon_{it} \quad (3')$$

$$\text{VOTEG2}_{it} = \alpha_0 + \beta_1 \text{CHODGB}_{it} + \alpha_1 \text{POPPTS}_t + u_i + \varepsilon_{it} \quad (4')$$

$$\text{VOTEG2}_{it} = \alpha_0 + \alpha_1 \text{POPPTS}_t + u_i + \varepsilon_{it} \quad (5')$$

Pour chaque estimation, nous testons l'hypothèse d'homoscédasticité et l'hypothèse d'absence d'autocorrélation des erreurs à l'ordre 1. Pour faire ces tests, nous appliquons les MCO aux données transformées (estimateur MCQG) comme indiqué dans Wooldridge (2002, p. 286-287) et on utilise les résidus obtenus (Wooldridge, 2002, p. 288). Pour les estimations (1') à (5'), le test d'homoscédasticité de Breusch-Pagan<sup>13</sup> montre que l'on peut accepter l'hypothèse nulle d'homoscédasticité au seuil statistique de 5 % ; par exemple, pour les estimations (2') à (4'), nous avons  $\text{NR}^2 = 1,87 < \chi^2_{0,05}(3) = 7,82$  ;  $\text{NR}^2 = 2,95 < \chi^2_{0,05}(2) = 5,99$  et  $\text{NR}^2 = 0,61 < \chi^2_{0,05}(2) = 5,99$ . Pour chaque estimation (1') à (5'), nous acceptons également l'hypothèse d'absence d'autocorrélation des erreurs à l'ordre 1 au seuil statistique de 5 %.

Le coefficient de la variable LOC est positif dans les estimations (1) à (5) mais n'est pas significativement différent de 0 au seuil statistique de 10 %. Pour tester la significativité de la variable LOC, nous effectuons le test du rapport de vraisemblance et observons qu'on peut accepter pour tous les modèles (1) à (5) l'hypothèse  $H_0 : \alpha_2 = 0$  au seuil statistique de 5% ; pour les estimations (1) et (1'), (2) et (2'), (3) et (3'), (4) et (4'), (5) et (5'), nous avons respectivement :  $\text{RV} = -2,48$  ;  $\text{RV} = -2,27$  ;  $\text{RV} = -2,39$  ;  $\text{RV} = -2,05$  et  $\text{RV} = -2,28$  (la valeur critique est :  $\chi^2_{0,05}(1) = 3,84$ ). Avec un modèle à effets aléatoires, les  $R^2$  et  $R^2_c$  des estimations (1) à (5) sont un peu moins élevés que ceux des estimations (1') à (5')<sup>14</sup>.

13. Version de Koenker (1981) présentée en français dans Araujo, Brun et Combes (2004, p. 91-92).

14. Avec un modèle à effets fixes, le  $R^2$  des estimations des modèles (2) à (5) est un peu plus élevé que celui des estimations (2') à (5') mais nous pouvons accepter l'hypothèse  $H_0 : \alpha_2 = 0$  au seuil statistique de 5 %.

Tableau 3 : *Estimations du vote 1988-1995  
(gauche au second tour)*

Variables explicatives	(1') EA	(2') EA	(3') EA	(4') EA	(5') EA
Constante	23,29 (29,27)***	24,73 (29,63)***	24,94 (29,98)***	24,28 (28,12)***	23,50 (30,07)***
CHODG <sub>it</sub>	-0,04 (-1,31)	-0,15 (-3,79)***			
CHODCG <sub>it</sub>		0,19 (4,16)***			
CHODCGA <sub>it</sub>			-0,09 (-3,99)***		
CHODCGB <sub>it</sub>				-0,08 (-2,05)**	
POPPS <sub>t</sub>	0,53 (42,90)***	0,50 (37,47)***	0,49 (37,87)***	0,50 (35,69)***	0,52 (45,08)***
R <sup>2</sup>	0,99	0,99	0,99	0,99	0,99
R <sup>2</sup> <sub>c</sub>	0,99	0,99	0,99	0,99	0,99
ETR	0,72	0,66	0,67	0,71	0,72
JB	7,74	10,11	22,59	29,78	15,74
LM	88,02	87,89	87,78	86,84	88,08
H	0,14	1,45	2,33	2,00	0,00
EMPN	0,02	0,05	0,05	0,03	0,02
EMPD	0,56	0,49	0,48	0,52	0,55

Nous avons pour chaque estimation le coefficient de la variable explicative et en-dessous le t de Student : \*\*\* significatif au seuil statistique de 1 %, \*\* significatif au seuil statistique de 5 %, \* significatif au seuil statistique de 10 %.

Les R<sup>2</sup> et R<sup>2</sup><sub>c</sub> des estimations (1') à (5') sont donnés avec les effets aléatoires. Sans les effets aléatoires, le R<sup>2</sup> est respectivement égal à 0,92 ; 0,93 ; 0,93 ; 0,92 ; 0,91 et le R<sup>2</sup><sub>c</sub> à 0,91 ; 0,93 ; 0,93 ; 0,92 ; 0,91. ETR : écart-type résiduel avec les effets aléatoires. Sans les effets aléatoires, ETR est respectivement égal à 1,01 ; 0,93 ; 0,94 ; 1,00 ; 1,01.

