

## Analyses électorales Dialogue entre statisticiens et politistes

A la veille de plusieurs échéances électorales importantes, la Société française de statistique (groupe « enquêtes, modèles ») et l'AFSP ont organisé le 21 novembre 2013 à l'Institut Henri Poincaré un séminaire d'une demi-journée, centré sur les transformations des analyses électorales en France, tenant tant à celles de leur objet, le vote, que de leurs méthodes. Son objectif a été de faire dialoguer politistes et statisticiens sur les enjeux clés que sont la participation électorale, la spatialisation des choix, le niveau d'analyse pertinent et les nouvelles techniques d'enquêtes.

### Programme & contributions

13h30-13h45	Introduction : Nonna Mayer (CEE-Sciences Po-CNRS) – Jean Chiche (CEVIPOF-Sciences Po-CNRS)
13h45-14h30	Inscription/non inscription/mal inscription Modérateur : Bruno Cautrès (CEVIPOF-Sciences Po-CNRS)
13h45-14h	Céline Braconnier (Université de Cergy Pontoise)
14h00-14h15	Alexandre Kych (Centre Maurice Halbwachs, CNRS-EHESS-ENS)
14h15-14h30	Dialogues
14h30-15h15	Les nouveaux modes de collecte d'enquêtes Modérateur : Benoit Riandey (INED)
14h30-14h45	ELIPSS : Anne Cornilleau (CDSP-Sciences Po)
14h45-15h00	DYNAMOB : Florent Gougou (CEE-Sciences Po)
15h00-15h15	Dialogues
15h15-16h00	Analyses contextuelles et multiniveaux Modérateur : Elisabeth Morand (INED)
15h15-15h30	Joel Gombin (Université de Picardie/ CURAPP)
15h30-15h45	Valérie Golaz (INED-UMR 196 CEPED)

15h45-16h00  
16h00-16h15  
16h15-17h00

#### Dialogues

#### Pause

#### Les territoires du vote

Modérateur : France Guerin Pacé (INED)

Michel Bussi (Université de Rouen)

Avner Bar-Hen (Université Paris Descartes, MAP5)

#### Dialogues

Table ronde : Les défis pour 2017 : forces/faiblesses/

Modérateur : Jean Chiche (CEVIPOF-Sciences Po-

CNRS)

François Héran (INED)

Patrick Lehingue (Université de Picardie/ CURAPP)

Jean-François Royer (INSEE)

Nonna Mayer (CEE-Sciences Po-CNRS)



# Analyses électorales : Dialogue entre statisticiens et politistes

Avner Bar-Hen

SFdS/AFSP: 20 Novembre 2013

- Des origines
  - Statistique (*Microscopium Statisticum*, Helenus Politanus, 1672) : *Staat* constitution de l'État
  - Politique : science des affaires de la Cité
  - De l'arithmétique politique à la statistique
- À nos jours
  - Statistiques
  - Sondages (dernier numéro de *Statistique et Société*)
  - Statistique et débat public

- En statistique : modélisation de la dépendance entre les données,
- En politique : compréhension spatiale des phénomènes politiques

Les données, les questions ou les possibilités des logiciels rendent les interactions directes nécessaires et donne de l'autonomie aux intermédiaires

# Pour ouvrir la discussion

- Données publiques, données ouvertes ? ▶ Exemple
- Sources très hétérogènes quand les données existent : avantages et inconvénients
- Contradiction entre données précises et confidentialité.
- Quelles échelles pour quelles questions ?
- Des questions génériques : du cas particulier à la problématique scientifique
- Apport des logiciels
- exemple ▶ Exemple
- Cadre de collaboration

Merci aux organisateurs

Profitons de cette journée pour démarrer des choses concrètes

# Des progrès et de leur limite...

Retour

Insee - Territoire - Données carroyées sur la population - Données carroyées à 200 mètres - Mozilla Firefox

Fichier Édition Affichage Historique Marque-pages Outils 2

< DENTAL VITY... quizz chi-deu... Chapter 06... Deug B - Deu... quizz sample... Sampling Dist... 16th ICID | I... quizz euclide... STAT3401 pl... statistique sp... joel gombin - ... Joel G

www.insee.fr/themes/detail.asp?reg\_id=0&ref\_id=donnees-carroyees&page=donnees-detaillees/donnees-carroyees/donnees-carroyees ☆ Joel Gombin



Institut national de la statistique  
et des études économiques

Mesurer pour comprendre

Mobile | Actualités | Agendas | Contactez-nous | Aide | FAQ | Liens | English | Home page

Chercher sur le site  Google

Accueil

Thèmes

Bases de données

Publications et services

Régions

Définitions et méthodes

Accès par public

L'Insee et la statistique publique

## Thèmes

- ▣ Agriculture
- ▣ Commerce
- ▣ Comptes nationaux - Finances publiques
- ▣ Conditions de vie - Société
- ▣ Conjoncture
- ▣ Économie
- ▣ Enseignement - Éducation
- ▣ Entreprises
- ▣ Industrie - IAA - Construction
- ▣ Population
- ▣ Revenus - Salaires
- ▣ Santé
- ▣ Services - Tourisme - Transports
- > Territoire

Accueil > Thèmes > Territoire > Régions, dé... > Données car... > Données carroyées à 200 mètres

## Données carroyées à 200 mètres

Mise à jour : 20 novembre 2013

18 variables sur la structure par âge des individus, les caractéristiques des ménages (locataire/propriétaire, etc.) et les revenus au 31 décembre 2010.

Afin de respecter la règle de diffusion des données sur les revenus fiscaux des ménages, aucune information statistique (à l'exception du nombre total d'individus) n'est diffusée sur des carreaux de moins de 11 ménages. Ces carreaux de faibles effectifs sont donc regroupés en rectangles de taille plus importante et satisfaisant à cette règle des 11 ménages minimum.

Par ailleurs, un certain nombre de variables considérées comme « à risque » ont été traitées afin que tout risque de rupture de confidentialité soit évité.

L'utilisation correcte de ces données carroyées impose donc une lecture attentive de la [documentation complète sur les données carroyées à 200 mètres](#).

## Documentation

- ▣ [Documentation synthétique sur les données carroyées à 200 mètres](#)
- ▣ [Documentation complète sur les données carroyées à 200 mètres](#)
- ▣ **Modes opératoires :**
  - ▣ [Sous Mapinfo, mode opératoire pour la métropole \(tables en \\_m\)](#) (Format PDF - 637 Ko)
  - ▣ [sous QGIS](#) (Format PDF - 494 Ko)

## Pour en savoir plus

- ▣ [Documentation générale](#)

