

Benoît MULKAY

« Analyse statistique des reports de vote
entre les deux tours de l'élection présidentielle de 2017 »

WP MRE 2019.5

Montpellier Recherche en Economie EA 7491 – Faculté d'Économie
Université de Montpellier - MUSE « Montpellier Université d'Excellence »
Contact : alain.marciano@umontpellier.fr

ANALYSE STATISTIQUE DES REPORTS DE VOTE

ENTRE LES DEUX TOURS DE L'ÉLECTION PRÉSIDENTIELLE DE 2017¹

Benoît MULKAY²

RESUME

Dans cet article, une méthode alternative d'estimation des reports de vote entre les deux tours de l'élection présidentielle 2017. Contrairement à la méthode traditionnelle basée sur des sondages à la sortie des urnes, cette méthode est basée sur une analyse statistique des résultats dans 68 260 bureaux de vote.

On estime une matrice de report de vote, incluant l'abstention et le vote blanc et nul entre les deux tours de l'élection présidentielles de 2017. Les résultats présentés confirment qu'une grande partie de l'électorat des candidats éliminés au premier tour, a voté Macron au second tour, sauf pour Dupont-Aignan. Les reports vers Macron sont d'autant plus important que la commune est dans une grande aire urbaine, ou qu'elle a une richesse importante. On peut noter que les consignes de vote de Mélenchon et de Hamon ont été largement suivies par leurs électeurs, alors que 10 % de l'électorat de Fillon s'est reporté vers Le Pen. D'autre part, l'alliance entre Le Pen et Dupont-Aignan a eu pour résultats qu'un peu plus de la moitié de ses électeurs (53 %) ont suivi cette consigne de vote.

Cette contribution montre qu'une analyse statistique de sciences politiques sur la base de données très nombreuses récoltées à un niveau très fin dans les collectivités territoriales.

¹ Cet article a été préparé pour le livre en hommage au Professeur Jacques Aben. Je remercie Julien Malizard pour sa lecture attentive d'une première version de cette contribution. De même, cette analyse provient de nombreuses discussions avec Amaury Barra qui ont inspirés ce travail.

² Université de Montpellier, MRE.
Adresse : MRE, Faculté d'Économie, Avenue Raymond Dugrand, CS 79606, 34 960 Montpellier Cedex 2,
messagerie : benoit.mulkey@umontpellier.fr

L'élection présidentielle en France est organisée selon un scrutin uninominal majoritaire à deux tours. En 2017, il y avait 11 candidats au premier tour, le 23 avril 2017. Seuls les deux candidats ayant récoltés le plus de vote sont qualifiés pour le deuxième tour qui a eu lieu le 7 mai 2017. Les analystes politiques s'intéressent beaucoup au comportement de report de vote d'un candidat vers un autre entre les deux tours de l'élection. Notamment on peut étudier si les consignes de vote données par les candidats éliminés sont respectées par l'électeur et dans quelle proportion. La plupart des études publiées dans la presse, sont basées sur des sondages (IFOP-FIDUCIAL³, BVA⁴, IPSOS-Sopra-Steria⁵ ou OpinionWay⁶). Ceux-ci sont effectués le soir du second tour, par enquêtes sur Internet (*Computer Assisted Web Interview*) par différents instituts avec des échantillons d'une taille allant de 2 000 à 8 000 électeurs. De plus, ils concernent l'ensemble du territoire. Ainsi on ne peut pas étudier le comportement de report de vote par régions ou par catégories de communes du fait de la faible taille des échantillons.

Dans cet article, nous proposons d'utiliser une méthode alternative pour estimer le report de vote entre les deux tours de l'élection présidentielle basée sur une analyse statistique. Etant organisée au niveau national, on peut disposer des résultats par bureau de vote ou par communes. Cette information est mise à disposition par le Ministère de l'Intérieur⁷ qui organise les élections. Cela donne un échantillon de presque 70 000 bureaux de vote et permet ainsi une estimation statistique très précise de la matrice de report de vote entre les deux tours de l'élection. On pourra ici considérer si les reports de vote sont différents selon la région, le type, la taille ou encore la richesse de la commune. Cependant les reports estimés doivent être considérés comme des comportements moyens des électeurs d'un bureau, et non des comportements individuels. En effet, ils n'indiquent pas des comportements individuels, tels qu'ils sont déclarés dans les enquêtes des instituts de sondage, mais ils montrent les comportements moyens des électeurs au sein des bureaux de vote.

On peut également s'intéresser au comportement d'abstention ou le vote « Blanc ou Nul » qui représente deux alternatives supplémentaires pour les électeurs au-delà des candidats qui se présentent. Ainsi on peut estimer la part des électeurs de chaque candidat éliminé au premier tour qui se sont tournés vers l'abstention ou le vote « blanc ou nul ». De même, on peut calculer la part des abstentionnistes du premier tour qui sont allés voter au deuxième tour.

Dans un premier temps, on présentera la méthodologie statistique utilisée. Elle sera tout d'abord appliquée à l'estimation des reports de vote pour la France entière. On effectuera ensuite une distinction selon le type de commune classée en aires urbaines ou en zone rurale. Enfin, on s'intéressera aux différences provenant de la richesse des communes.

³ IFOP-Fiducial : « Le profil des électeurs et les clefs du second tour de l'élection présidentielle », 7 mai 2017, web : www.ifop.com/publication/le-profil-des-electeurs-et-les-clefs-du-second-tour-de-lelection-presidentielle/

⁴ BVA : « Présidentielle 2017 – 2nd tour : Décrypter le vote des Français », 7 mai 2017,

web : www.bva-group.com/sondages/presidentielle-2017-2nd-tour-decrypter-le-vote-des-francais/

⁵ IPSOS-Sopra-Steria : « 2nd tour – Sociologie des électors et profil des abstentionnistes », 7 mai 2017,

web : www.ipsos.com/fr-fr/2nd-tour-presidentielle-2017-sociologie-des-electors-et-profil-des-abstentionnistes.

⁶ OpinionWay : « Sondage jour du vote Second tour de l'élection présidentielle 2017 », mai 2017,

web : opinionlab.opinion-way.com/dokumenty/OpinionWay-SondageJourduVote-Tour2Presidentielle20177Mai2017.pdf

⁷ Site web gouvernemental : www.data.gouv.fr

1. La méthodologie statistique

L'étude porte sur les bureaux de vote de la métropole et des départements d'outre-mer. Le niveau « bureau de vote » a été choisi parce que nous avons davantage d'observations, mais aussi qu'ils sont plus homogènes que les communes en termes de taille : aucun bureau de vote ne dépasse 5 000 inscrits, ce qui est très différents de l'amplitude de taille des communes par exemple). Nous n'avons pas retenu les bureaux de vote des Français de l'étranger, ni des collectivités d'outre-mer. De même, quelques communes ou quelques bureaux de vote⁸ ont été éliminées du fait de données aberrantes entre les deux tours. Finalement l'échantillon est composé de 68 260 observations (bureaux de vote) répartis dans 35 375 communes, ce qui représente 96,2% des électeurs inscrits pour cette élection⁹.