JB : test de Jarque-Bera

LM : test de Breusch et Pagan (multiplicateur de Lagrange) ; H : test d'Hausman

EMPN : erreur moyenne de prévision nationale (valeur absolue)

EMPD : erreur moyenne de prévision départementale (valeur absolue)

Nous pouvons rejeter l'hypothèse de normalité des résidus (test de Jarque-Bera) au seuil statistique de 5 % pour les estimations (1') à (5') (la valeur critique est :  $\chi^2_{0,05}(2) = 5,99$ ) donc les résultats sont à interpréter avec précaution<sup>15</sup>.

On obtient les estimations (tableau 3)<sup>16</sup>.

L'estimation (1') montre que l'hypothèse de responsabilité de la majorité parlementaire sortante par rapport à la situation économique ne convient pas pleinement, car même si le coefficient de la variable CHODG est négatif (signe attendu), il n'est

15. Si on utilise un modèle à effets fixes, on ne peut pas rejeter l'hypothèse de normalité des résidus (test de Jarque-Bera) au seuil statistique de 5 % pour les estimations des modèles (1') et (2').

16. Les 96 valeurs des effets aléatoires par département ( $u_{it}$ ) ne sont pas reproduites ici.

Tableau 4 : *Prévisions nationales ex post en pourcentage de voix 1988-1995 (gauche au second tour)*

Élection	Valeur prédite (2')	Valeur prédite (3')	Valeur prédite (4')	Valeur réalisée	Erreur (2') (3') (4')
1988	53,92	53,92	53,94	53,98	0,06 / 0,06 / 0,04
1995	47,37	47,37	47,35	47,32	0,05 / 0,05 / 0,03

pas significativement différent de 0 au seuil statistique de 10 %. Les différentes estimations ont des indicateurs statistiques proches mais ceux des estimations (2') et (3') sont un peu meilleurs.

Selon l'estimation (2'), les coefficients des variables CHODG et CHODCG indiquent que, pendant une période de cohabitation avec un président sortant de gauche, candidat à sa réélection (élection présidentielle de 1988), si le taux de croissance des demandes d'emploi dans un département a baissé (resp. augmenté) de 5 points lors de l'élection présidentielle, alors le vote départemental en faveur de la gauche a augmenté (resp. baissé) de :  $(-0,15 + 0,19) \times 5 = 0,20$  point. D'après les estimations (3') et (4'), cette augmentation (resp. baisse) est de 0,45/0,40 point. L'estimation (2') montre également que, pendant une période de cohabitation avec un président sortant de gauche qui n'est pas candidat à sa réélection (élection présidentielle de 1995), si le taux de croissance des demandes d'emploi dans un département a baissé (resp. augmenté) de 5 points lors de l'élection présidentielle, le vote départemental pour la gauche diminue (resp. augmente) de 0,75 point. D'après l'estimation (3'), cette baisse (resp. augmentation) est de 0,45 point.

Selon les estimations (2') à (5'), le coefficient de la variable POPPS montre que la popularité du parti socialiste joue un rôle important et une augmentation (resp. une baisse) de 5 points de son indice de popularité entraîne une augmentation (resp. une baisse) du vote pour la gauche égale à environ 2,5 points.

Le tableau 4 ci-après donne, pour les élections présidentielles de 1988 et 1995, les prévisions nationales *ex post* en pourcentage de voix pour la gauche au second tour obtenues avec les estimations (2'), (3') et (4')<sup>17</sup>.

Les erreurs sont très faibles. L'erreur moyenne sur les deux élections est respectivement égale à 0,05/0,05/0,03 point avec chaque estimation. Les prévisions nationales *ex post* en pourcentage de voix au second tour pour la gauche sont donc très satisfaisantes<sup>18</sup>.

17. Les prévisions nationales *ex post* obtenues avec des estimations de modèles à effets fixes sont presque toutes identiques.

18. Selon les estimations (2') à (5'), Lionel Jospin aurait obtenu 50,72 %, 49,77 %, 49,86 % ou 50,43 % en France métropolitaine s'il avait été qualifié pour le deuxième tour de l'élection présidentielle de 2002 (on calcule la variable popularité du parti socialiste en remplaçant le résultat de l'enquête de popularité publiée entre les deux tours par celui de la dernière enquête publiée avant le premier tour).

Tableau 5 : *Variables pour le chômage (2007)*

	CHODG	CHODCG	CHODCGA	CHODCGB
2007	-CHOD	0	CHOD	CHOD

Tableau 6 : *Situation politique (2007)*

	Majorité présidentielle	Président sortant se représentant ?	Majorité parlementaire
2007	Droite	Non	Droite

Au niveau départemental, l'erreur moyenne de prévision avec les estimations (2'), (3') et (4') est respectivement égale à 0,49/0,48/0,52 sur la période 1988-1995.

## 4. Prévisions pour 2007

Les prévisions *ex ante* en voix pour la gauche au second tour de l'élection présidentielle de 2007 sont estimées avec l'hypothèse d'un duel classique gauche/droite (Royal/Sarkozy). Les prévisions *ex ante* sont obtenues à partir des estimations (2'), (3') et (4').