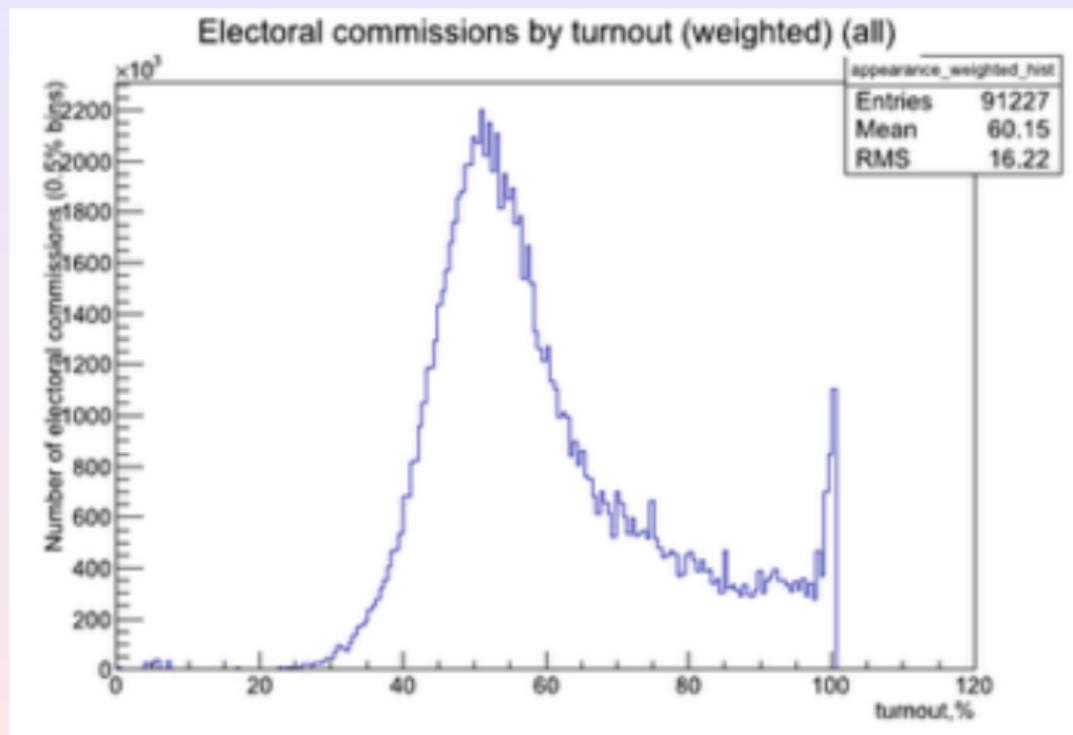
81 x



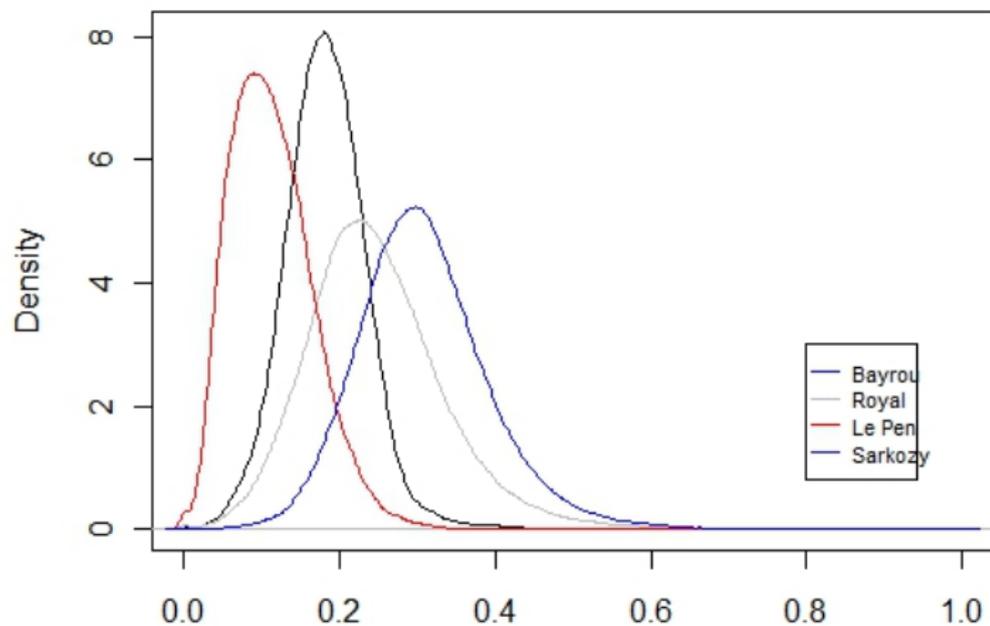
# Un exemple



# Un exemple

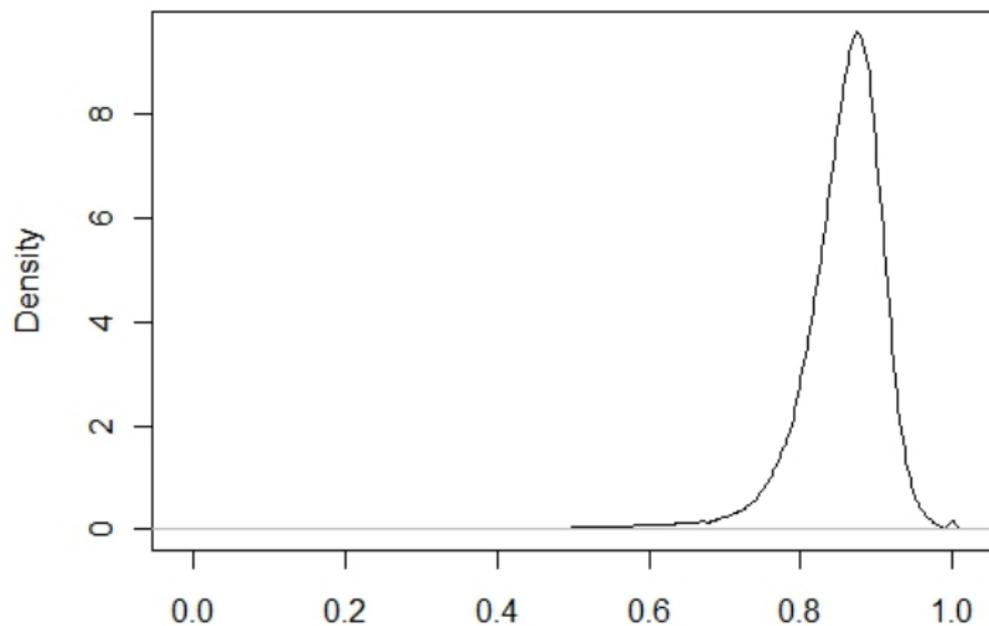


## Vote pour les candidats



N = 65608 Bandwidth = 0.004852

## participation par bureau de vote

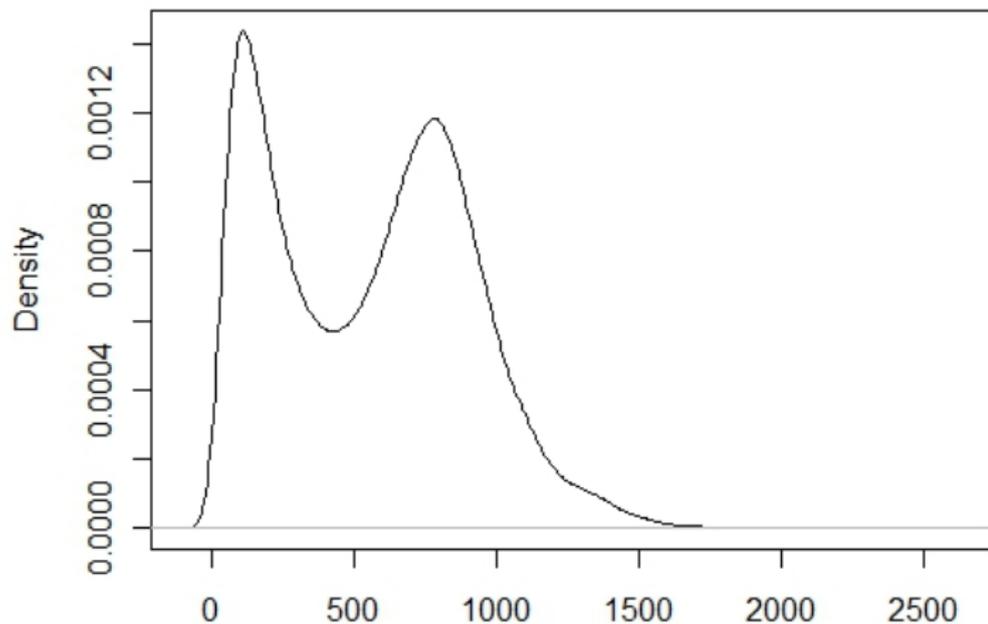


N = 65617 Bandwidth = 0.004369

# Un exemple

Retour

taille des bureaux de vote



N = 65617 Bandwidth = 34.64

21 novembre 2013, Institut National de la Statistique

**Non inscription, malinscription, abstention**  
**De l'observation localisée à l'expérimentation de terrain**

**Céline Braconnier, Université de Cergy-Pontoise, CEPEL**  
**Jean-Yves Dormagen, Université de Montpellier 1, CEPEL**

# 1-L'essentiel de la participation se joue dans l'inscription l'hypothèse énoncée depuis les Cosmonautes



## **Non inscription et malinscription : l'ampleur de la distance aux urnes aux Cosmonautes**

**Un quartier populaire de 1400 habitants**

<b>➤ Taux de participation municipales de 2001</b>	<b>41%</b>
<b>➤ Taux de participation premier tour présidentielle 2002</b>	<b>59%</b>
<b>➤ Taux de non inscription en 2002</b>	<b>25%</b>
<b>➤ Taux de malinscription en 2004</b>	<b>25%</b>

# Les effets de la malincription sur la participation aux Cosmonautes

Boîte aux lettres et présence vérifiée par enquête de voisinage (2002/2004)

	Boîte aux lettres et présence vérifiée	Pas de boîtes aux lettres et absence vérifiée
Premier tour de la présidentielle 2002	81,1	15,2
Second tour de la présidentielle 2002	87,6	48,5
Premier tour des législatives 2002	61,9	15,6
Second tour des législatives 2002	54,9	25
Premier tour des régionales 2004	66	12,5
Second tour des régionales 2004	66	8,3
Elections Européennes de 2004	35,9	4,3
Moyenne	64,8	18,5

## 2-Tester l'hypothèse énoncée aux Cosmonautes par l'expérimentation : le modèle de Yale

### The Effects of Canvassing, Telephone Calls, and Direct Mail on Voter Turnout: A Field Experiment

ALAN S. GERBER and DONALD P. GREEN *Yale University*

*We report the results of a randomized field experiment involving approximately 30,000 registered voters in New Haven, Connecticut. Nonpartisan get-out-the-vote messages were conveyed through personal canvassing, direct mail, and telephone calls shortly before the November 1998 election. A variety of substantive messages were used. Voter turnout was increased substantially by personal canvassing, slightly by direct mail, and not at all by telephone calls. These findings support our hypothesis that the long-term retrenchment in voter turnout is partly attributable to the decline in face-to-face political mobilization.*

During the last half-century, a dramatic transformation has occurred in the manner in which voters are mobilized. The election campaigns described by Gosnell (1937), Sayre and Kaufman (1960, chap. 6), and Wolfinger (1974, chap. 4) relied heavily on face-to-face contact between voters and those seeking their support. Notably absent from such accounts are professional campaign consultants, direct mail vendors, and commercial phone banks, all of which have gradually replaced work performed by party activists. The advent of modern campaign tactics (Broder 1971; Ware 1985) has coincided with a decline in the proportion of adults who report working for a political party. Based on an annual aggregation of Roper surveys between 1973 and 1994, Putnam (2000, 41) reports a steady decline in this proportion: Whereas 6% of the public reported working for a political party in the early 1970s, just 3% did so in the mid-1990s.

At the same time, there has been a marked decline in the size and vitality of nonpartisan organizations. In the mid-1960s, 2.4 of every 1,000 women over the age of 20 belonged to the League of Women Voters, compared to .79 in 1998 (Putnam 2000, 438–44). A similar fate has befallen such civic organizations as the Lions, Rotary, and Kiwanis Clubs, which have experienced sharp membership declines since the 1960s. Due to the changing character of both partisan and nonpartisan organizations, voter mobilization has become increasingly impersonal, and messages that once might have been delivered in person are now communicated using mass marketing techniques.

The decline of personal mobilization has arguably contributed to the erosion of voter turnout in the United States since the 1960s. This hypothesis is related to, yet distinct from, Rosenstone and Hansen's (1993) contention that diminishing rates of turnout are a result of a decline in the volume of mobilization

activity. As Abramson, Aldrich, and Rohde (1998, 85) point out in their discussion of Rosenstone and Hansen, there has been no clear trend over time in the proportion of American National Election Study (ANES) respondents who report some form of contact with political parties or campaigns, whether personal or impersonal. When read in conjunction with the trend lines compiled by Putnam, the ANES data are consistent with the view that campaigns are reaching as many people as ever but through less personal means.

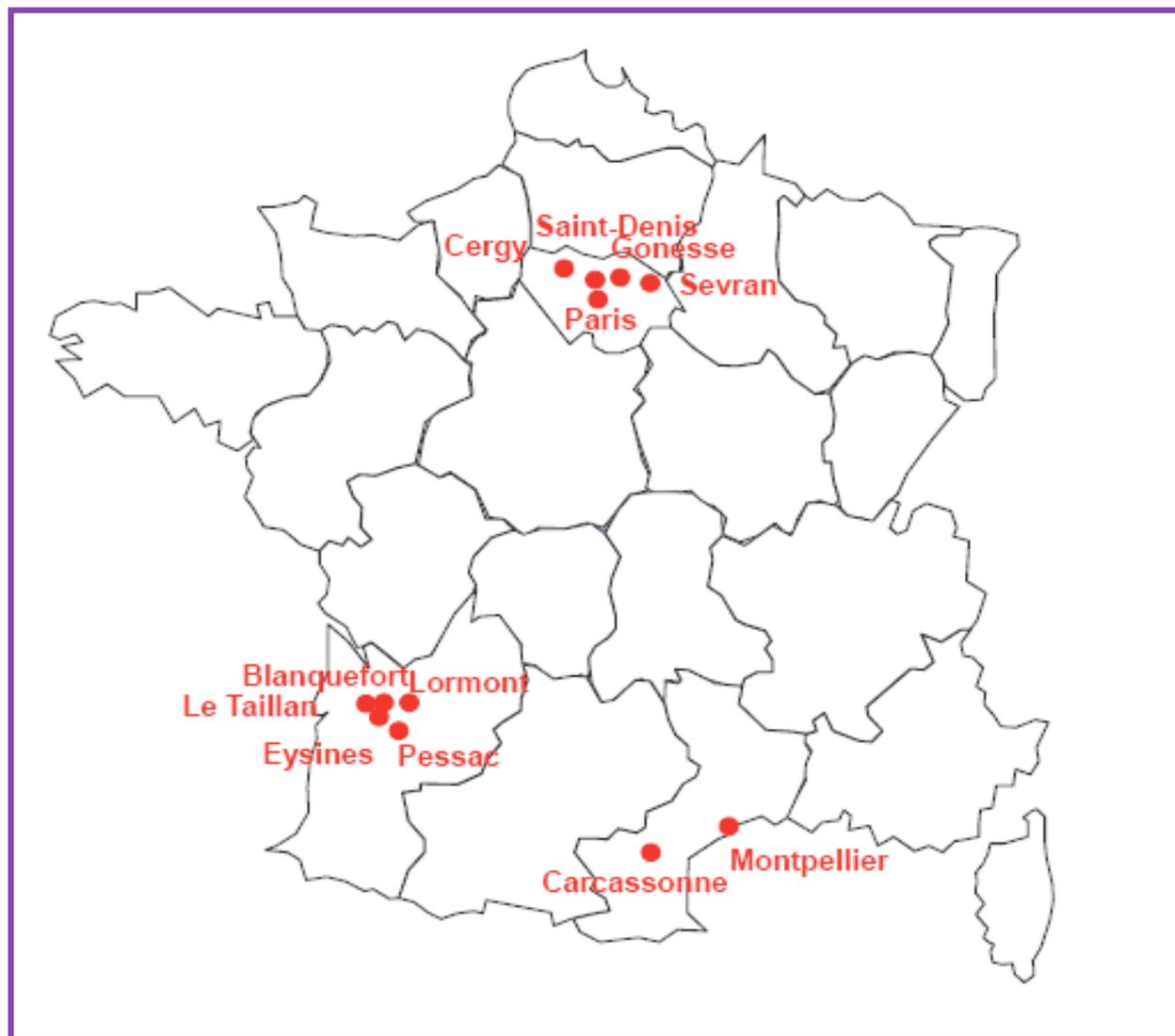
Our hypothesis about declining turnout rates rests on the claim that personal canvassing mobilizes voters more effectively than other modes of contact that have taken its place, such as direct mail or telephone appeals. The literature on collective action and prosocial behavior supports this conjecture. Studies of blood donations, recycling, and "good deeds" underscore the importance of delivering urgent requests and making vivid the obligation to act (Christensen et al. 1998; Wang and Katzev 1990), and these blandishments seem particularly effective if delivered in person (Spaccarelli, Zolik, and Jason 1989). The special force of face-to-face contact is illustrated by Reams and Ray (1993) and Jason et al. (1984), whose experiments demonstrate that recycling and blood donations are particularly responsive to in-person appeals.