Le modèle de report de vote est le suivant : soit X_i le vecteur des votes au premier tour incluant les abstentions et les votes « Blancs ou Nuls »¹⁰ (BNL) dans un bureau de vote i , et Y_i le vecteur des votes au deuxième tour dans ce bureau de vote. X_i est de dimension 13×1 (abstention, Blanc ou Nul, et les 11 candidats¹¹), alors que Y_i est de dimension 4×1 : Abstention, Blanc ou Nul, Macron et Le Pen. La matrice de passage B entre les votes du premier et du deuxième tour est de dimension 4×13 et respecte l'équation :

$$Y_i = BX_i + E_i, \quad i = 1, 2, \dots, N = 68\,260.$$

avec E_i : un vecteur 4×1 d'erreurs, non corrélés avec les votes du premier tour X_i . La matrice B est composée de paramètres $\beta_{k,l}$ ($k = 1, 2, \dots, 4$ et $l = 1, 2, \dots, 13$) à estimer qui indiquent la part du vote du candidat l qui se reporte sur le candidat k au second tour. Par définition, ces parts sont bornées entre 0 et 1 : $0 \leq \beta_{k,l} \leq 1$ pour tout k et l . Une transformation logistique est utilisée pour estimer des paramètres non-contraints $\gamma_{k,l}$ sur l'intervalle $[-\infty, +\infty]$ qui sont reliés aux $\beta_{k,l}$ par la relation :

$$\beta_{k,l} = \frac{1}{1 + \exp(-\gamma_{k,l})}$$

Ainsi une estimation de $\gamma_{k,l}$ inférieure à -10 indique que quasiment aucun électeur d'un candidat du premier tour ne s'est reporté sur le candidat au second tour ($\beta_{k,l} \cong 0$). Inversement un paramètre $\gamma_{k,l}$ supérieur -10 indique que quasiment tous les électeurs du candidat l reportent leur vote sur le candidat k au second tour ($\beta_{k,l} \cong 1$).

De plus, il faut que la somme de chaque colonne de la matrice B soit égale à 1, parce que tous les électeurs du candidat l au premier tour doivent être aloués dans un vote au second tour, y compris l'abstention ou le vote « blanc ou nul ». En conséquence dans notre modèle, on aura 13 paramètres contraints :

⁸ Par exemple, les bureaux où le vote a été annulé par le Conseil Constitutionnel du fait d'une irrégularité. Dans ceux-ci, tous les votes sont considérés comme nuls, et aucun des candidats n'a de voix.

⁹ La perte principale provient de l'absence des Français de l'étranger et des Collectivités d'Outre-Mer qui ne sont pas retenues dans cette étude.

¹⁰ Ceux-ci ont été agrégés du fait de la difficulté à faire une différence entre ces deux types de comportement, et que les effectifs sont relativement faibles.

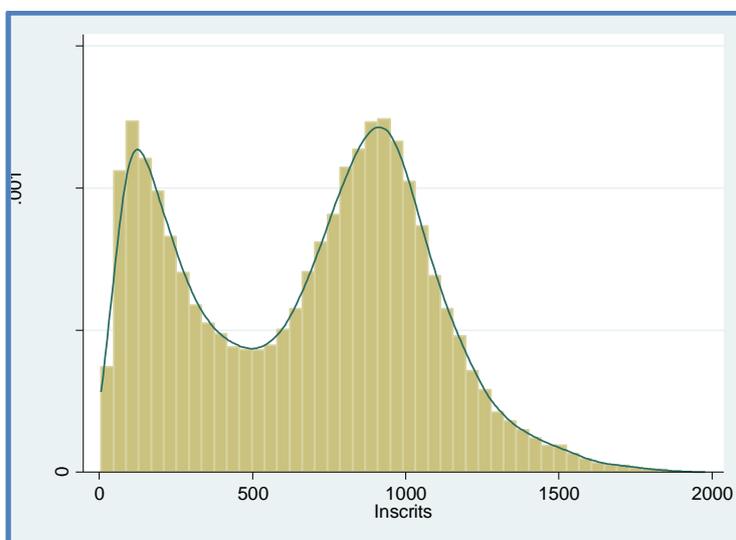
¹¹ Les 11 candidats sont pris dans l'ordre de leur score au premier tour : Macron, Le Pen, Fillon, Mélenchon, Hamon, Dupont-Aignan, Lassalle, Poutou, Asselineau, Arthaud, Cheminade.

$$\beta_{k,1} + \beta_{k,2} + \beta_{k,3} + \beta_{k,4} = 1, \text{ pour tout } k = 1,2, \dots, 13.$$

Ce qui implique pour tout $k = 1,2, \dots, 13$:

$$\frac{1}{1 + \exp(-\gamma_{k,4})} = 1 - \frac{1}{1 + \exp(-\gamma_{k,1})} - \frac{1}{1 + \exp(-\gamma_{k,2})} - \frac{1}{1 + \exp(-\gamma_{k,3})}.$$

Ce modèle est non-linéaire dans les paramètres. Il comporte alors 39 ($= 4 \times 3 - 13$) paramètres à estimer. On utilise la procédure d'estimation NLSUR¹² de Stata qui effectue une estimation par moindres carrés quasi-généralisés non-linéaires d'un système d'équation apparemment indépendante parce que les 4 équations de report de vote sont reliées par les corrélations sur leurs termes d'erreur. De plus, comme le total des choix pour les différentes modalités du second tour est égale aux électeurs inscrits, une des 4 équation est redondante parce qu'elle peut être retrouvée à partir des 3 autres. Les estimations convergent en quelques itérations¹³. Afin de tenir compte de l'importance de chaque bureau de vote, les observations ont été pondérées par le nombre d'inscrits dans le bureau de vote. Le graphique 1 donne l'histogramme des électeurs inscrits dans les bureaux de vote. La distribution est très clairement bimodale du fait qu'au-delà d'un certain seuil d'électeurs dans un bureau de vote, celui-ci est dédoublé.



Graphique 1 : Nombre d'électeurs par bureaux de vote.