Pour chaque prévision, le taux de croissance est calculé sur un an des demandes d'emploi au niveau départemental (utilisation des données de janvier 2006, avril 2006, juillet 2006, octobre 2006, janvier 2007, février 2007 et de celles du mois de l'année précédente). Les données de février 2007, par exemple, ont été diffusées pendant la première quinzaine du mois d'avril 2007.

Les variables chômage sont définies dans le tableau 5 et la situation politique est synthétisée dans le tableau 6.

La période 2002-2007 est une période sans cohabitation.

Pour la popularité du parti socialiste (variable POPPS), les données utilisées ci-après proviennent de la SOFRES (POPSS = 42,67 % en avril 2006, POPPS = 46 % en juillet 2006, POPPS = 47 % en octobre 2006, POPPS = 50,33 % en janvier 2007, POPPS = 47,33 % en avril 2007 et POPPS = 46 % en mai 2007<sup>19</sup>). La popularité du PS a augmenté au cours de l'année 2006 et a baissé au début de l'année 2007 (elle est égale à 49 % en mars 2007, à 44 % en avril 2007 et à 45 % en mai 2007).

---

Pour mémoire, dans la dernière enquête de la SOFRES réalisée avant le premier tour (17-18 avril 2002), Lionel Jospin recueillait 50% des intentions de vote.

19. Moyenne des données des trois derniers mois : mars, avril et mai 2007 calculée pour la prévision de mai 2007. L'enquête de popularité de mai 2007 a été publiée le 4 mai sur le site internet de la SOFRES et le 5 mai dans le *Figaro magazine*.



Tableau 7 : *Prévisions nationales ex ante en pourcentage de voix pour 2007 (gauche au second tour)*

Mois	Valeur prédite (2)	Valeur prédite (3)	Valeur prédite (4)
avril-06	45,16	45,43	45,39
juil-06	46,45	46,83	46,88
oct-06	46,60	47,10	47,20
janv-07	48,33	48,79	48,92
avr-07	46,82	47,30	47,40
mai-07	46,04	46,55	46,64

Ségolène Royal a obtenu 46,58 % au second tour de l'élection présidentielle de 2007 en France métropolitaine.

Le tableau 7 présente les prévisions nationales *ex ante* en pourcentage de voix pour la gauche au second tour de l'élection présidentielle de 2007 obtenues avec les estimations (2'), (3') et (4')<sup>20, 21</sup>.

Le résultat du second tour de l'élection présidentielle de 2007 s'annonce en avril et en mai 2007 moins serré qu'en janvier 2007 car la popularité du parti socialiste a baissé en avril 2007 (44 %) et s'est stabilisée en mai 2007 (45 %). On prévoit bien la victoire de Nicolas Sarkozy en cas de duel Royal/Sarkozy : l'erreur de prévision des estimations (2'), (3') et (4') est respectivement égale à 0,54/0,03/0,06 point.

## 5. Après l'élection présidentielle de 2007

Pour la fonction de vote, notre étude couvre la période 1988-2007 hormis l'élection présidentielle de 2002 (trois élections présidentielles : 1988, 1995 et 2007). Ainsi, NT = 288 observations.

Pour la variable LOC, nous avons : LOC = 1 pour les Deux-Sèvres, LOC = -1 pour les Hauts-de-Seine et LOC = 0 dans les autres départements pour l'élection présidentielle de 2007<sup>22</sup>. L'influence du localisme est bien prise en compte lorsqu'on intègre l'élection présidentielle de 2007.

20. Pour calculer ces prévisions, nous utilisons le nombre d'électeurs inscrits dans chaque département de la France métropolitaine : pour les prévisions d'avril 2006 à avril 2007, nous avons recours au nombre d'inscrits aux élections européennes de 2005 et pour la prévision de mai 2007, au nombre d'inscrits au premier tour de l'élection présidentielle de 2007 (si nous utilisons le nombre de suffrages exprimés au premier tour de l'élection présidentielle de 2007, nous obtiendrions des prévisions quasi identiques). Avec l'estimation (5'), on obtient 47,42 % comme prévision nationale *ex ante* en mai 2007.

21. Les prévisions *ex ante* obtenues avec des modèles à effets fixes sont presque identiques.

22. Nous avons également essayé : LOC = 1 pour les quatre départements de la région Poitou-Charente (Charentes, Charente-Maritime, Deux-Sèvres et Vienne) car depuis les élections régionales de 2004, Ségolène Royal est présidente de cette région mais les indicateurs statistiques des estimations (3), (4) et (4a) sont alors très légèrement moins bons.