There is good reason to suspect that personal canvassing is an effective means by which to mobilize voter turnout, but its effects have seldom been gauged reliably. Nonexperimental studies, beginning with Kramer (1970), tend to rely on survey data to examine the relationship between turnout and reported contacts with political organizations or candidates (Blydenburgh 1971; Cain and McCue 1985; Caldeira, Clausen, and Patterson 1990; Kramer 1970; Lupfer and Price 1972; Price and Lupfer 1973; Wielhouwer and Lockerbie 1994; see also Huckfeldt and Sprague 1992). Rosenstone and Hansen (1993), for example, regress reported voter turnout on reported contact with candidates or political parties. An important drawback to this approach is that political contact may not be an exogenous predictor of turnout. If parties direct their appeals disproportionately to committed partisans, those most likely to vote will also be most likely to receive contact, and the apparent link between contact and turnout may be spurious. Regression analyses generally include a host of control variables, but it is

Alan S. Gerber is Associate Professor of Political Science and Donald P. Green is Professor of Political Science, Yale University, 124 Prospect Street, New Haven, CT 06520-8301.

The authors are grateful to the Smith Richardson Foundation and the Institution for Social and Policy Studies at Yale, which funded this research but bear no responsibility for the content of this report. We are grateful to the League of Women Voters, New Haven Chapter, to Barry McMillion and Claire Nee, who coordinated the voter mobilization campaign.

**10 villes et 48 bureaux de vote pour mieux comprendre  
l'exclusion électorale par l'inscription  
(expérimentation réalisée avec Vincent Pons)**



## L'échantillon

-48 bureaux de votes avec une faible participation, 4118  
Bâtiments , 38 000 citoyens

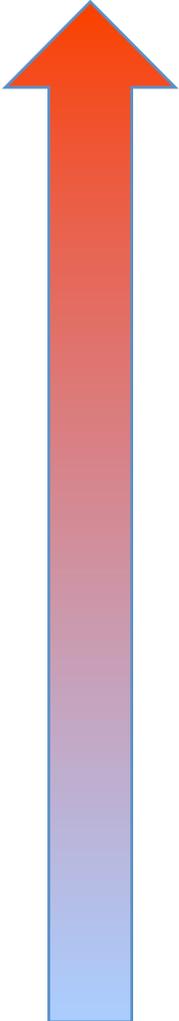
-Identification des foyers où résident des citoyens non inscrits  
et mal inscrits par le relevé des boîtes aux lettres

-230 mobilisateurs pour frapper aux portes des non inscrits et  
malinscrits : étudiants / militants de partis politiques /  
membres d'associations

# Deux grands types d'intervention

Nombre de NI et MI  
touchés en 1 heure

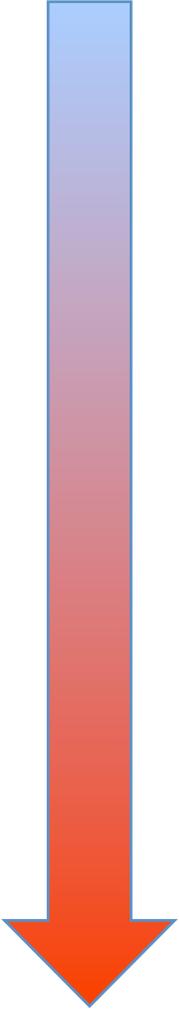
Proportion des NI et  
MI touchés qui  
s'inscrivent



## Porte-à-porte simple

- ✓ Vérifier si l'interlocuteur et autres membres du ménage sont inscrits et inscrits à cette adresse
- ✓ Encourager les non inscrits (NI) et mal-inscrits (MI) à s'inscrire
- ✓ Leur expliquer comment s'inscrire (pièces nécessaires, horaires d'ouverture de la mairie, etc.) et laisser le tract

## Aide à l'inscription à domicile

- ✓ Vérifier si l'interlocuteur et autres membres du ménage sont inscrits et bien inscrits
  - ✓ Encourager les NI et MI à s'inscrire
  - ✓ Leur proposer de les aider à s'inscrire: remplir avec eux la demande d'inscription; prendre des photos d'une pièce d'identité et justificatif de domicile (ou leur demander de faire des photocopies)
  - ✓ Prendre RdV pour compléter les dossiers incomplets et aider d'autres membres du ménage absents à s'inscrire
- 

# La répartition dans les groupes d'intervention

## Méthode: « randomisation »

- Foyers avec malinscrits et non inscrits répartis aléatoirement en plusieurs groupes à l'échelle du bâtiment
- Chaque groupe est caractérisé par les interventions reçues pendant le 1<sup>er</sup> et le 2<sup>nd</sup> passages
- Tous les groupes sont symétriques. Les différences mesurées en termes de nombre de nouveaux inscrits peuvent donc être attribuées aux interventions, à l'exclusion de tout biais

	1 <sup>er</sup> passage	2 <sup>nd</sup> passage
Groupe 1 (contrôle)		
Groupe 2	Porte-à-porte	
Groupe 3		Porte-à-porte
Groupe 4	Inscription à domicile	
Groupe 5		Inscription à domicile
Groupe 6	Porte-à-porte	Inscription à domicile
Groupe 7	Inscription à domicile	Inscription à domicile

## **Les questions structurantes :**

**Que se passe t-il quand on allège ou supprime le coût de l'inscription ?**

- Le porte à porte est-il efficace pour réduire la non inscription et la malinscription ?**
- L'aide à l'inscription à domicile donne-t-elle de meilleurs résultats qu'une simple campagne d'information et de mobilisation citoyenne ?**
- Ces interventions sont-elles plus efficaces en décembre que dans la période précédente ?**
- Les inscrits à domicile participent-ils moins aux élections (moins politisés, moins investis, plus passifs...) ?**
- L'inscription à domicile peut-elle nuire à l'appropriation de la citoyenneté (thématique de l'empowerment) ?**

# La campagne de mobilisation

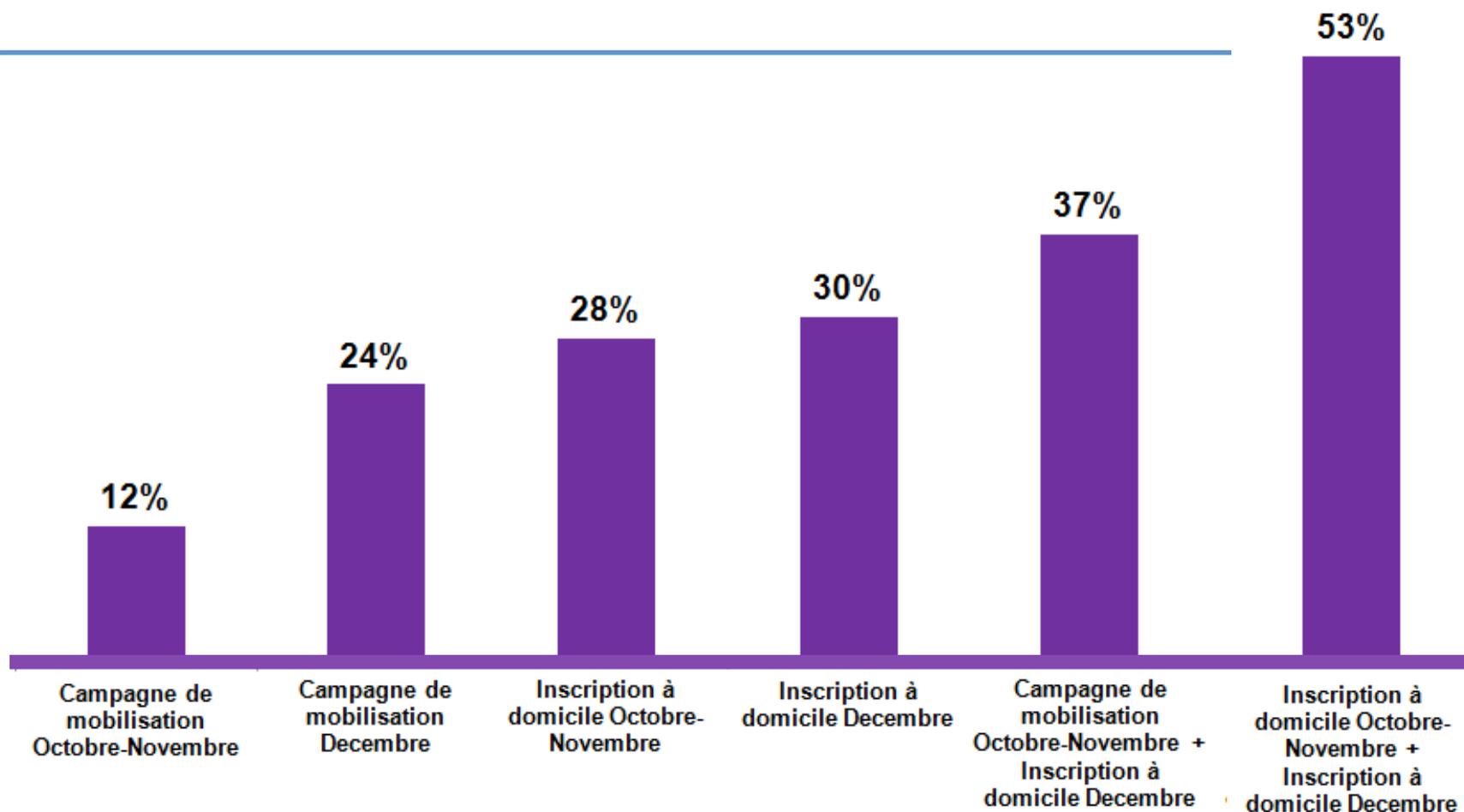
Les étudiants de Montpellier, de Cergy, de l'ENS, des IEP de Bordeaux et Paris