¹² Nonlinear Seemingly Unrelated Regression Equation. Voir l'article initial de Arnold Zellner (1962) : "An efficient method of estimating seemingly unrelated regressions and tests for aggregation bias", Journal of the American Statistical Association Vol. 57, p. 348-368.

¹³ L'estimation sur l'ensemble de l'échantillon prend une vingtaine de minutes pour une douzaine d'itérations sur un ordinateur avec un processeur I7 tournant à 4.2 Ghz, avec 16 Go de mémoire.

2. Estimation générale pour la France

Les résultats sont indiqués dans le tableau 1 pour l'ensemble des 68 260 bureaux de vote que l'on considère comme la France entière. Par soucis de clarté, la matrice estimée de report de vote \hat{B} est présentée de manière transposée.

Les résultats se lisent de la manière suivante : pour une ligne correspondant à un des candidats du premier tour, on aura les proportions des électeurs de ce candidat qui se reportent sur les 4 modalités du deuxième tour. Par exemple, pour le candidat « Fillon », il y a en moyenne 14,8% de ses électeurs qui en moyenne se sont abstenus au deuxième tour, 12,5% qui ont voté « Blancs ou Nuls », 10,0% qui ont voté Le Pen et 62,6% qui ont voté pour Macron au second tour. Dans le tableau, les paramètres surlignés correspondent aux paramètres contraints qui sont calculés à partir des 3 autres paramètres de la ligne.

Le modèle très simple (parce qu'il ne fait intervenir que les résultats des candidats au premier tour) est assez bon du fait que les régressions ont des coefficients de détermination supérieurs à 96%, sauf pour le vote « Blanc ou Nul » qui a un R^2 de 76%. Les reports de vote estimés sont très précis : tous les écarts-type estimés sont tous inférieurs au millième, soit 0.1%.

<i>2ème tour</i>				
<i>1er tour</i>	ABSTENTION	BLANCS /NULS	MACRON	LE PEN
ABSTENTION	0.844	0.000	0.093	0.062
BLANCS / NULS	0.000	0.685	0.026	0.290
MACRON	0.000	0.000	1.000	0.000
LE PEN	0.000	0.000	0.000	1.000
FILLON	0.148	0.125	0.626	0.100
MELENCHON	0.265	0.166	0.570	0.000
HAMON	0.000	0.000	1.000	0.000
DUPONT-AIGNAN	0.005	0.466	0.000	0.529
LASSALLE	0.000	0.602	0.070	0.328
POUTOU	0.000	0.710	0.157	0.134
ASSELINEAU	0.569	0.000	0.431	0.000
ARTHAUD	0.000	0.000	0.000	1.000
CHEMINADE	0.000	0.498	0.249	0.253
R^2	0.970	0.759	0.977	0.963
s	18.319	14.506	29.580	19.562

68 260 bureaux de vote, 45 771 789 inscrits.

Estimation NLSUR pondérée par le nombre d'inscrits dans le bureau de vote

Tableau 1 : Estimation des reports de vote – France entière

Environ un abstentionniste du premier tour sur 6 s'est déplacé pour voter au second tour. Sur ces nouveaux votants, environ 60% ont choisi Macron au second tour, contre environ 40% pour Le Pen. 31,6% des électeurs qui avaient voté « Blancs ou Nuls » au premier tour, se sont reportés sur un des deux candidats au second tour avec seulement 2,6% pour Macron et 29,0% pour le Pen !

En moyenne, tous les électeurs de Macron ou de Le Pen au premier tour ont confirmé leur choix au second tour. De même Macron bénéficie de tous les électeurs de Hamon, ce qui correspond à la consigne de vote donné par ce dernier au soir de son élimination au premier tour. Cependant, de manière surprenante, Le Pen bénéficie de l'entière de l'électorat de Arthaud au premier tour. Il faut mentionner que les reports de vote des « petits » candidats du premier tour (Poutou, Asselineau, Arthaud et Cheminade) ne sont pas bien pris en compte par le modèle qui privilégie les grands effectifs de votants.

L'électorat de Fillon et de Mélenchon s'est largement reporté sur Macron au second tour, mais dans une proportion légèrement moindre pour ce dernier : 57,0% au lieu de 62,6% pour Fillon. Alors que leur électorat s'est reporté quasiment dans la même proportion vers le vote « Blanc ou Nul » (12,5% et 16,6%) avec davantage de vote « Blanc ou Nul » dans l'électorat de Mélenchon. La différence s'est fait dans l'abstention : l'électorat de Mélenchon s'y est davantage réfugié (26,5%) que l'électorat de Fillon (14,8%). On peut noter que, suivant les consignes de vote de Mélenchon, aucun de ses électeurs ne s'est reporté sur Le Pen, contrairement à 10,0% de l'électorat de Fillon.

Entre les deux tours, Dupont-Aignan a fait alliance avec Le Pen. Un peu plus de la moitié de ses électeurs (52,9%) ont suivi cette consigne de vote. L'autre partie de ses électeurs ont voté « Blanc ou Nul » (46,6%), alors qu'à peine 0,5% se sont abstenus, et aucun n'a voté en moyenne pour Macron. Les autres reports de vote concernant les 5 candidats ayant obtenu le moins de voix sont très volatiles : ils concernent de petits effectifs d'électeurs. Notons seulement que 32,8% des électeurs de Lassalle et 25,3% des électeurs de Cheminade ont voté Le Pen au deuxième tour. Comme on l'a mentionné ci-dessus, on estime que tous les électeurs de Arthaud se sont reporté sur Le Pen au second tour, mais seulement 15,4% des électeurs de Poutou. Macron reçoit le vote 43,1% des électeurs d'Asselineau, de 24,9% des électeurs de Cheminade et de 15,7% des électeurs de Poutou. Mais la plupart des électeurs de ces candidats (hormis Arthaud) se sont soit abstenus, soit ont voté « Blancs ou Nuls ».

On a effectué une comparaison avec l'enquête de l'Institut BVA¹⁴. Celle-ci a été réalisée par sondage sur Internet au soir du vote du deuxième tour auprès de 2 877 inscrits. Du fait de l'échantillon relativement réduit, les résultats des reports de vote pour les « petits » candidats ne sont pas publiés. Néanmoins cette étude nous semble la plus complète, même si elle omet les reports de vote provenant des abstentionnistes et des électeurs qui ont voté Blanc ou Nul au premier tour. Le Tableau 2 donne la comparaison pour les principaux candidats des reports de vote. L'enquête de BVA qui donne le comportement des abstentionnistes et des électeurs avec un vote Blanc ou Nul a été retravaillée pour donner des informations comparables à celles des estimations ci-dessus.