Nous commençons par préciser le choix du modèle. Un test de Fisher est effectué pour choisir entre un modèle à effets fixes (EF) et un modèle sans effets ; pour tous les modèles, les résultats sont favorables au modèle à effets fixes au seuil statistique de 5 % et, par exemple, pour les modèles (3) et (4), nous avons respectivement :  $F(95,189) = 11,78$  et  $F(95,189) = 11,73$  (la valeur critique est 1,32). Afin de choisir entre un modèle à effets aléatoires (EA) et un modèle sans effets, nous utilisons le test de Breusch et Pagan ; pour tous les modèles, les résultats sont favorables au modèle à effets aléatoires au seuil statistique de 5 % et, à titre d'exemple, pour les estimations (3) et (4), nous obtenons respectivement  $LM = 174,66$  et  $LM = 174,69$  (la valeur critique est :  $\chi_{0,05}^2(1) = 3,84$ ). Pour choisir entre un modèle à effets fixes (EF) et un modèle à effets aléatoires (EA), on effectue le test d'Hausman et pour tous les modèles, nous choisissons le modèle à effets aléatoires au seuil statistique de 5 % ; par exemple, pour les estimations (3) et (4), cela donne respectivement :  $H = 3,44$  et  $H = 7,74$  (la valeur critique est :  $\chi_{0,05}^2(3) = 7,81$ ). Comme nous retenons pour le modèle (4), le modèle à effets fixes au seuil statistique de 10 % (la valeur critique est :  $\chi_{0,05}^2(3) = 6,25$ ), nous présentons pour ce modèle deux estimations : l'une avec des effets aléatoires (EA) et l'autre avec des effets fixes (EF). Le critère de Schwartz ou bayésien que Hsiao et Sun (2000) recommandent d'utiliser montre qu'il faut choisir le modèle à effets aléatoires ; pour l'estimation (4) avec des effets aléatoires, nous avons :  $\ln BIC = 1,40$  alors que pour l'estimation (4a) avec des effets fixes,  $\ln BIC = 3,24$ .

Le tableau D de l'annexe présente des statistiques descriptives sur les variables utilisées. Le tableau E, en annexe également, présente les coefficients de corrélation simples entre les variables explicatives : le  $R^2$  des cinq estimations étant supérieur aux carrés des coefficients de corrélation simples, d'après le test de Klein (voir Bourbonnais, 2002, p. 108-109), il n'y a pas de présomption de multicolinéarité. Le tableau F de l'annexe présente une détection de la multicolinéarité : les  $R^2$  obtenus à partir des régressions auxiliaires sont inférieurs au  $R^2$  de chaque estimation (1) à (5). Si l'on se réfère à Gujarati (2004, p. 365) et à la règle empirique de Klein, « la multicolinéarité n'est pas un problème fâcheux ».

Pour le modèle économétrique, nous choisissons par conséquent un modèle à effets aléatoires (les effets aléatoires sont calculés par la méthode de Swamy-Arora) mais pour le modèle (4), nous proposons également l'estimation (4a) avec des effets fixes.

Pour expliquer le vote obtenu par la gauche au second tour des élections présidentielles (1988-2007 sans 2002), nous estimons les modèles à effets aléatoires suivants (1) à (5) et le modèle à effets fixes (4a) :

$$VOTEG2_{it} = \alpha_0 + \beta_1 CHODG_{it} + \alpha_1 POPPS_t + \alpha_2 LOC_{it} + u_i + \varepsilon_{it} \quad (1)$$

$$VOTEG2_{it} = \alpha_0 + \beta_1 CHODG_{it} + \beta_2 CHODCG_{it} + \alpha_1 POPPS_t + \alpha_2 LOC_{it} + u_i + \varepsilon_{it} \quad (2)$$

$$VOTEG2_{it} = \alpha_0 + \beta_1 CHODGA_{it} + \alpha_1 POPPS_t + \alpha_2 LOC_{it} + u_i + \varepsilon_{it} \quad (3)$$

$$\text{VOTEG2}_{it} = \alpha_0 + \beta_1 \text{CHODGB}_{it} + \alpha_1 \text{POPPTS}_t + \alpha_2 \text{LOC}_{it} + u_i + \varepsilon_{it} \quad (4)$$

$$\text{VOTEG2}_{it} = \gamma_i + \beta_1 \text{CHODGB}_{it} + \alpha_1 \text{POPPTS}_t + \alpha_2 \text{LOC}_{it} + \varepsilon_{it} \quad (4a)$$

$$\text{VOTEG2}_{it} = \alpha_0 + \alpha_1 \text{POPPTS}_t + \alpha_2 \text{LOC}_{it} + u_i + \varepsilon_{it} \quad (5)$$

Pour chaque estimation (1) à (5) et (4a), le test d'homoscédasticité de Breusch-Pagan<sup>23</sup> montre que l'hypothèse nulle d'homoscédasticité est rejetée au seuil statistique de 5 % ; par exemple, pour les estimations (3), (4), et (4a), nous obtenons :  $\text{NR}^2 = 30,43 > \chi_{0,05}^2(3) = 7,82$  ;  $\text{NR}^2 = 31,20 > \chi_{0,05}^2(3) = 7,82$  et  $\text{NR}^2 = 51,49 > \chi_{0,05}^2(3) = 7,82$ . L'hypothèse d'absence d'autocorrélation des erreurs à l'ordre 1 est, quant à elle, rejetée pour les estimations (1), (2) et (4) et acceptée pour les estimations (3) et (5) au seuil statistique de 5 % mais rejetée pour les estimations (3) et (5) au seuil statistique de 10 %. Pour l'estimation (4a), le DW adapté par Bhargava *et al.* (1982) au modèle à effets fixes est égal à 1,60. L'hypothèse d'absence d'autocorrélation des erreurs à l'ordre 1 est, par conséquent, rejetée au seuil statistique de 5 %.