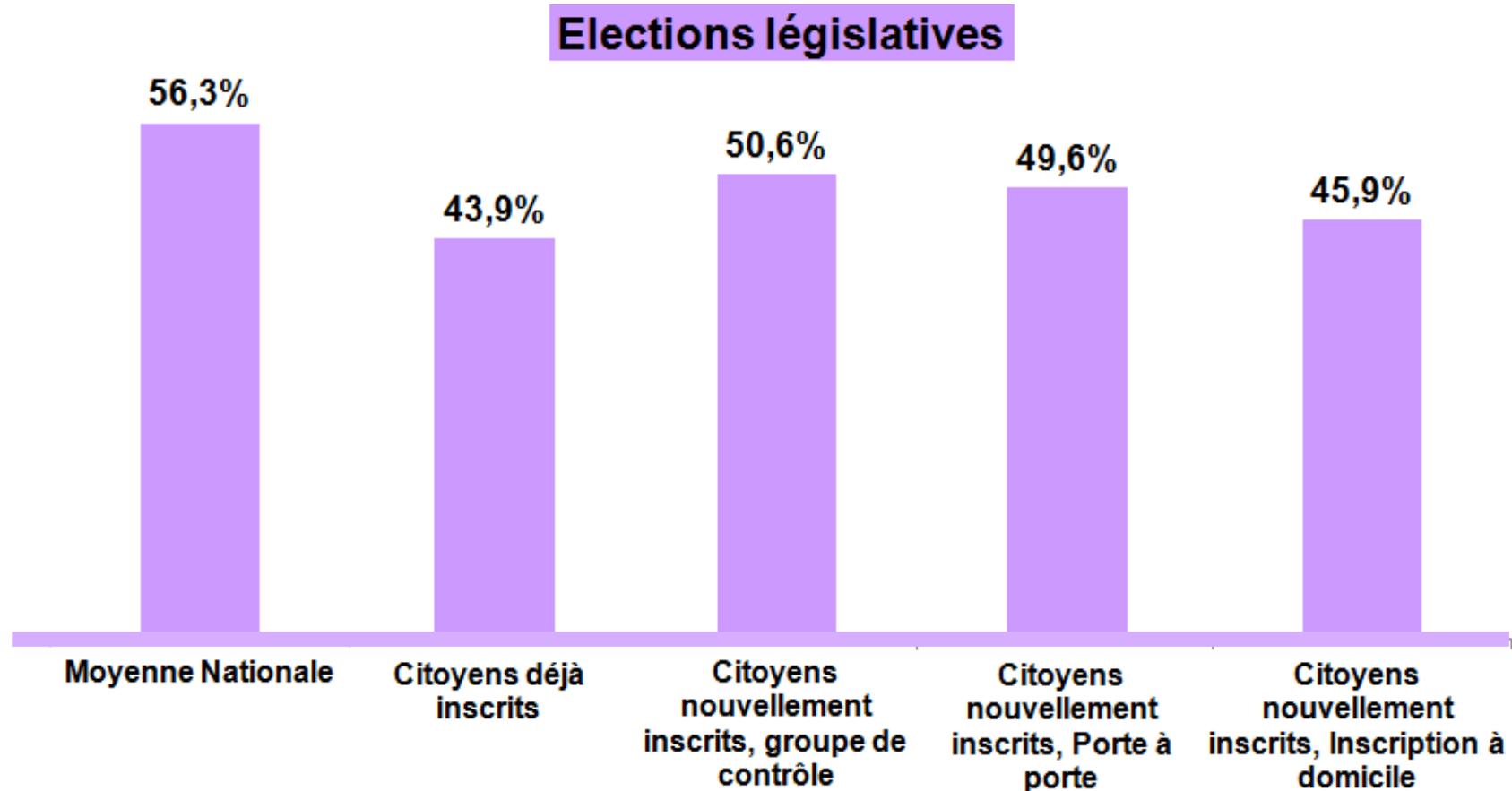
## Les retraitées de la MGEN inscrivent les Dyonisiens à domicile



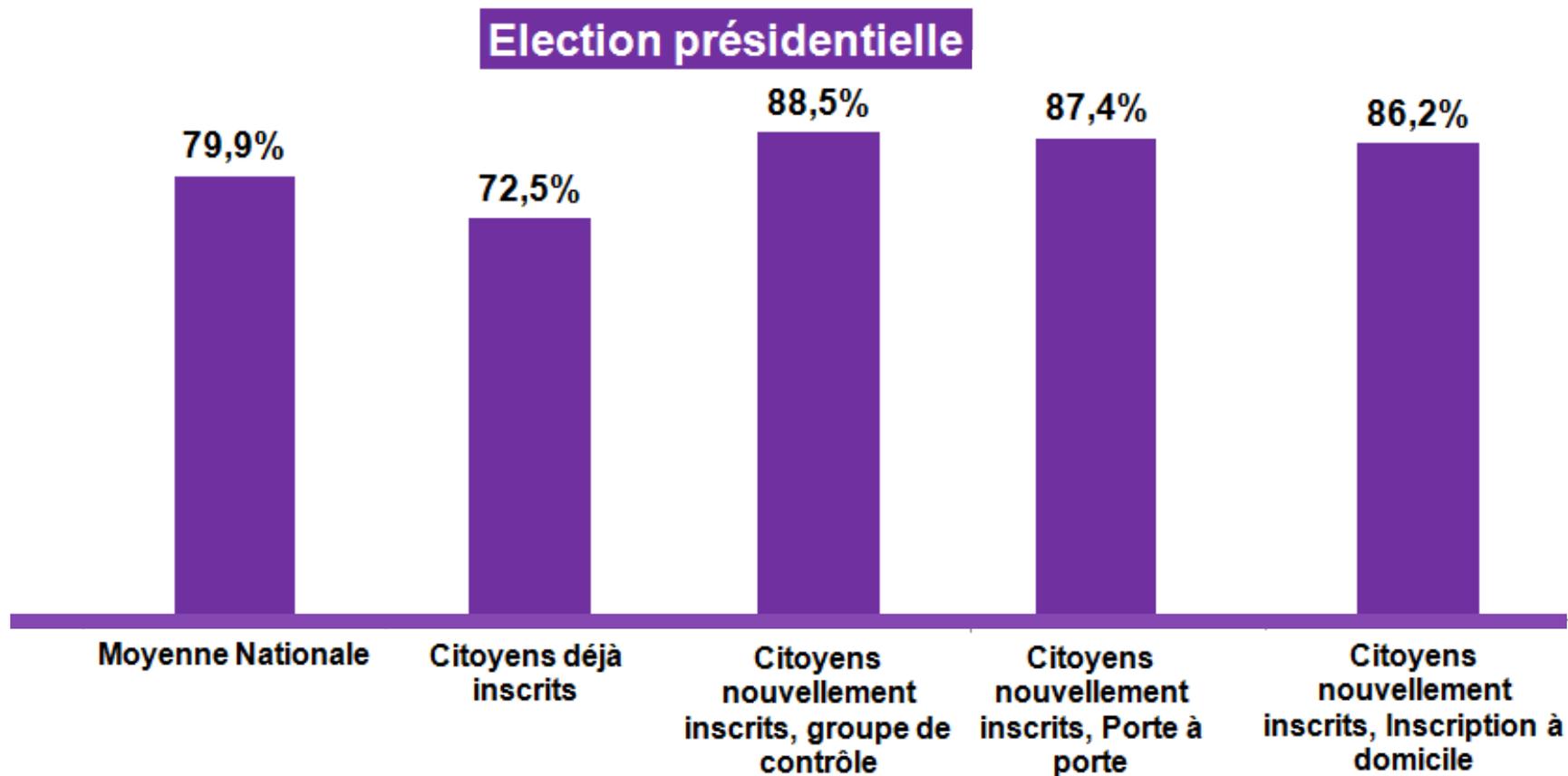
## Résultats : hausse des taux de nouvelles inscriptions sur les listes électorales des bureaux de vote, comparés à celles du groupe de contrôle



## Résultats : participation des nouveaux inscrits par grands types d'intervention et par comparaison avec le groupe de contrôle et les anciens inscrits



## Résultats : participation des nouveaux inscrits par grands types d'intervention et par comparaison avec le groupe de contrôle et les anciens inscrits



## **Mesurer la diffusion de la malinscription dans l'espace social**

L' étude en cours avec Xavier Niel (Insee)

- Etablir l'ampleur de la malinscription au niveau national en mesurant la distance qui sépare le lieu de vote des individus identifié grâce à l'enquête participation 2012 (40 000 Inscrits) de leur lieu de résidence effectif établi par bulletins de recensement 2010.
- Etablir les déterminants sociaux de la malinscription
- Etablir les effets de la malinscription sur la participation aux scrutins présidentiel et législatif en mettant en évidence les facteurs décisifs de leur variation par une modélisation (degrés d'éloignement, catégories d'âges, statut professionnel, niveau de diplôme, taille des agglomérations, genre)



# Territoires, analyses géographiques du vote , *Cartelec, le futur des analyses territoriales.*

*Michel Bussi, UMR CNRS IDEES  
Université de Rouen,  
Pour l'équipe Cartelec*

# Approche contextuelle ou analyse spatiale

- La géographie électorale a été souvent cantonnée à l'analyse cartographique des résultats électoraux
- La justification de l'emploi des sondages d'opinions pour expliquer les votes s'est opérée dans un premier temps par la critique de l'approche écologique (son usage fut hégémonique avant que les sondages ne la supplantent comme outil d'analyse).

Depuis une dizaine d'années, l'approche cartographique retrouve un intérêt majeur, profitant de **l'essor des SIG** et de l'informatisation des données.

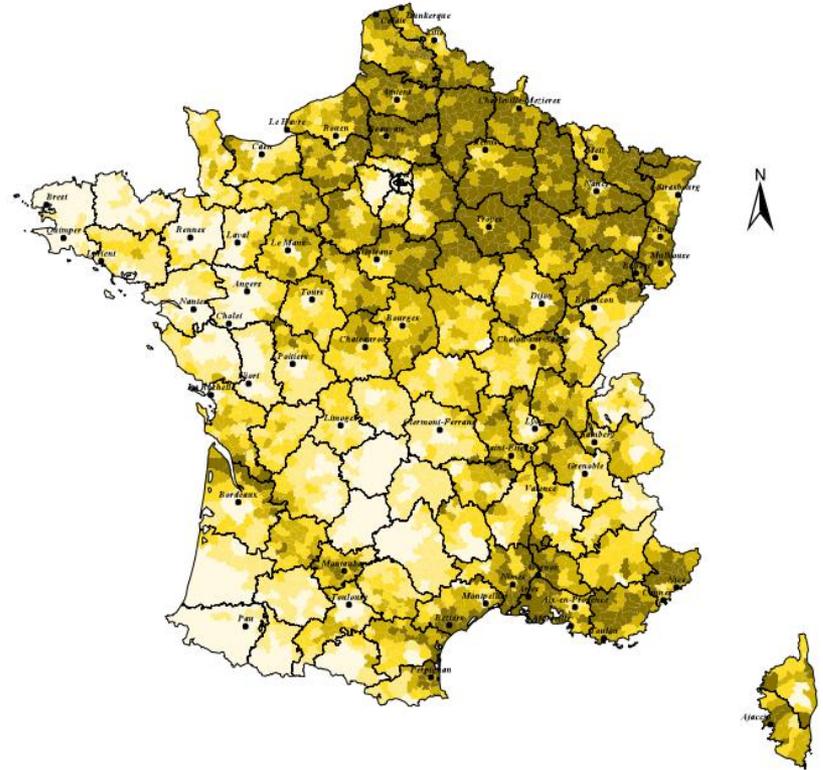
- le ministère de l'intérieur fournit désormais les données électorales sur support informatique pour les 36 000 communes françaises dans la semaine qui suit chaque scrutin.
- Mais ces cartes survalorisent les espaces ruraux et rendent invisibles les enjeux urbains

La survalorisation d'une grille de lecture « ruralocentrée » tient notamment à une question technique : les données électorales agrégées sont connues à l'échelle des principaux zonages républicains (communes, cantons, départements pour la France).

Tous écrasent la diversité urbaine, en découpant d'autant plus finement les territoires qu'ils sont peu densément peuplés.

Ainsi, en France, les cartes électorales dévoilent avant tout des phénomènes ruraux.

Vote pour Marine Le Pen  
(en pourcentage des suffrages exprimés)



- Principales villes françaises
- Limites départementales
- Score de Marine Le Pen :
- 5.00 - 15.05
- 15.05 - 18.27
- 18.27 - 21.37
- 21.37 - 24.84
- 24.84 - 28.31

0 200 400 Km.