¹⁴ BVA : « Présidentielle 2017 – 2nd tour : Décrypter le vote des Français », 7 mai 2017, web : www.bva-group.com/sondages/presidentielle-2017-2nd-tour-decrypter-le-vote-des-francais/

Enquête BVA					Estimation NLSUR				
1er tour	2ème tour				1er tour	2ème tour			
	Abstention	Blanc / Nul	Macron	Le Pen		Abstention	Blanc / Nul	Macron	Le Pen
Macron	5%	1%	93%	1%	Macron	0%	0%	100%	0%
Le Pen	9%	2%	3%	86%	Le Pen	0%	0%	0%	100%
Fillon	26%	12%	45%	17%	Fillon	15%	13%	63%	10%
Mélenchon	33%	14%	44%	9%	Mélenchon	26%	17%	57%	0%
Hamon	18%	7%	71%	3%	Hamon	0%	0%	100%	0%
Dupont-Aignan	21%	12%	29%	38%	Dupont-Aignan	1%	47%	0%	53%

Tableau 2 : Comparaison de l'estimation NLSUR avec l'enquête de BVA

Dans l'enquête BVA, on n'a en revanche aucune information sur le comportement au second tour des abstentionnistes et des électeurs avec un vote Blanc ou Nul au premier tour. Cependant on a proposé aux électeurs d'un candidat du premier tour, ces modalités de choix lors du second tour. On remarque que les votes sont plus dispersés au second tour dans l'enquête BVA, que dans nos estimations. Par exemple, l'enquête indique que 93% des électeurs de Macron confirme leur choix au second tour. Mais seulement 86% des électeurs de Le Pen continuent à voter pour elle au second tour. Nos estimations montrent en revanche que la totalité des électeurs de ces deux candidats se reportent sur le même candidat au second tour.

On obtient aussi dans l'estimation des reports de vote plus importants pour Macron : 63% au lieu de 45% dans l'enquête pour l'électorat de Fillon, 57% contre 44% pour l'électorat de Mélenchon, et 100% contre 71% pour l'électorat de Hamon. Cela au détriment de l'abstention et du vote Le Pen qui sont plus importants dans l'enquête BVA. En revanche, l'enquête donne un report de vote de 29% de l'électorat de Dupont-Aignan vers Macron, alors qu'on l'estime nul ; et un report de 38% vers Le Pen, alors qu'il est de 53% dans l'estimation.

En général les résultats des estimations sont plus tranchés que ceux de l'enquête BVA. Néanmoins, cette dernière n'assure pas une cohérence entre les résultats agrégés des deux tours, ce qui est une des caractéristiques de notre estimation. En effet, par construction de la méthode d'estimation, si on applique la matrice de report de vote aux résultats du premier tour, on obtient les résultats du second tour, ce qui n'est pas le cas dans l'enquête BVA, même si on impute les transferts de vote des autres candidats.

3. Estimation par catégorie de communes

Cette analyse concerne l'ensemble des communes de France (excepté les collectivités d'outre-mer ou les Français de l'étranger). Avec la grande taille de l'échantillon, on peut étudier les reports de vote selon quelques stratifications des communes. On a prolongé l'analyse en considérant une première stratification en utilisant le découpage territorial en zones urbaine effectués en 2010 par l'INSEE¹⁵. En effet, beaucoup d'observateurs de l'élection présidentielle de 2017 ont souligné la différence entre les villes et les zones rurales où les habitants ont l'impression d'être moins bien considérés. On a réparti les observations sur les bureaux de vote selon 4 catégories de communes : les villes centres (grands pôles urbains - GPU), la banlieue (ou espace péri-urbain et communes multi-polarisées - EPU), les communes isolées (ISO) ; et les pôles ruraux (PRU) comportant les autres communes. On donne ci-dessous les caractéristiques de ces 4 sous-échantillons :

AIRES URBAINES 2010		Communes		Bureaux de Vote		Inscrits	
		Nombre	Pourcent	Nombre	Pourcent	Nombre	Pourcent
GAU	Grandes aires urbaines	3 231	9.1%	26 580	38.9%	25 224 349	55.1%
EPU	Espace péri-urbain	15 698	44.4%	20 899	30.6%	11 942 170	26.1%
PRU	Pôle ruraux	9 347	26.4%	12 716	18.6%	6 234 820	13.6%
ISO	Communes isolées	7 096	20.1%	8 065	11.8%	2 370 450	5.2%

Tableau 3 : Répartition en aires urbaines

Plus de la moitié de la population considérée ici vit dans une grande aire urbaine, alors qu'un quart de la population se trouve dans un espace périurbain. Le dernier quart se réparti dans des pôles ruraux et des communes isolées qui représentent quasiment la moitié des communes et 30% des bureaux de vote. Les résultats électoraux dans ces 4 catégories de communes sont donnés dans le tableau 4.

Macron, et dans une moindre mesure Mélenchon et Hamon, ont un électorat plus urbain que les autres candidats, ce qui est confirmé pour Macron au second tour où il obtient 10 points de plus dans les grandes aires urbaines, par rapport aux autres catégories de communes. En contrepartie, Le Pen a davantage d'électeurs dans les communes rurales ou isolées. Fillon a un profil électoral moins marqué avec une proportion d'électeurs moins importantes dans les grandes aires urbaines, mais plus fortes dans les communes isolées, alors qu'il a un déficit d'électeurs dans l'espace péri-urbain. On note aussi que le profil des abstentionnistes est identique pour les deux tours, même s'il est légèrement plus élevé au second tour qu'au premier. Le vote Blanc et Nul augmente très fortement au second tour, d'autant plus que la commune est rurale ou isolée. Cela peut provenir de la disparition des « petits » candidats au second tour, alors qu'ils avaient davantage d'électeurs dans ces communes.

¹⁵ Voir par exemple : Chantal Brutel : « Le nouveau zonage en aires urbaines de 2010 », INSEE Première N°1374, 18 octobre 2011.