Pour toutes les estimations, il n'est pas possible de rejeter l'hypothèse de normalité des résidus au seuil statistique de 5 % (la valeur critique est :  $\chi_{0,05}^2(2) = 5,99$ ).

Les estimations suivantes sont obtenues avec correction de l'hétéroscédasticité et de l'autocorrélation des erreurs à l'ordre 1 à l'aide de la méthode de Newey-West pour les estimations (1), (2), (4) et (4a) et avec correction de l'hétéroscédasticité à l'aide de la méthode de White pour les estimations (3) et (5) (tableau 8)<sup>24</sup>.

L'estimation (1) montre que l'hypothèse de responsabilité de la majorité parlementaire sortante par rapport à la situation économique ne convient pas pleinement, car même si le coefficient de la variable CHODG est négatif (signe attendu), il n'est pas significativement différent de 0 au seuil statistique de 10 %. L'estimation (2) montre que, contrairement à la période 1988–1995, on ne peut pas retenir l'hypothèse d'une responsabilité partielle du président sortant pendant une période de cohabitation, s'il se représente, car  $0,20 - 0,08 = 0,12 > 0,08$  et le président sortant dans cette situation ne peut pas être davantage récompensé (resp. sanctionné) si la situation économique s'améliore (resp. se dégrade) que la majorité parlementaire sortante pendant une période de cohabitation si le président sortant ne se représente pas (élection présidentielle de 1995) ou pendant une période sans cohabitation (élection présidentielle de 2007).

On peut retenir les estimations (3) et (4) car leurs indicateurs statistiques sont un peu meilleurs que ceux de l'estimation (5). Dans les estimations (4) et (4a), nous constatons que les coefficients des variables CHODCGB et POPPS sont presque identiques alors que ceux de la variable LOC sont un peu différents.

23. Version de Koenker (1981).

24. Les 96 valeurs des effets aléatoires par département ( $u_i$ ) et les 96 valeurs des effets fixes par département  $\gamma_i$  ne sont pas reproduites ici.

Tableau 8 : *Estimations du vote 1988-2007  
(gauche au second tour)*

Variables explicatives	(1) EA	(2) EA	(3) EA	(4) EA	(4a) EF	(5) EA
Constante	22,91 (22,02)***	25,37 (20,98)***	25,23 (18,02)***	24,87 (21,17)***		22,10 (17,09)***
CHODG <sub>it</sub>	-0,05 (-1,63)	-0,08 (-2,22)**				
CHODCG <sub>it</sub>		0,20 (3,25)***				
CHODCGA <sub>it</sub>			-0,09 (-2,79)***			
CHODCGB <sub>it</sub>				-0,08 (-2,60)***	-0,09 (-2,59)***	
POPPS <sub>t</sub>	0,53 (30,40)***	0,48 (21,44)***	0,49 (18,48)***	0,49 (23,23)***	0,49 (23,04)***	0,55 (24,95)***
LOC <sub>it</sub>	6,03 (3,58)**	6,05 (3,52)***	6,06 (3,56)***	6,22 (3,50)***	6,10 (2,39)**	5,94 (3,83)***
R <sup>2</sup>	0,89	0,90	0,90	0,90	0,89	0,89
R <sup>2</sup> <sub>c</sub>	0,89	0,90	0,90	0,90	0,85	0,89
ETR	2,00	1,97	1,96	1,97	2,36	2,01
JB	1,32	0,80	0,89	1,12	1,32	4,69
LM	172,05	171,27	174,66	174,69		171,48
H	0,36	5,87	3,43	7,69	7,69	0,21
EMPN	0,16	0,10	0,09	0,08	0,08	0,34
EMPD	1,58	1,53	1,54	1,54	1,50	1,59

Pour chaque estimation, nous indiquons le coefficient de la variable explicative et en-dessous le t de Student corrigé éventuellement à l'aide de la méthode de Newey-West (1, 2 et 4a) ou de celle de White (3 et 4) : \*\*\* significatif au seuil statistique de 1 %, \*\* significatif au seuil statistique de 5 %, \* significatif au seuil statistique de 10 %.

Les effets fixes individuels peuvent également s'écrire :  $\alpha_{0it} = \alpha_0 + f_i$  et la constante estimée de l'estimation (4a) est alors égale à 25,01.

Si on corrige le t de Student des coefficients pour les estimations (3) et (5) à l'aide de la méthode de Newey-West, il est respectivement égal à : 21,52 ; -2,87 ; 22,95 ; 3,50 et 18,60 ; 27,84 ; 3,81.

Les R<sup>2</sup> et R<sup>2</sup><sub>c</sub> des estimations (1) à (5) sont donnés avec les effets aléatoires. Sans les effets aléatoires, le R<sup>2</sup> est respectivement égal à 0,66 ; 0,67 ; 0,67 ; 0,67 ; 0,65 et le R<sup>2</sup><sub>c</sub> à 0,65 ; 0,66 ; 0,66 ; 0,66 ; 0,65.