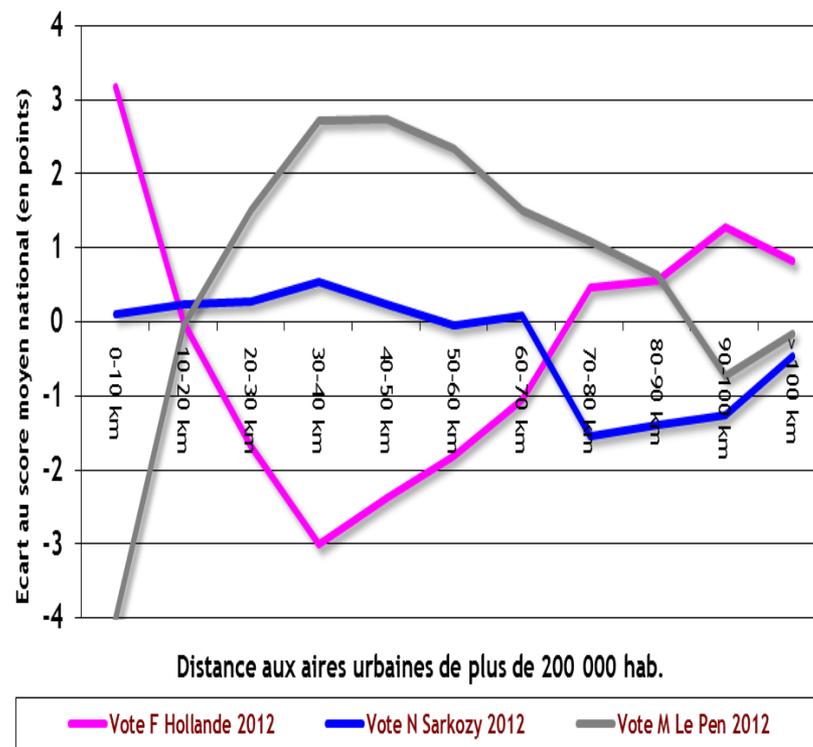
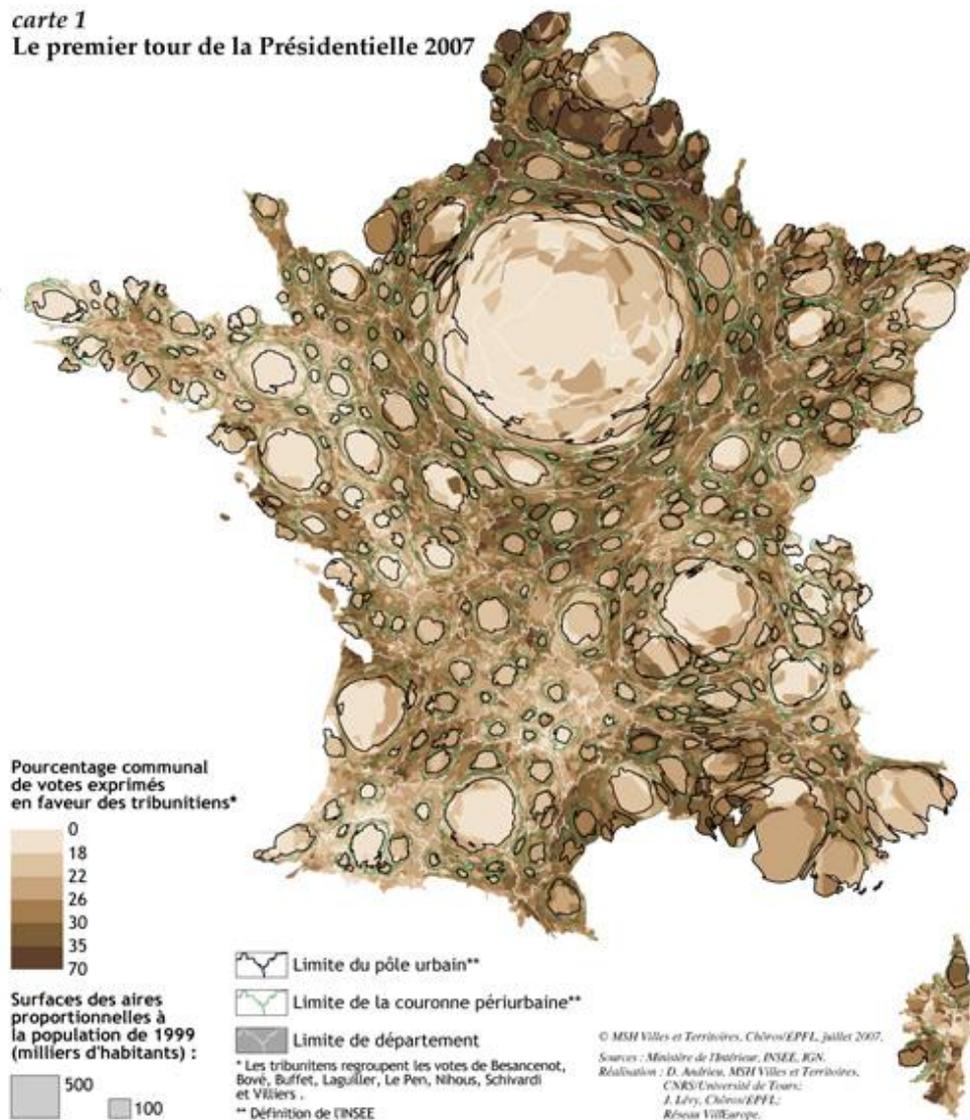
Conception, Réalisation : Céline COLANGE et Jean-Paul GOSSET.  
UMR 6266 IDEES, Université de Rouen.

Source : Ministère de l'Intérieur, avril 2012

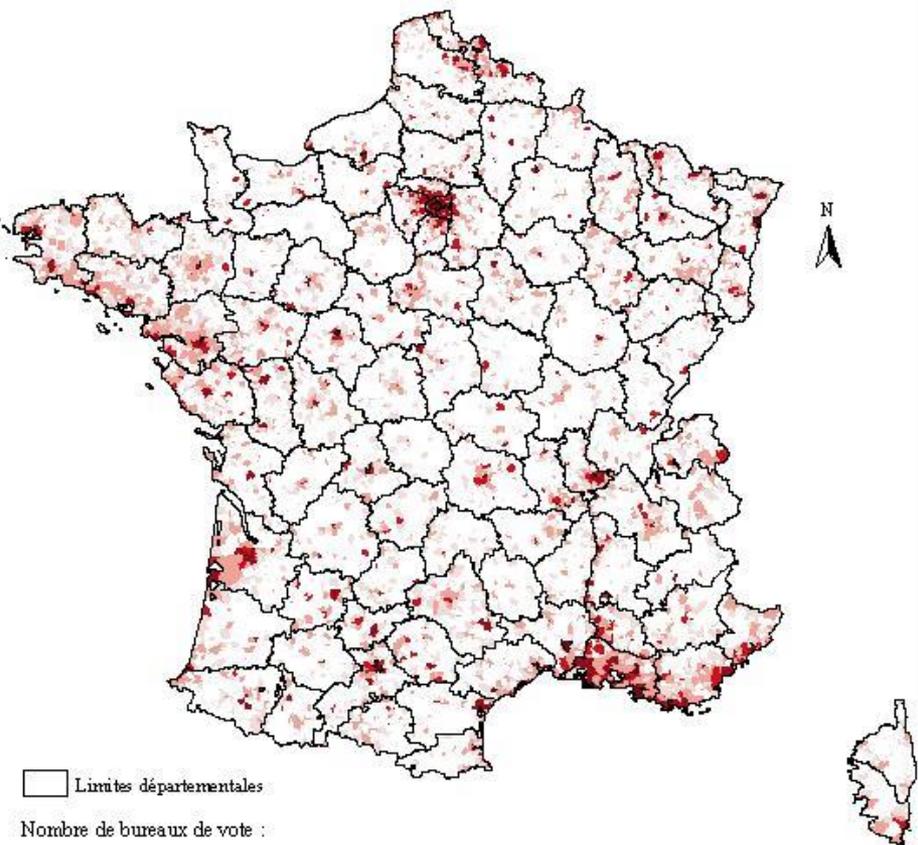


Pourtant, la relation politique à la ville soulève des débats sociétaux, tel le statut de l'urbanité selon J.Lévy dans Libération, ou la mesure du « gradient d'urbanité...

carte 1  
Le premier tour de la Présidentielle 2007



### Nombre de bureaux de vote par commune en 2002

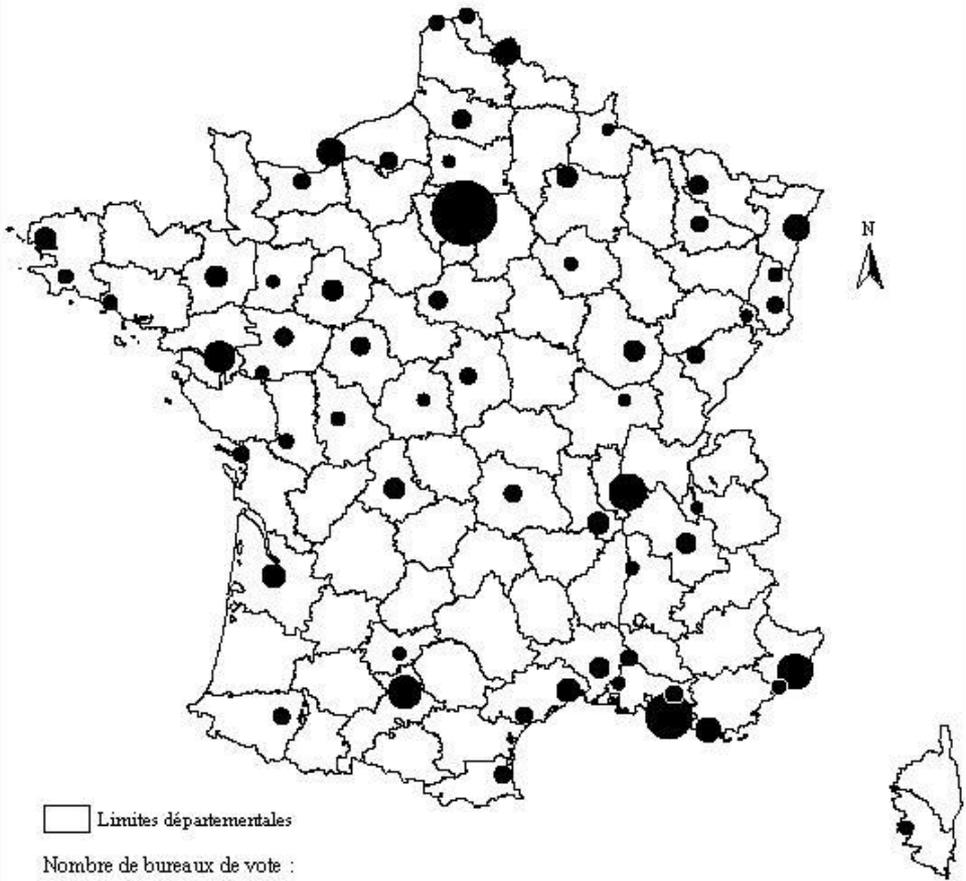


- Limites départementales
- Nombre de bureaux de vote :
- un seul bureau de vote
- deux bureaux de vote
- de 3 à 6 bureaux de vote
- de 7 à 14 bureaux de vote
- de 15 à 49 bureaux de vote
- de 50 à 914 bureaux de vote
- absence de données

0                      200                      400 Km

*Conception, Réalisation : Céline COLANGE et Michel BUSSI  
Laboratoire MTG - UMR 6266 IDEES  
Université de Rouen  
Source : Ministère de l'Intérieur.*