1er tour					
	GAU	EPU	PRU	ISO	TOTAL
Abstentions	22.9%	17.4%	20.1%	19.6%	20.9%
Blanc ou Nul	2.4%	2.6%	2.9%	3.0%	2.6%
MACRON	25.6%	22.1%	21.4%	20.8%	23.8%
LE PEN	18.2%	25.8%	24.9%	23.8%	21.5%
FILLON	20.1%	18.9%	20.1%	21.5%	19.8%
MELENCHON	21.4%	17.8%	17.8%	17.5%	19.7%
HAMON	7.1%	5.5%	5.7%	5.4%	6.3%
DUPONT-AIGNAN	4.1%	5.8%	5.3%	5.3%	4.8%
AUTRES CANDIDATS	3.6%	4.2%	4.9%	5.7%	4.1%

2ème tour					
	GAU	EPU	PRU	ISO	TOTAL
Abstentions	26.7%	20.9%	22.8%	22.0%	24.4%
Blanc ou Nul	10.8%	12.3%	12.7%	13.5%	11.6%
MACRON	70.7%	59.6%	60.3%	60.7%	65.7%
LE PEN	29.3%	40.4%	39.7%	39.3%	34.3%

68 260 bureaux de vote de la France Métropolitaine et des Départements d'Outre-Mer.

45 771 789 inscrits.

Les pourcentages pour les candidats se rapportent au nombre de votes exprimés.

Tableau 4 : Vote par catégorie de communes

Les reports de vote dans ces 4 catégories de communes sont donnés en détail dans l'Annexe 1. On se concentre ici uniquement sur les reports de vote de Fillon et de Mélenchon entre le premier et le second tour de l'élection présidentielle de 2017. Les autres reports de vote étant très proche de ceux de la France entière, ou trop instables du fait des faibles effectifs des électeurs de certains candidats.

Les reports de vote de Fillon vers Macron s'améliorent fortement lorsque la commune appartient à une Aire Urbaine quelle que soit sa taille, par rapport à une commune isolée. Même dans l'espace péri-urbain ou dans les pôles ruraux, environ 64 % des électeurs de Fillon se reportent vers Macron au deuxième tour. Il n'y a une chute des reports de voix que dans les communes isolées. Dans les zones rurales (pôles ruraux ou communes isolées, le report vers Le Pen augmente légèrement, de même que les reports vers le vote Blanc et Nul. Les abstentions ou les votes blancs et nuls représentent partout de 20% à 30,7% des votes pour Fillon.

Les écarts entre catégories de communes sont plus importants pour les électeurs de Mélenchon au premier tour. En effet, si 61,6% de ses électeurs se reportent vers Macron dans les Grandes Aires Urbaines, ce pourcentage tombe à 38,8% dans les pôles ruraux et 40,2% dans les communes isolées, soit une différence de plus de 20 points. Dans les pôles ruraux, on estime à 16,5% les électeurs de Mélenchon qui votent Le Pen au deuxième tour. Ils sont 8,6% dans les communes

isolées, mais aucun électeur de Mélenchon ne se reportent sur Le Pen au deuxième tour dans les Grandes Aires Urbaines, et à peine 2,6% dans l'espace péri-urbain. En revanche il y a une forte poussée des abstentionnistes au second tour dans l'électorat de Mélenchon, sauf dans les pôles ruraux, où il y a beaucoup plus de vote Blanc ou Nul.

		ABST.	B / N	MACRON	LE PEN
FILLON	GAU	15.5%	11.3%	64.0%	9.2%
	EPU	12.1%	13.4%	63.7%	10.8%
	PRU	0.0%	20.0%	64.2%	15.8%
	ISO	15.7%	15.0%	56.8%	12.5%
		ABST.	B / N	MACRON	LE PEN
MELENCHON	GAU	27.6%	10.8%	61.6%	0.0%
	EPU	21.8%	26.9%	48.7%	2.6%
	PRU	0.0%	44.7%	38.8%	16.5%
	ISO	19.0%	32.2%	40.2%	8.6%

Tableau 5 : Reports des votes de Fillon et de Mélenchon par catégorie de communes

Comme mentionné par de nombreux commentateurs, le vote pour Macron semble ainsi beaucoup plus important dans les Grandes Aires Urbaines, mais aussi les reports de vote des deux autres grands candidats (Fillon et Mélenchon). A l'inverse le report de vote vers Le Pen est légèrement plus important en zone rurale, mais ces reports sont amoindris par un recours plus important soit à l'abstention, soit au vote blanc ou nul.

4. Estimation selon la richesse des communes

Enfin nous avons distingué les communes selon leur richesse en se basant sur le revenu net moyen par foyer fiscal de la commune en 2011¹⁶. Celui-ci est défini comme le revenu net déclaré en moyenne par foyer fiscal. Ce revenu net moyen est attribué à tous les bureaux de vote de la commune¹⁷ parce qu'on ne peut pas disposer de cette information au niveau des découpages infra-communaux des bureaux de vote. La moyenne est de 24 352 €, alors que la médiane est de 22 788 € indiquant une asymétrie vers la droite de la distribution de ces revenus.

¹⁶ Les données proviennent de la source FiLoSoFi 2011 (Dispositif sur les revenus fiscaux et sociaux localisés) de l'INSEE sur la base d'informations provenant de la Direction Générales des Finances Publiques.

¹⁷ Le revenu net moyen ne peut être calculé dans 930 communes de l'échantillon du fait du secret statistique qui empêche de communiquer le revenu net déclaré dans les très petites communes. Les bureaux de vote de ces communes ont été exclus de l'analyse.

Sur cette base, on distingue les communes pauvres avec un revenu net moyen inférieur à 20 000 € (ce qui correspond environ au premier quartile : 19 963 € ; les communes riches où ce revenu net moyen est supérieur à 27 000 € (ce qui correspond environ au troisième quartile : 26 975 €) ; et les communes moyennes entre ces deux bornes qui concernent en moyenne la moitié des bureaux de vote. Le Tableau 6 donne la répartition des communes, des bureaux de vote et des inscrits pour ces 3 classes de communes.

Revenu net moyen	Communes		Bureaux de Vote		Inscrits	
	Nombre	Pourcent	Nombre	Pourcent	Nombre	Pourcent
Inférieur à 22 000 €	7 128	21.7%	17 028	25.3%	9 355 340	20.5%
22 000 € à 27 000 €	16 679	50.7%	33 538	49.8%	22 927 440	50.1%
Supérieur à 27 000 €	9 073	27.6%	16 744	24.9%	13 441 903	29.4%

Tableau 6 : Répartition par revenu net moyen

Des commentateurs ont souligné les différences dans les votes entre les communes pauvres et les communes riches. Le Tableau 7 présente les résultats obtenus par les différents candidats au premier et au second tour dans nos données.