ETR : écart-type résiduel avec les effets aléatoires. Sans les effets aléatoires, ETR est respectivement égal à 2,39 ; 2,37 ; 2,36 ; 2,37 ; 2,41 ;

JB : test de Jarque-Bera ;

LM : test de Breusch et Pagan (multiplicateur de Lagrange) ; H : test d'Hausman ;

EMPN : erreur moyenne de prévision nationale (valeur absolue) ;

EMPD : erreur moyenne de prévision départementale (valeur absolue).

Dans les estimations (3) et (4), tous les coefficients ont le signe attendu et sont significativement différents de 0 au seuil statistique de 5 %. Par rapport aux estimations (3') et (4') faites sur la période 1988-1995, il apparaît que les coefficients des variables CHODCGA, CHODCGB et POPPS sont très stables. Selon les estimations (3) et (4), les coefficients des variables CHODCGA et CHODCGB

Tableau 9 : *Prévisions nationales ex post en pourcentage de voix 1988-2007 (gauche au second tour)*

Élection	Valeur prédite (3)	Valeur prédite (4)	Valeur prédite (4a)	Valeur réalisée	Erreur (3) (4) (4a)
1988	53,88	53,89	53,90	53,98	0,10 / 0,09 / 0,08
1995	47,36	47,37	47,40	47,32	0,04 / 0,05 / 0,08
2007	46,70	46,69	46,66	46,58	0,12 / 0,11 / 0,08

indiquent que, pendant une période de cohabitation avec un président sortant de gauche candidat à sa réélection (élection présidentielle de 1988), si le taux de croissance sur un an des demandes d'emploi dans un département a baissé (resp. augmenté) de 5 points, alors le vote départemental pour la gauche a augmenté (resp. baissé) de 0,45/0,40 point ; pendant une période hors cohabitation (élection présidentielle de 2007 avec la gauche représentant l'opposition), si le taux de croissance sur un an des demandes d'emploi dans un département a baissé de 5 points, le vote départemental pour la gauche a baissé de 0,45/0,40 point<sup>25</sup>. D'après l'estimation (3), pendant une période de cohabitation avec un président sortant de gauche qui n'est pas candidat à sa réélection (élection présidentielle de 1995), si le taux de croissance sur un an des demandes d'emploi dans un département a baissé (resp. augmenté) de 5 points, alors le vote départemental pour la gauche a baissé (resp. augmenté) de 0,45 point. D'après les estimations (3) et (4), une augmentation (resp. une baisse) de 5 points de l'indice de popularité du parti socialiste entraîne une augmentation (resp. une baisse) du vote pour la gauche égale à 2,45 points. Les estimations (3) et (4) montrent que l'influence du localisme (département dans lequel un candidat du second tour est un élu important) entraîne un gain électoral ou une perte électorale pour la gauche de 6,06/6,22 points.

Le tableau 9 donne, pour les élections présidentielles de 1988, 1995 et 2007, les prévisions nationales *ex post* en pourcentage de voix pour la gauche au second tour obtenues avec les estimations (3), (4) et (4a)<sup>26</sup>.

Les erreurs sont faibles. L'erreur moyenne sur les trois élections est respectivement égale à 0,09/0,08/0,08 point. Les prévisions nationales *ex post* en pourcentage de voix au second tour pour la gauche sont donc très satisfaisantes.

Au niveau départemental, l'erreur moyenne de prévision avec les trois estimations (3), (4) et (4a) est respectivement égale à 1,54/1,54/1,50 sur la période

25. Avant le second tour de l'élection présidentielle de 2007, le taux de croissance sur un an des demandes d'emploi a varié entre -15,49 % et -3,33 % suivant les départements (données de février 2007 et 2006).

26. Les prévisions nationales *ex post* obtenues avec l'estimation d'un modèle (3) avec des effets fixes sont presque identiques.

1988-2007 (sans 2002) et respectivement égale à 1,20/1,22/1,19 ; 1,21/1,21/1,20 et 2,20/2,19/2,11 pour chaque élection présidentielle de 1988, 1995 et 2007.

## 6. Conclusion

Le rôle joué par la popularité du parti socialiste sur le vote obtenu par la gauche au second tour des élections présidentielles est, dans cet article, mis en évidence au niveau départemental et au niveau national (1988-1995 et 1988-2007 sans 2002). La situation économique (taux de croissance sur un an des demandes d'emploi, au niveau départemental, calculé avant chaque élection présidentielle) a également une influence sur le vote pour la gauche. Nous en concluons que, sur les périodes 1988-1995 et 1988-2007 sans 2002, pendant une période de cohabitation, si le président sortant est candidat à sa réélection (élection présidentielle de 1988), il (et donc la gauche en 1988) est récompensé (resp. sanctionné) si la situation économique départementale s'améliore (resp. se dégrade) ; en l'absence de cohabitation (élection présidentielle de 2007), la majorité parlementaire est récompensée et la gauche est sanctionnée car la situation économique départementale s'est améliorée dans chaque département. Pendant une période de cohabitation, si le président sortant ne se représente pas, deux modèles produisent des résultats voisins : l'un suppose que la majorité parlementaire est récompensée (resp. sanctionnée) et, par conséquent, que la gauche est sanctionnée (resp. récompensée) si la situation économique départementale s'améliore (resp. se dégrade) et l'autre suppose que la situation économique départementale n'a alors pas d'influence sur le vote. Les estimations sur la période 1988-2007 (sans 2002) permettent de bien intégrer l'influence du localisme.