### Nombre de bureaux de vote dans les grandes communes en 2002



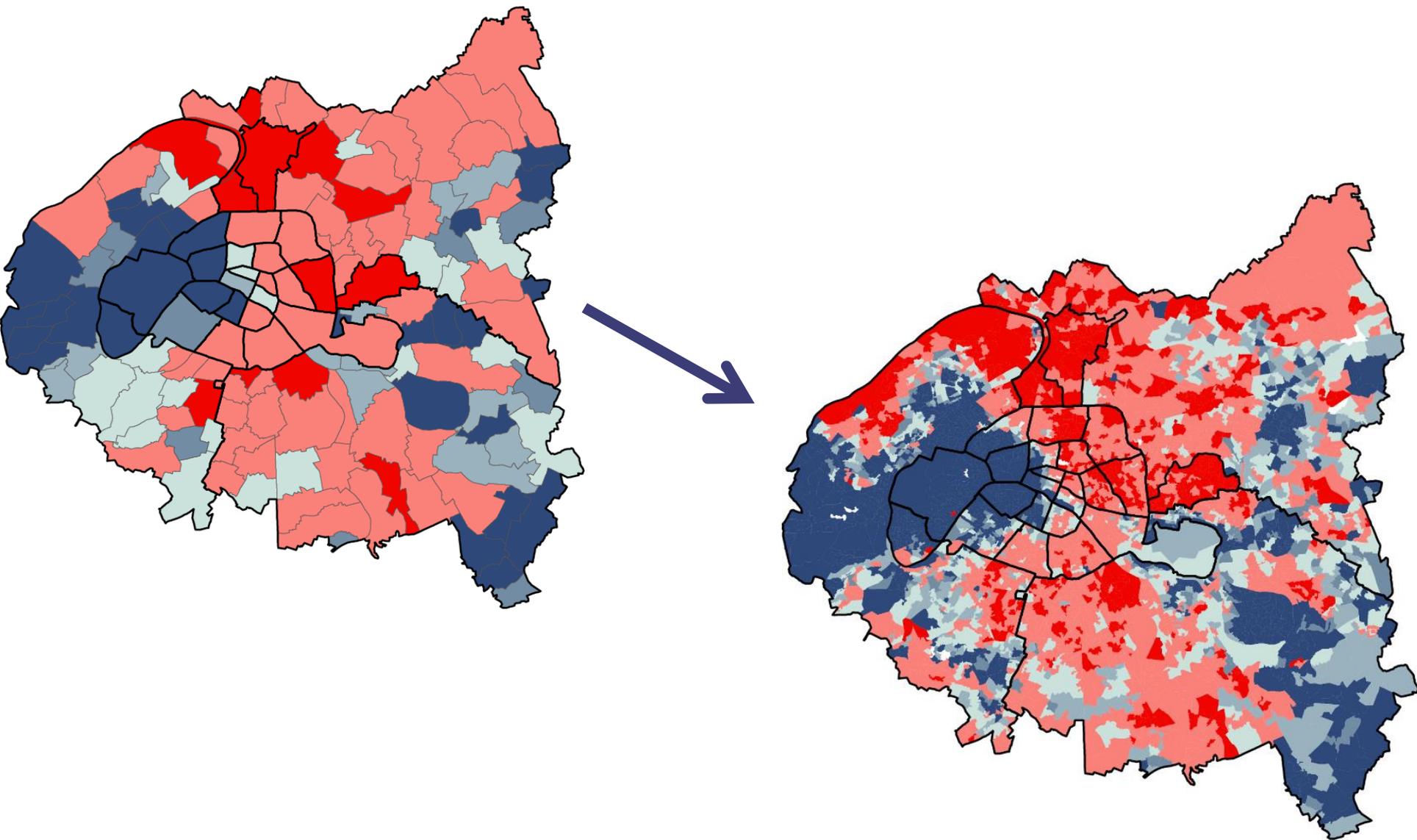
- Limites départementales
- Nombre de bureaux de vote :
- 28
- 150
- 444
- 914

0                      200                      400 Km

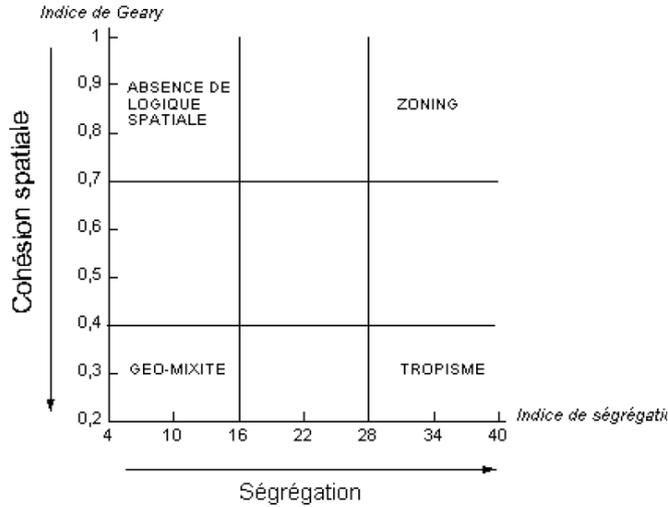
*Conception, Réalisation : Céline COLANGE et Michel BUSSI  
Laboratoire MTG - UMR 6266 IDEES  
Université de Rouen  
Source : Ministère de l'Intérieur.*

S = Q

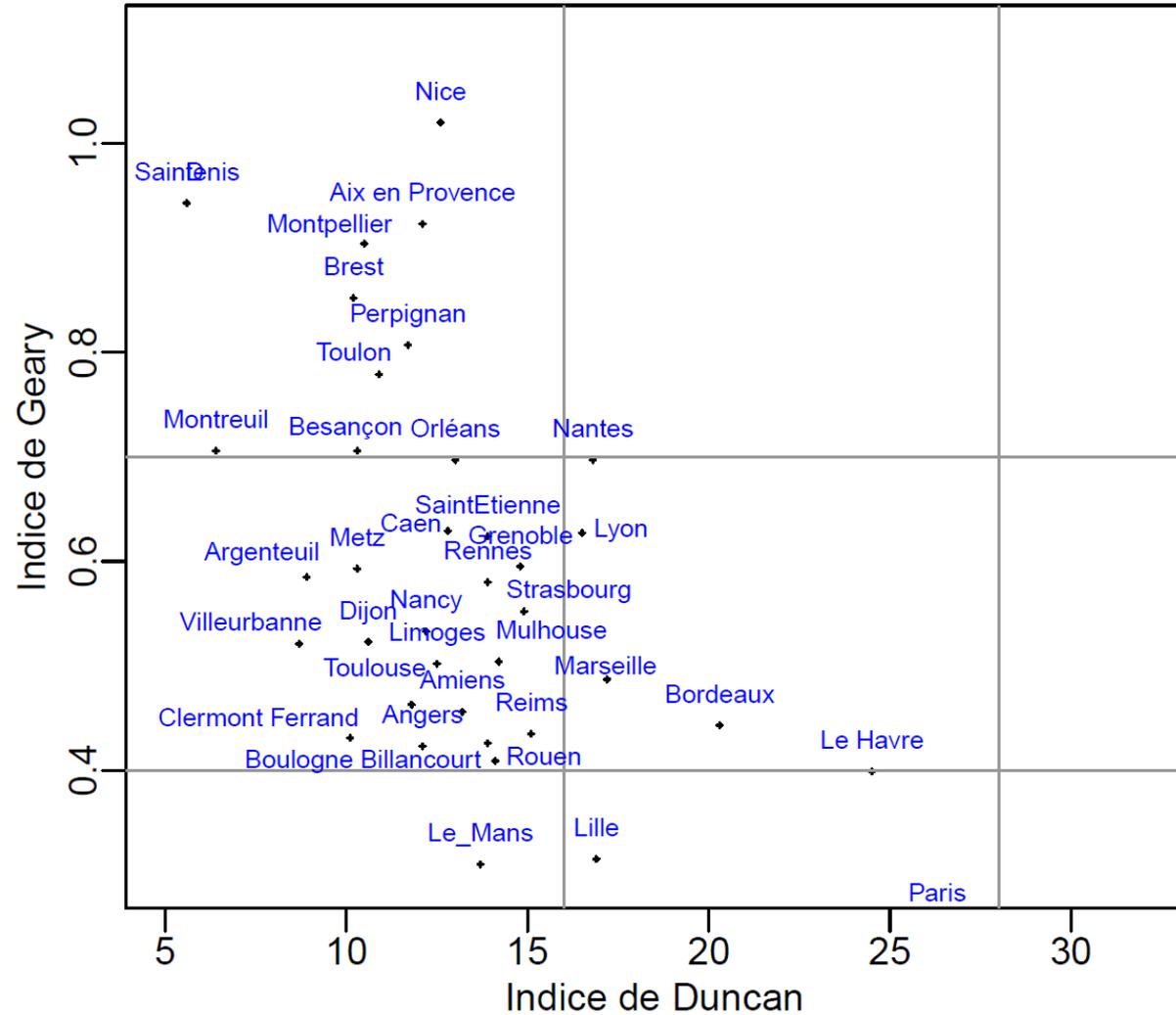
Le résultat : de 143 communes à 3347 bureaux



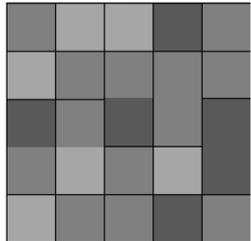
# Sarkozy, Présidentielles 2007 - T1



Source : F. Girault, 2000.