	Revenu net moyen			TOTAL
	Inférieur à 22 000 €	De 22 000 € à 27 000 €	Supérieur à 27 000 €	
Premier tour				
Abstentions	27.0%	20.7%	17.1%	20.9%
Blanc et Nul	3.3%	2.6%	2.1%	2.6%
MACRON	20.2%	22.9%	27.4%	23.8%
LE PEN	25.7%	22.7%	16.9%	21.5%
FILLON	16.2%	18.5%	24.2%	19.8%
MELENCHON	22.3%	20.4%	17.0%	19.7%
HAMON	6.2%	6.4%	6.4%	6.3%
DUPONT-AIGNAN	4.4%	4.9%	4.8%	4.8%
AUTRES CANDIDATS	5.0%	4.2%	3.3%	4.1%
Second tour				
Abstentions	28.8%	24.3%	21.4%	24.4%
Blanc et Nul	12.3%	11.8%	10.8%	11.6%
MACRON	59.7%	64.1%	72.1%	65.7%
LE PEN	40.3%	35.9%	27.9%	34.2%

67 310 bureaux de vote de la France Métropolitaine et des Départements d'Outre-Mer.

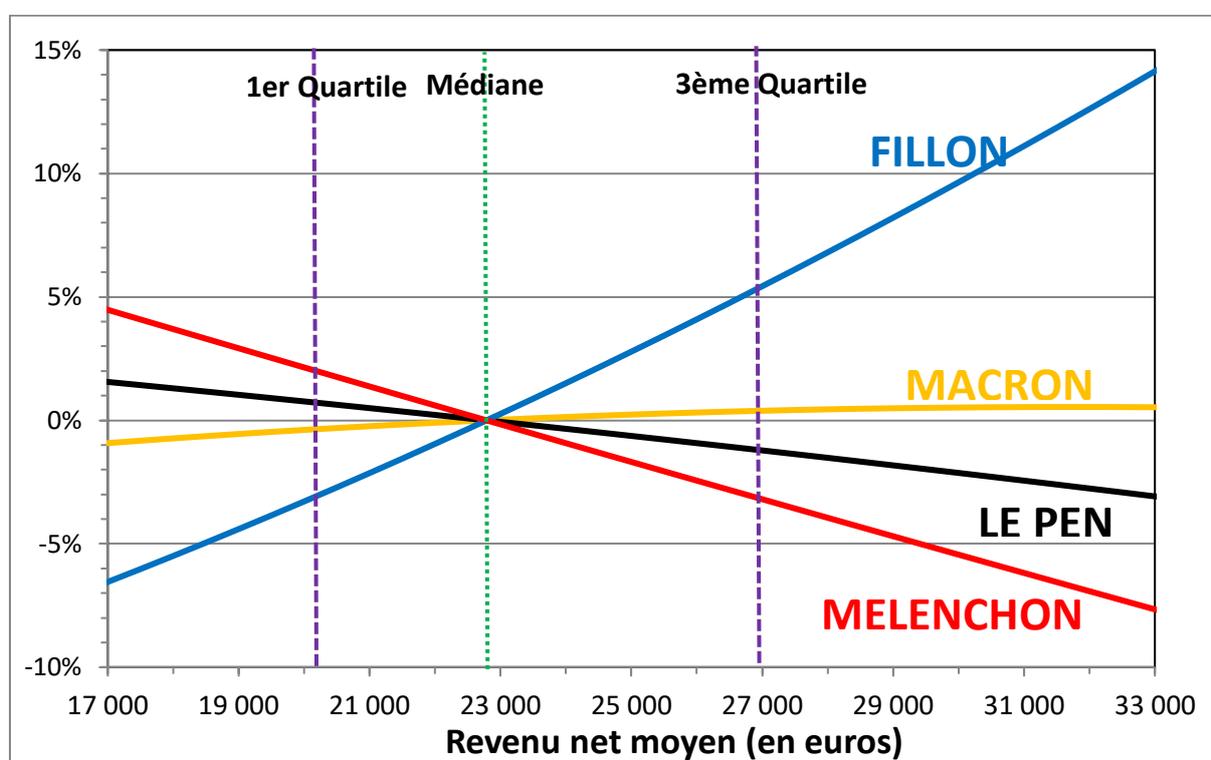
45 724 683 inscrits.

Les pourcentages pour les candidats se rapportent au nombre de vote exprimés.

Tableau 7 : Vote selon le revenu net moyen de la commune

Ces résultats confirment que les électeurs de Macron et Fillon sont davantage présents dans les communes « riches » alors qu'en contrepartie les électors de Le Pen et de Mélenchon se retrouve plus dans les communes « pauvres ». En conséquence, Macron obtient un score supérieur de 12,4 % dans les communes « riches » par rapport aux communes « pauvres », et inversement pour Le Pen. Les votes pour Hamon et pour Dupont-Aignan sont en revanche plus homogènes entre les types de communes. Les abstentionnistes et les votes blancs et nuls sont aussi plus nombreux dans les communes « pauvres », même si l'écart se réduit au deuxième tour.

Afin de confirmer l'effet du revenu net moyen de la commune sur les résultats électoraux, on a effectué une régression¹⁸ du score obtenu par chaque candidat, sur une fonction quadratique du revenu net moyen, en tenant compte de nombreuses variables de contrôle : structure par âge et par sexe de la population, catégories socio-professionnelles, niveau d'éducation, type de famille, type de logement, proportion d'étrangers et d'immigrés, taille de la commune et taille de l'aire urbaine, ...). L'effet du revenu net moyen est indiqué dans le graphique 2 où celui-ci est par construction nul au niveau de la médiane de ce revenu net moyen. Celui-ci est considéré entre le premier décile des observations (environ 17 000 €) et le neuvième décile (environ 33 000 €).



Graphique 2 : Effet du Revenu net moyen sur le vote du premier tour

¹⁸ Cette régression est effectuée par moindres carrés ordinaires sur les 37 710 communes de l'échantillon. Il n'y a aucun gain à effectuer un modèle à équation apparemment indépendantes (SURE) du fait que les variables explicatives sont identiques dans toutes les équations.