Pour l'élection présidentielle de 2007, nous prévoyons que la gauche (Ségolène Royal) obtient *ex ante* 46,04 % à 46,64 % (suivant le choix du modèle) au second tour en France métropolitaine dans l'hypothèse d'un duel Royal/Sarkozy.

Les trois estimations retenues donnent des résultats très proches sur la période 1988-2007 (hormis 2002) et permettront de présenter des prévisions avant l'élection présidentielle de 2012 en cas de duel classique gauche/droite au second tour.

Le vote en faveur du Front national a été notamment étudié par Jérôme et Jérôme-Speziari (2004a, 2004b) et Auberge (2008). Des recherches futures pourront s'intéresser à l'étude du vote obtenu par les candidats des principaux partis au premier tour des élections présidentielles (*PS, UMP, UDF* ou *MODEM*).

## Références

- Araujo, C., Brun J. F. et J.L. Combes. 2004. *Économétrie*. Bréal, coll. Amphi économie.
- Auberger, A. 2001. *Popularité, cycles et politique économique*. Thèse de doctorat en sciences économiques, université Paris 2.
- Auberger, A. 2004. « Les Fonctions de vote : un survol de la littérature », *L'actualité Économique – revue d'analyse économique*, 80, 95-107.
- Auberger, A. 2008. « The National Front Vote and Turnout in the French National Elections », *French Politics*, 6, 94-100 (section Data, Methods and Measures).
- Auberger, A. et E. Dubois. 2003. « Situation politico-économique et résultats des élections législatives françaises », *Revue Économique*, 54, 551-560.
- Auberger, A. et E. Dubois. 2005. “The Influence of Local and National Economic Conditions on French Legislative Elections”, *Public Choice*, 125, 363-383.
- Baltagi, B. H. 2003. *Econometric Analysis of Panel Data*, 2<sup>nd</sup> edition. John Wiley and Sons, Ltd.
- Bélangier, E., Fauvelle-Aymar, C. et M.S. Lewis-Beck. 2007. « Iowa couronne Royal », *Libération*, 28 février.
- Bhargava, A., Franzini, L. et W. Narendranathan. 1982. “Serial Correlation and Fixed Effects Model”, *Review of Economic Studies*, 49, 533-549.
- Bourbonnais, R. 2002. *Économétrie, manuel et exercices corrigés*. 4<sup>e</sup> édition. Paris : Dunod.
- Courbis, R. 1995. « La Conjoncture économique, la popularité politique et les perspectives électorales dans la France d'aujourd'hui », *Journal de la société de statistique de Paris*, 136 (1), 47-70.
- Downs, A. 1957. *An Economic Theory of Democracy*. New York : Harper and Brothers Publishers.
- Dubois, É. 2002. « Un modèle de prévision par département pour les élections présidentielles françaises », *Mimeo LAEP*, université Paris 1.
- Dubois, É. 2007. « Les Déterminants économiques du vote. 1976-2006 : trente ans de fonctions de vote en France », *Revue d'Économie Politique*, 117, p. 243-270.
- Fair, R.C. 1978. “The Effects of Economic Events on Votes for Presidents”, *Review of Economics and Statistics*, 60, p. 159-172.

Fauvelle-Aymar, C., Lafay J. D. et M. Servais. 2000. "The Impact of Turnout on Electoral Choices: An Econometric Analysis of the French Case", *Electoral Studies*, 19, 393-412.

Gujarati, D. N. 2004. *Économétrie*. Bruxelles : De Boeck Université, traduction de la 4<sup>e</sup> édition de *Basic Econometrics*.

Hausman, J. A. 1978. "Specification Tests in Econometrics", *Econometrica*, 46 (6), 1251-1271.

Hsiao, C. 2003. *Analysis of Panel Data*. 2<sup>nd</sup> edition. Cambridge University Press.

Hsiao, C. et B. Sun. 2000. "To Pool or Not To Pool Panel Data", chapitre 9 in Krishnakumar J. et E. Ronchetti, eds., *Panel Data Econometrics: Future Directions*, North-Holland, 181-198.

Jérôme, B. et V. Jérôme-Speziari. 2004a. "A Le Pen Vote Function for the 2002 Presidential Election: A Way to Reduce Uncertainty", *French Politics*, 1, 247-251.

Jérôme, B. et V. Jérôme-Speziari. 2004b. "Forecasting the 2002 Presidential Elections : Lessons from a Political Economy Model", in Lewis-Beck M. S. (ed.), *The French Voter. Before and After the 2002 Elections*, Palgrave Macmillan, 178-204.

Jérôme, B., Jérôme-Speziari V. et M.S. Lewis-Beck. 1999. "Polls Fails in France: Forecasts of the 1997 Legislative election", *International Journal of Forecasting*, 15, 163-174.