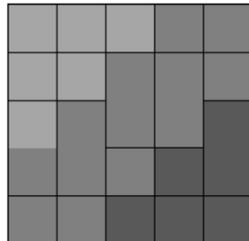
Absence de logique spatiale



C=1.1, ID = 9.6

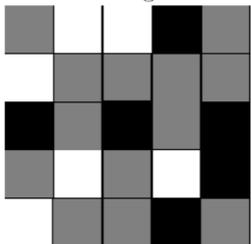
Zonage

Géo-mixité



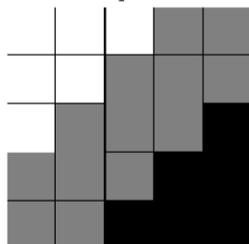
C = 0.31, ID = 9.6

Tropisme



C = 1.01, ID = 38.4

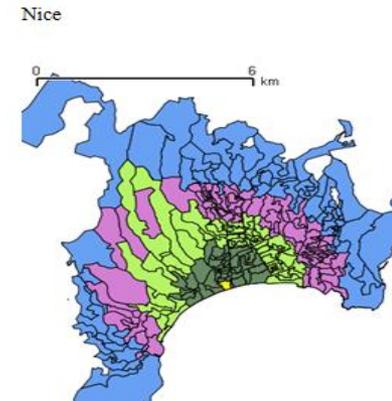
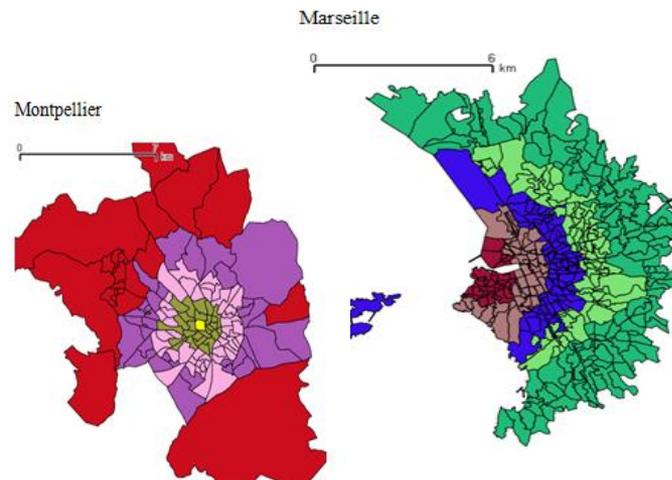
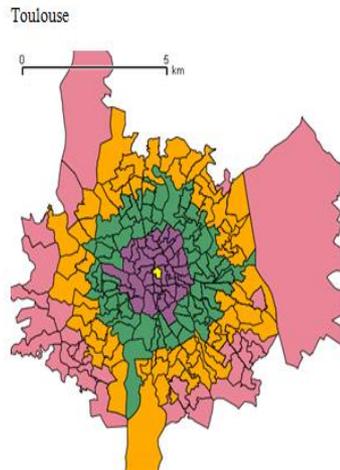
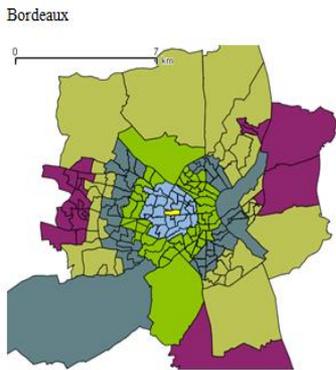
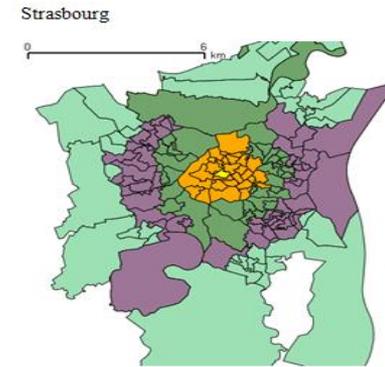
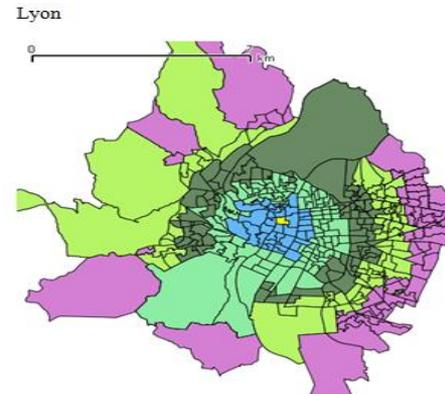
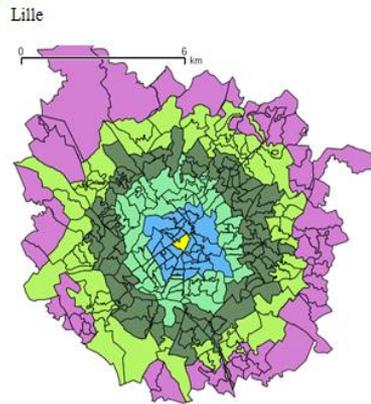
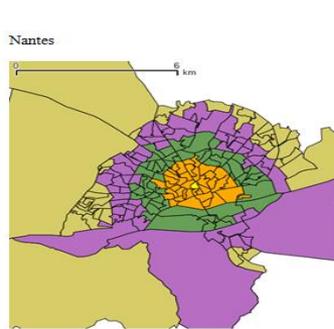
Tiré de Girault et Bussi, 2001.



C = 0.31, ID = 38.4

# Vote et distance au centre ville : tests sur l'intra-urbain

- Point au centre de la commune (BV central, par exemple : Toulouse, Capitole ; Strasbourg : juste au dessus de la petite France ; Lille : centre – Mairie)
- 4 ou 5 couronnes, en fonction de l'étendue de la commune, avec pas de 1000 m

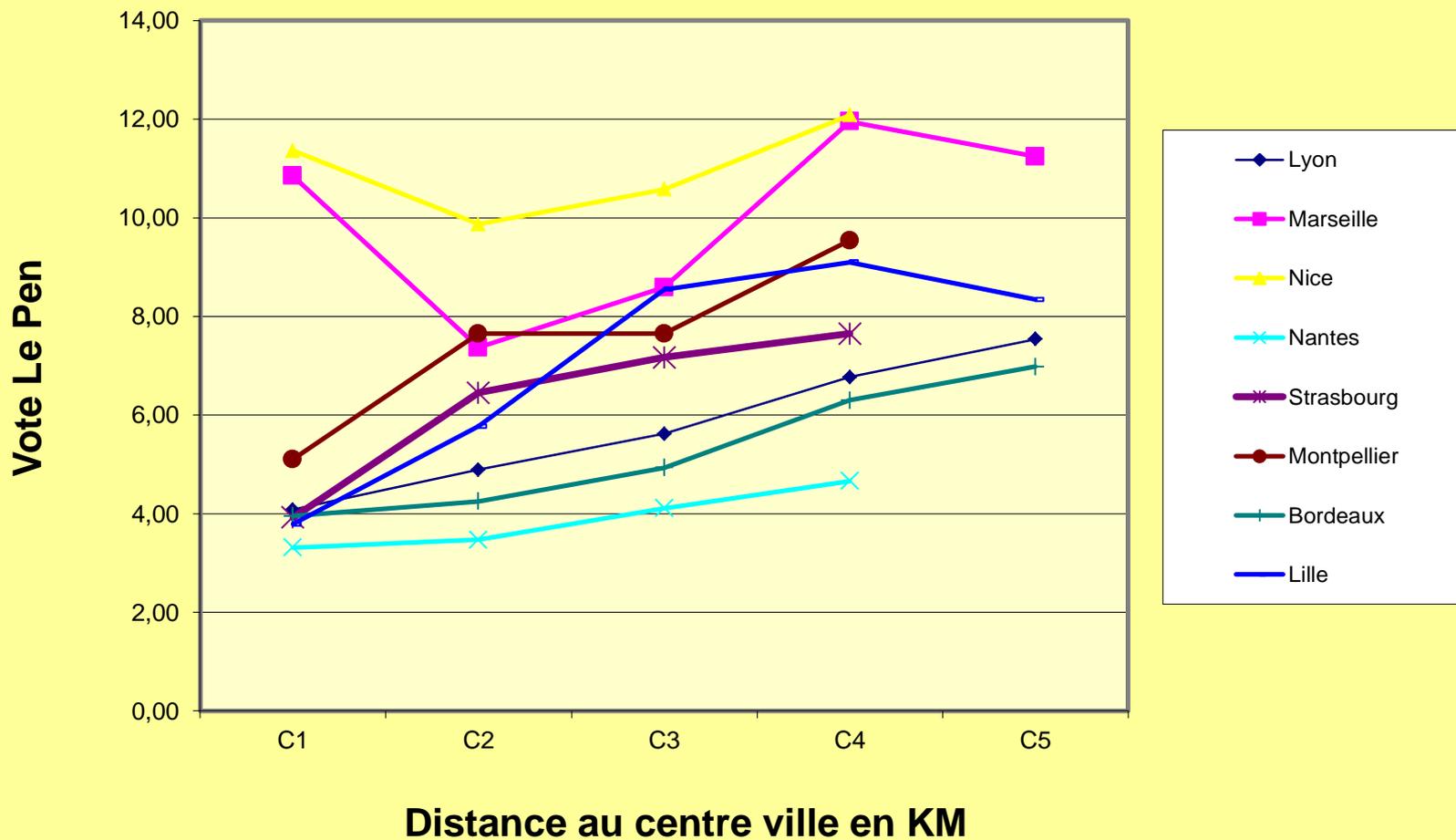




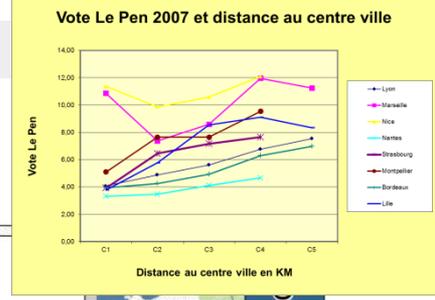
## ***Hypothèse centrale : le vote comme indicateur de ségrégations urbaines***

- Il existe un gradient de vote au sein des communes-centre des agglomération : gentrification des centres, relégation en périphérie des classes sociales dominées
- Ce gradient centre-périphérie s'explique notamment par celui du coût de l'immobilier dans les centres villes
- L'implantation des partis est plus ou moins sensible à ce gradient

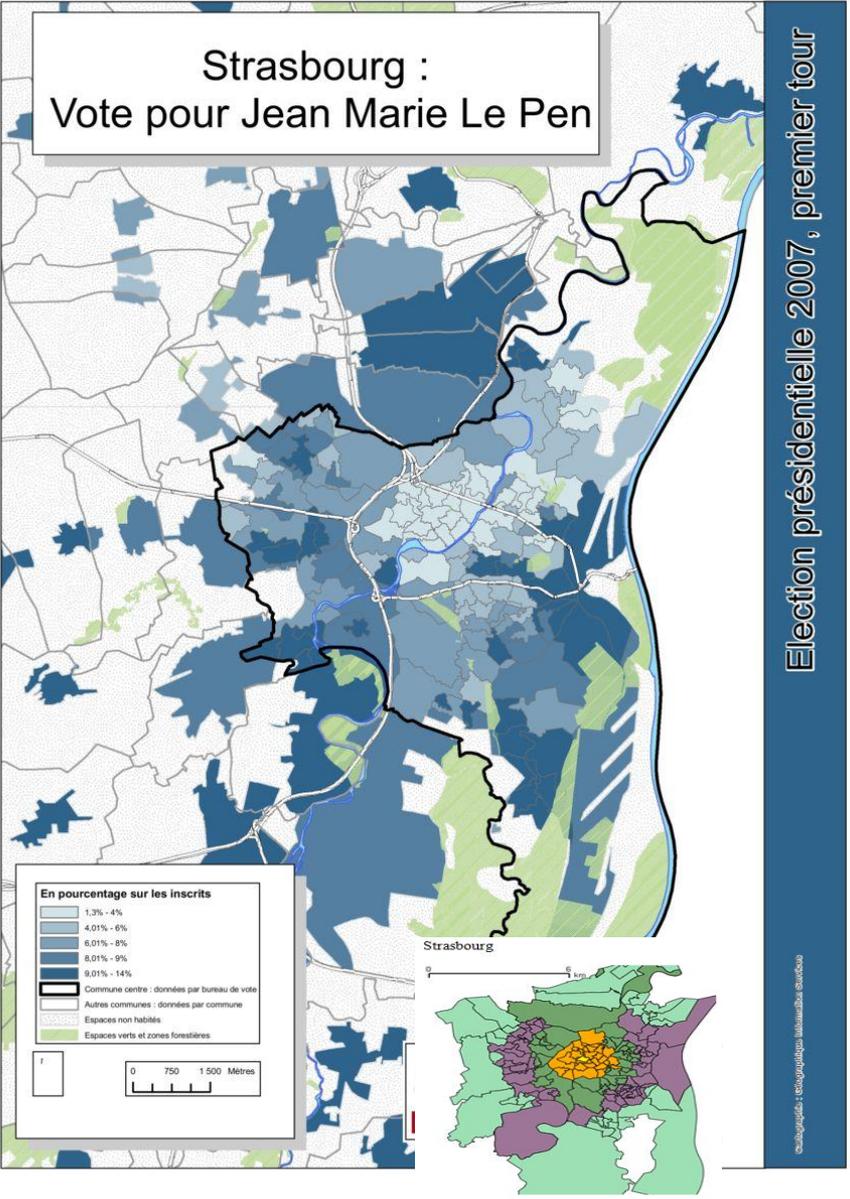
# Vote Le Pen 2007 et distance au centre ville