Les effets quadratiques, bien que significatifs, sont très peu marqués dans ces courbes où il semble davantage y avoir des effets linéaires du revenu net moyen sur le score de chaque « grand » candidat. L'effet est le plus important pour Fillon qui perd 3,3% au niveau du premier quartile, alors qu'il en gagne 5,4% au niveau du troisième quartile. Au neuvième décile, Fillon gagne même 14,1 % des voix. Le vote Macron est très stable quelle que soit le revenu net moyen de la commune, avec seulement une perte de 0,4 % au premier quartile, contre un gain de 0,4 % au troisième quartile. En revanche l'effet du revenu net moyen est négatif pour le vote Le Pen et encore plus pour le vote Mélenchon. Au premier quartile, Le Pen et Mélenchon aurait respectivement un gain de 0,8% et de 2,1 %, alors qu'il perdrait respectivement 1,2 % et 3,2 % au troisième quartile. Dans les 10% des communes les plus « riches », Mélenchon perdrait ainsi plus de 7,7% des voix, et Le Pen plus de 3,1 % des voix. En conclusion, même en contrôlant pour de nombreux effet, le revenu net moyen a un effet très négatif sur l'électorat de Mélenchon, et un peu moins sur l'électorat de Le Pen. En revanche son effet est largement positif pour Fillon. Il est à noter qu'avec ces variables de contrôle, la richesse de la commune n'influence que très peu le vote pour Macron au premier tour, contrairement au Tableau 7, où on ne prenait pas en compte les effets des co-variables.

Les reports de vote selon le niveau de richesse de la commune sont donnés en détail dans l'Annexe 2. On reprend ici encore, le tableau des reports de vote pour les deux « grands » candidats éliminés au premier tour : Fillon et Mélenchon. Pour ces deux candidats, le report de vote pour Macron augmente avec le revenu net moyen de la commune. Le report de vote de Fillon vers Macron passe de 55,2% dans une commune « pauvre » à 66 % dans une commune « riche », alors que le report de Mélenchon vers macron est encore plus faible dans les communes « pauvres » (47,3 %), mais il augmente très fortement, au-delà de celui de Fillon, dans les communes « riches » (72,7 %), soit une augmentation de 25 points !

	REVENU	ABST.	B / N	MACRON	LE PEN
FILLON	Inférieur à 22 000 €	16.5%	17.3%	55.2%	11.0%
	de 22 000 € à 27 000 €	14.1%	14.0%	60.1%	11.8%
	Supérieur à 27 000 €	13.3%	11.0%	66.0%	9.7%
	ENSEMBLE	15.3%	11.5%	63.9%	9.3%
	REVENU	ABST.	B / N	MACRON	LE PEN
MELENCHON	Inférieur à 22 000 €	34.4%	15.5%	47.3%	2.8%
	de 22 000 € à 27 000 €	23.1%	14.1%	62.8%	0.0%
	Supérieur à 27 000 €	17.9%	9.4%	72.7%	0.0%
	ENSEMBLE	26.6%	13.4%	60.0%	0.0%

Tableau 8 : Reports des votes de Fillon et de Mélenchon selon le revenu net moyen de la commune

En contrepartie, le report de vote vers Le Pen est pratiquement stable autour de 10% de la part de l'électorat de Fillon, même si on constate une légère baisse en fonction de du revenu net

moyen de la commune. On constate le même affaiblissement du report de Fillon vers l'abstention ou le vote blanc et nul lorsque le revenu net moyen de la commune augmente. En revanche, ces reports vers l'abstention ou le vote blanc et nul de la part de l'électorat de Mélenchon se réduit très fortement avec le niveau de revenu net moyen de la commune. Ainsi le report vers l'abstention qui était de 34,4% dans une commune « pauvre » passe à 17,9 % dans une commune « riche ». De même, le report vers le vote blanc et nul se réduit de 15,5 % dans une commune pauvre vers 9,4 % dans une commune « riche ». Cet effet positif du revenu net moyen sur le report de vote semble ici une marque importante du scrutin présidentiel de 2017, au-delà du type et des motivations respectives des électors de ces deux candidats.

5. Conclusions

La méthode d'estimation statistique des reports de vote présentée ici permet d'analyser ces reports en fonction de différentes catégorisations de la commune ou du bureau de vote. Cette méthode est basée sur l'analyse d'un grand nombre de données sur les bureaux de vote, plutôt que sur les réponses d'un échantillon d'individus après le vote. Les résultats sont parfois différents parce que la méthodologie utilisée n'est pas identique. Dans une enquête, on s'intéresse au comportement individuel avec les biais classiques dus aux sondages et à la révélation des votes des individus. D'autant plus que les questions ne sont généralement pas complètes en ce sens qu'elles ne prennent pas souvent en compte les abstentions ou les votes blancs et nuls.

La méthode statistique utilisée ici repose sur l'analyse du comportement moyen des électeurs dans un bureau de vote. On utilise ici un critère des moindres carrés pour déterminer le comportement moyen en forçant les reports à être compris entre 0 % et 100 % avec une transformation logistique et à être complet pour un candidat donné. Ces contraintes impliquent d'estimer un modèle non linéaire. Une alternative pourrait être d'estimer un modèle linéaire, mais avec des contraintes d'inégalités sur les paramètres, ce qui impliquerait d'utiliser des techniques spéciales basées sur la programmation quadratique.

D'autre part, la méthode présentée ici estime les reports de vote inconditionnels quelles que soient les caractéristiques de l'électorat des bureaux de vote. On a présenté ici deux stratifications des bureaux de vote selon un zonage en aires urbaines d'une part, et sur la base du revenu net moyen d'autre part. On pourrait aller plus loin, mais au prix d'une lourdeur technique supplémentaire, en conditionnant les reports de vote selon quelques caractéristiques du bureau de vote ou de la commune : par exemple, la taille de la commune, la répartition par âge ou par sexe, etc. Cependant elle ne peut pas prendre en compte les informations, les comportements ou la campagne électorale entre les deux tours des élections qui viendraient à modifier le comportement de vote ou de report de vote.

Les résultats présentés confirment qu'une grande partie de l'électorat des candidats éliminés au premier tour a voté Macron au second tour, sauf pour Dupont-Aignan. Les reports vers

Macron sont d'autant plus important que la commune est dans une grande aire urbaine, ou qu'elle a une richesse importante.

Avec la méthodologie développée ici, on pourrait croiser les résultats du vote et les reports de vote avec d'autres variables ou des élections nationales antérieures permettant de mieux classer les comportements. Par exemple, il semble qu'au-delà de certaines caractéristiques des communes, l'attitude envers la construction européenne puisse être un acteur influençant les reports de vote pour un candidat pro-européen par rapport à une candidate plutôt opposée à la construction européenne.