Jérôme, B., Jérôme-Speziari V. et M.S. Lewis-Beck. 2003. "Reordering the French Election Calendar : Forecasting the Consequences for 2002", *European Journal of Political Research*, 42, 425-440.

Jérôme, B., Lafay, J. D. et M.S. Lewis-Beck. 1993. « Élections législatives de mars 1993 : prévisions politico-économétriques par région », *Mimeo LAEP*, université Paris 1.

Key, V. O. Jr. 1966. *The Responsible Electorate. Rationality in Presidential Voting, 1936-1960*. Cambridge : Harvard University Press.

Koenker, R. 1981. "A Note on Studentizing a Test of Heteroscedasticity", *Journal of Econometrics*, 17, 107-112.

Kramer, G. H. 1971. "Short-term Fluctuations in U.S. Voting Behavior, 1896-1964", *American Political Science Review*, 65, 131-143.

Lafay, J. D., Facchini F. et A. Auberger. 2007. « Modèles politico-économétriques et prévisions électorales pour mai 2007 », *Revue française d'économie*, 21, 145-164.



- Lewis-Beck, M. S. 1991. "French National Elections : Political Economic Forecasts", *European Journal of Political Economy*, 7, 487-496.
- Lewis-Beck, M. S. 1995. « Les Déterminants économiques de la popularité politique et la prévision électorale : un bilan général des résultats économétriques », *Journal de la Société de Statistique de Paris*, 136 (1), 29-45.
- Lewis-Beck, M. S. 1997. "Who's the Chef? Economic Voting under a Dual Executive", *European Journal of Political Research*, 31 (3), 315-325.
- Mueller, D. 2003. *Public Choice III*. Cambridge : Cambridge University Press.
- Paldam, M. 1981. "A Preliminary Survey of the Theories and Findings on Vote and Popularity Functions", *European Journal of Political Research*, 9 (2), 181-199.
- Sevestre, P. 2002. *Économétrie des données de panel*. Dunod, Paris, coll. Éco Sup.
- Stigler, G. J. 1973. "Short-Term Fluctuations in U.S. Voting Behavior, 1896-1964", *American Economic Review*, 63 (2), 131-143.
- Wooldridge, J. M. 2002. *Econometric Analysis of Cross Section and Panel Data*. The MIT Press, Cambridge, Massachusetts.

Tableau A : *Statistiques descriptives des variables (1988-1995)*

Variable	N	Moyenne	Écart-type	Min	Max
VOTEG2	192	50,60	5,92	34,52	64,67
CHODG	192	2,00	3,79	-11,75	15,92
CHODCG	192	1,43	3,06	-8,40	15,92
CHODCGA	192	-0,86	4,20	-15,92	10,20
CHODCGB	192	-1,43	3,06	-15,92	8,40
POPPS	192	52,00	6,33	45,67	58,33

Tableau B : *Corrélations simples entre les variables explicatives (1988-1995)*

	CHODG	CHODCG	CHODCGA	CHODCGB	POPPS	LOC
CHODG	1,00	0,74			0,23	-0,06
CHODCG		1,00			0,47	0,03
CHODCGA			1,00		-0,48	-0,09
CHODCGB				1,00	-0,47	-0,03
POPPS					1,00	0,03
LOC						1,00

Tableau C : *Détection de la multicolinéarité : régressions auxiliaires (1988-1995)*

Variables à expliquer	R <sup>2</sup>
CHODG	0,43
CHODCG	0,42
CHODCGA	0,23
CHODCGB	0,23
POPPS	0,25

Nous indiquons le R<sup>2</sup> le plus élevé (estimation d'un modèle à effets aléatoires et d'un modèle sans effets)

Tableau D : *Statistiques descriptives des variables (1988-2007 sans 2002)*

Variable	N	Moyenne	Écart-type	Min	Max
VOTEG2	288	49,34	6,06	31,92	64,67
CHODG	288	4,67	5,13	-11,75	15,92
CHODCG	288	0,95	2,59	-8,40	15,92
CHODCGA	288	2,77	6,37	-15,92	15,49
CHODCGB	288	2,38	6,15	-15,92	15,49
POPPS	288	50,00	5,89	45,67	58,33

Tableau E : *Corrélations simples entre les variables explicatives (1988-2007 sans 2002)*

	CHODG	CHODCG	CHODCGA	CHODCGB	POPPS	LOC
CHODG	1,00	0,24			-0,23	0,01
CHODCG		1,00			0,52	0,01
CHODCGA			1,00		-0,61	0,00
CHODCGB				1,00	-0,59	0,03
POPPS					1,00	0,00
LOC						1,00

Tableau F : *Détection de la multicolinéarité : régressions auxiliaires (1988-2007) (sans 2002)*

Variables à expliquer	R <sup>2</sup> (1) (2) (3) (4) (5)	R <sup>2</sup> (4a)
CHODG	0,23	
CHODCG	0,41	
CHODCGA	0,37	
CHODCGB	0,35	0,40
POPPS	0,41	0,37
LOC	0,42	0,53

Dans la première colonne, nous indiquons le R<sup>2</sup> le plus élevé (estimation d'un modèle à effets aléatoires et d'un modèle sans effets)