ANNEXE 1 : Estimation des Reports de vote par catégories de communes

GAU : Grandes Aires Urbaines

	Abstentions	Blanc ou Nul	MACRON	LE PEN
Abstentions	83.6%	0.0%	9.3%	7.0%
Blanc ou Nul	0.0%	83.6%	0.0%	16.4%
MACRON	0.0%	0.0%	100.0%	0.0%
LE PEN	0.0%	0.0%	0.0%	100.0%
FILLON	15.5%	11.3%	64.0%	9.2%
MELENCHON	27.6%	10.8%	61.6%	0.0%
HAMON	0.0%	0.0%	100.0%	0.0%
DUPONT-AIGNAN	4.9%	46.9%	0.0%	48.3%
LASSALLE	0.0%	79.7%	0.0%	20.3%
POUTOU	0.0%	80.0%	0.0%	20.0%
ASSELINEAU	68.4%	0.0%	31.6%	0.0%
ARTHAUD	0.0%	42.3%	0.0%	57.7%
CHEMINADE	34.1%	44.5%	0.0%	21.4%

PRU : Pôles Ruraux

	Abstentions	Blanc ou Nul	MACRON	LE PEN
Abstentions	61.4%	14.9%	2.2%	21.6%
Blanc ou Nul	0.0%	27.5%	72.5%	0.0%
MACRON	0.0%	0.0%	100.0%	0.0%
LE PEN	0.0%	43.3%	0.0%	56.7%
FILLON	0.0%	20.0%	64.2%	15.8%
MELENCHON	0.0%	44.7%	38.8%	16.5%
HAMON	0.0%	0.0%	100.0%	0.0%
DUPONT-AIGNAN	0.0%	53.1%	0.0%	46.9%
LASSALLE	0.0%	48.2%	32.2%	19.6%
POUTOU	0.0%	0.0%	100.0%	0.0%
ASSELINEAU	0.0%	42.4%	0.0%	57.6%
ARTHAUD	0.0%	100.0%	0.0%	0.0%
CHEMINADE	0.0%	0.0%	100.0%	0.0%

EPU : Espace Péri-Urbain

	Abstentions	Blanc ou Nul	MACRON	LE PEN
Abstentions	86.1%	0.0%	1.8%	12.1%
Blanc ou Nul	0.0%	60.2%	39.7%	0.1%
MACRON	0.0%	0.0%	100.0%	0.0%
LE PEN	0.0%	0.0%	0.0%	100.0%
FILLON	12.1%	13.4%	63.7%	10.8%
MELENCHON	21.8%	26.9%	48.7%	2.6%
HAMON	0.0%	0.0%	100.0%	0.0%
DUPONT-AIGNAN	17.9%	29.8%	9.8%	42.5%
LASSALLE	0.0%	44.7%	15.0%	40.4%
POUTOU	0.0%	29.1%	70.9%	0.0%
ASSELINEAU	62.0%	0.0%	14.5%	23.5%
ARTHAUD	0.0%	31.0%	15.2%	53.7%
CHEMINADE	0.0%	27.0%	67.1%	5.9%

ISO : Communes Isolées

	Abstentions	Blanc ou Nul	MACRON	LE PEN
Abstentions	83.0%	0.0%	8.0%	9.0%
Blanc ou Nul	0.0%	66.1%	22.0%	11.9%
MACRON	0.0%	0.0%	100.0%	0.0%
LE PEN	0.0%	0.0%	0.0%	100.0%
FILLON	15.7%	15.0%	56.8%	12.5%
MELENCHON	19.0%	32.2%	40.2%	8.6%
HAMON	0.0%	0.0%	100.0%	0.0%
DUPONT-AIGNAN	10.5%	20.2%	29.9%	39.4%
LASSALLE	8.2%	32.8%	43.8%	15.2%
POUTOU	19.9%	0.0%	80.1%	0.0%
ASSELINEAU	27.1%	22.0%	0.0%	50.8%
ARTHAUD	0.0%	0.0%	100.0%	0.0%
CHEMINADE	29.7%	30.0%	20.7%	19.5%

ANNEXE 2 : Estimation des reports de vote selon le revenu net moyen de la commune

Revenu net moyen				
<i>inférieur à 22 000 €</i>				
	Abstentions	Blanc ou Nul	MACRON	LE PEN
Abstentions	79.9%	0.0%	15.0%	5.1%
Blank or Null	0.0%	57.8%	14.4%	27.8%
MACRON	0.0%	0.0%	100.0%	0.0%
LE_PEN	0.0%	0.0%	0.0%	100.0%
FILLON	16.5%	17.3%	55.2%	11.0%
MELENCHON	34.4%	15.5%	47.3%	2.8%
HAMON	0.0%	0.0%	100.0%	0.0%
DUPONT-AIGNAN	0.0%	42.6%	0.0%	57.4%
LASSALLE	0.0%	69.8%	0.0%	30.2%
POUTOU	0.0%	72.9%	0.0%	27.1%
ASSELINEAU	34.9%	0.0%	65.1%	0.0%
ARTHAUD	0.0%	0.0%	0.0%	100.0%
CHEMINADE	0.0%	0.0%	100.0%	0.0%

de 22 000 € à 27 000 €				
	Abstentions	Blanc ou Nul	MACRON	LE PEN
Abstentions	89.7%	0.0%	4.9%	5.4%
Blank or Null	0.0%	74.6%	0.0%	25.4%
MACRON	0.0%	0.0%	100.0%	0.0%
LE_PEN	0.0%	0.0%	0.0%	100.0%
FILLON	14.1%	14.0%	60.1%	11.8%
MELENCHON	23.1%	14.1%	62.8%	0.0%
HAMON	0.0%	0.0%	100.0%	0.0%
DUPONT-AIGNAN	0.0%	42.8%	0.0%	57.2%
LASSALLE	0.0%	63.1%	0.0%	36.9%
POUTOU	0.0%	58.2%	41.8%	0.0%
ASSELINEAU	40.5%	0.0%	59.5%	0.0%
ARTHAUD	0.0%	50.3%	0.0%	49.7%
CHEMINADE	0.0%	46.9%	0.0%	53.1%

supérieur à 27 000 €				
	Abstentions	Blanc ou Nul	MACRON	LE PEN
Abstentions	93.3%	0.0%	2.2%	4.5%
Blank or Null	0.0%	80.6%	0.0%	19.4%
MACRON	0.0%	0.0%	100.0%	0.0%
LE_PEN	0.0%	0.0%	0.0%	100.0%
FILLON	13.3%	11.0%	66.0%	9.7%
MELENCHON	17.9%	9.4%	72.7%	0.0%
HAMON	0.0%	0.0%	100.0%	0.0%
DUPONT-AIGNAN	0.0%	44.0%	0.0%	56.0%
LASSALLE	0.0%	68.1%	0.0%	31.9%
POUTOU	0.0%	100.0%	0.0%	0.0%
ASSELINEAU	79.3%	0.0%	0.0%	20.7%
ARTHAUD	0.0%	100.0%	0.0%	0.0%
CHEMINADE	0.0%	100.0%	0.0%	0.0%