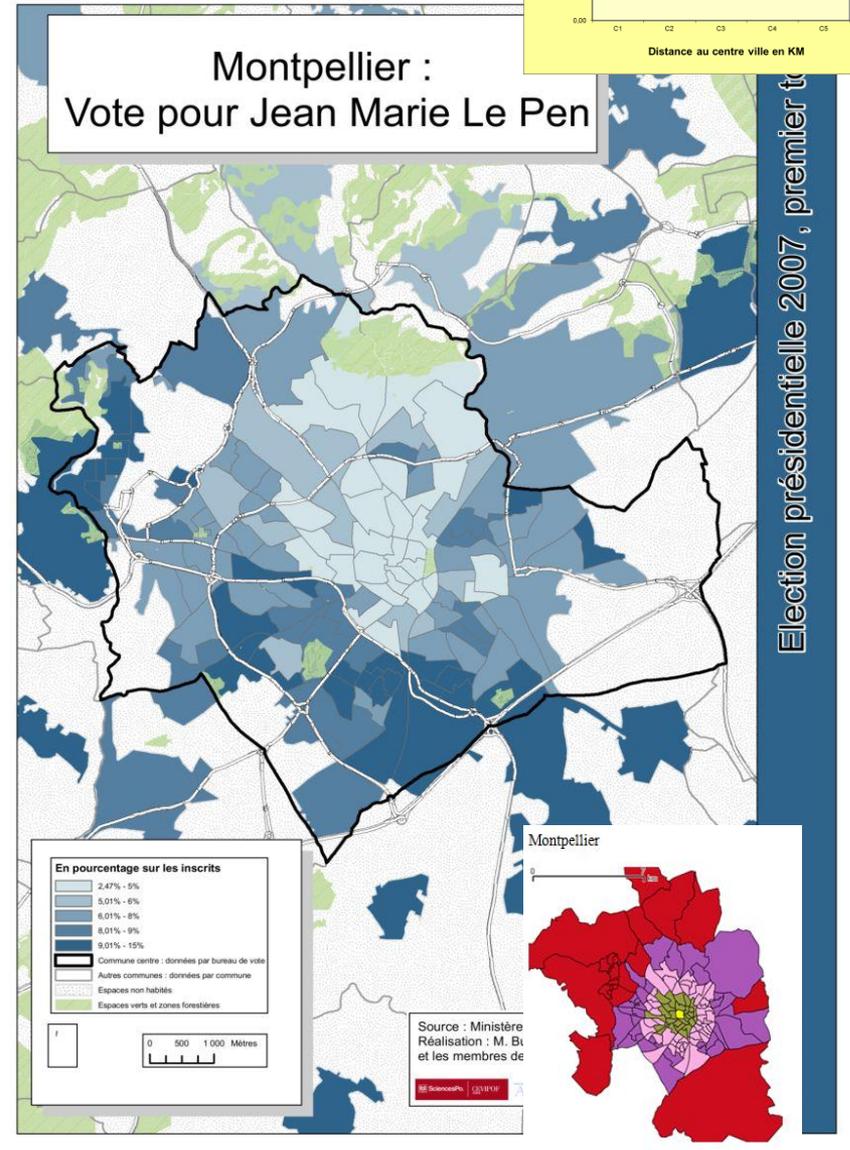
# L'effet centre-périérie apparaît souvent avec évidence



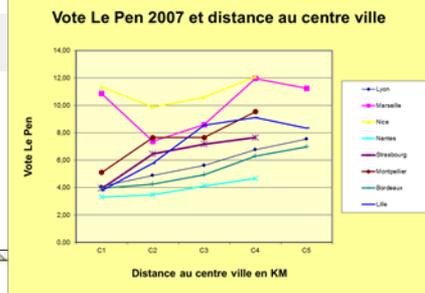
Strasbourg :  
Vote pour Jean Marie Le Pen



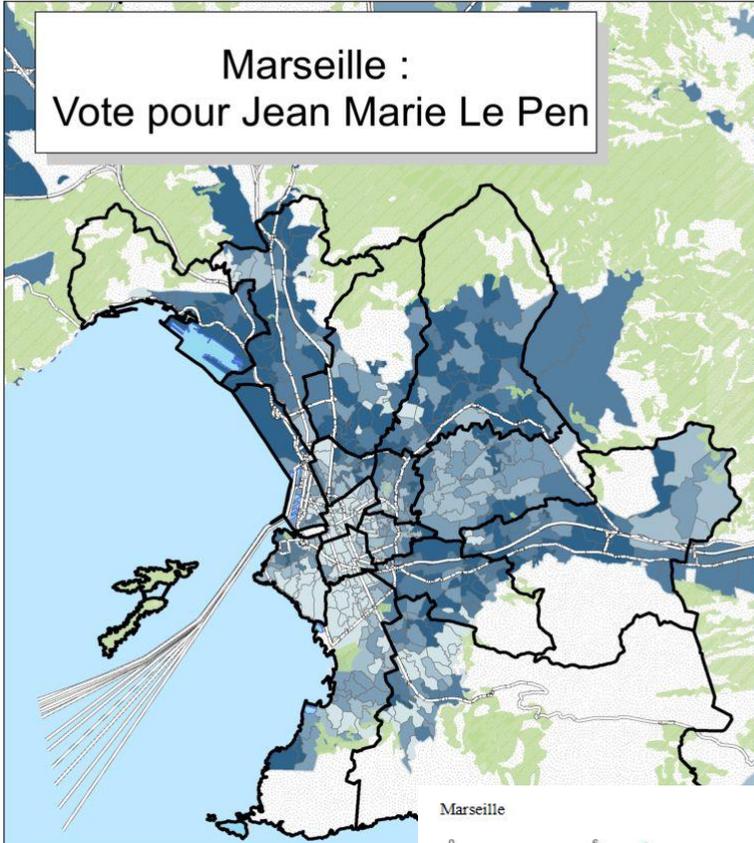
Montpellier :  
Vote pour Jean Marie Le Pen



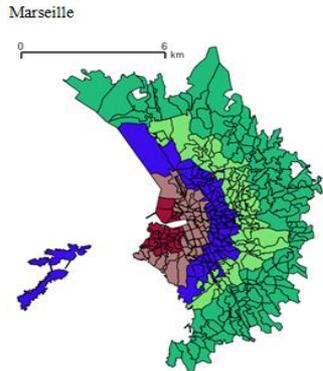
Même dans les cas de Nice et de Marseille, l'effet centre-périphérie semble réel, mais non mesurable par un simple modèle concentrique



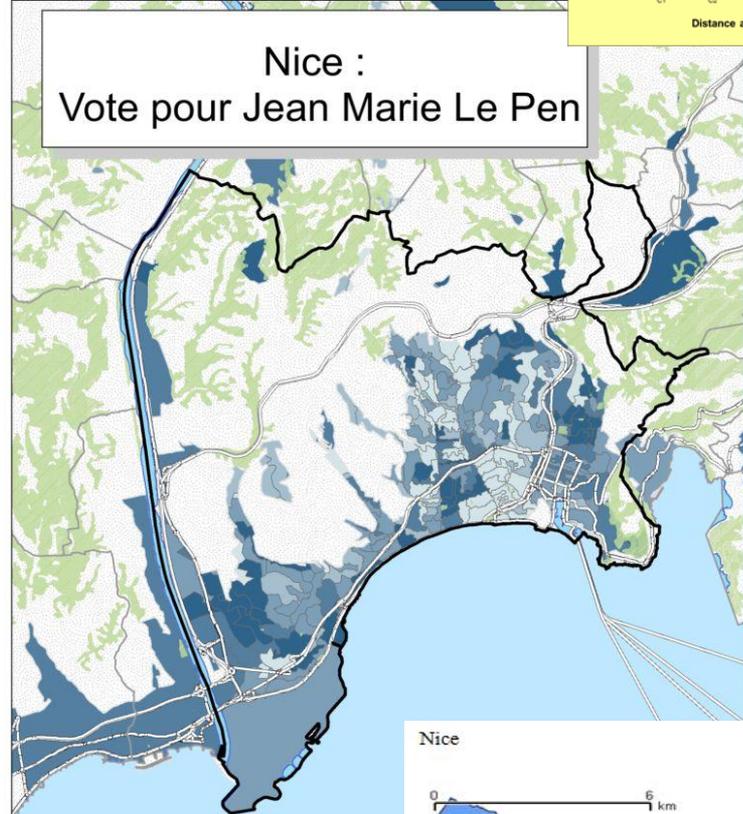
Marseille :  
Vote pour Jean Marie Le Pen



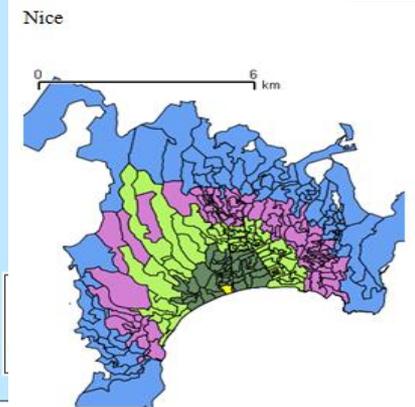
Election présidentielle 2007, premier tour



Nice :  
Vote pour Jean Marie Le Pen



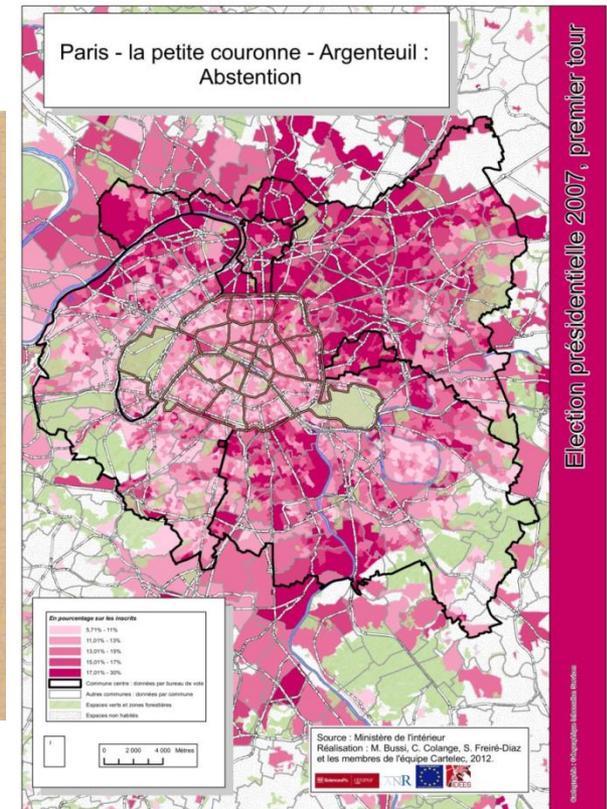
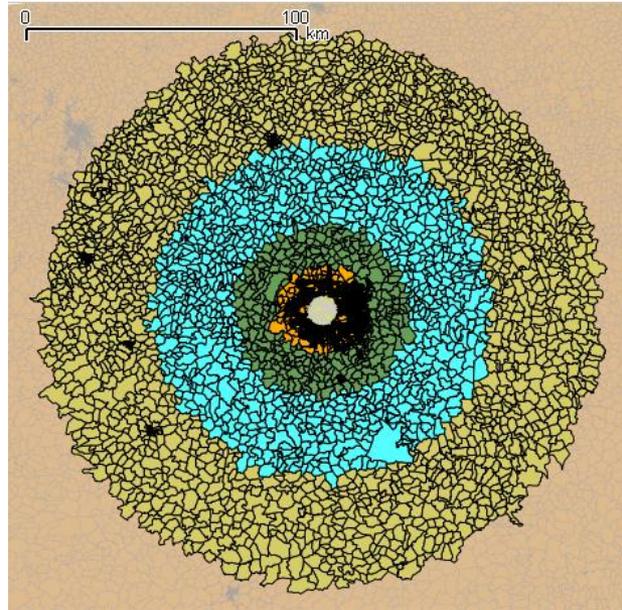
Election présidentielle 2007, premier tour



Nous avons effectué le même test sur **l'agglomération parisienne**,  
cette fois ci à partir de 5 couronnes :

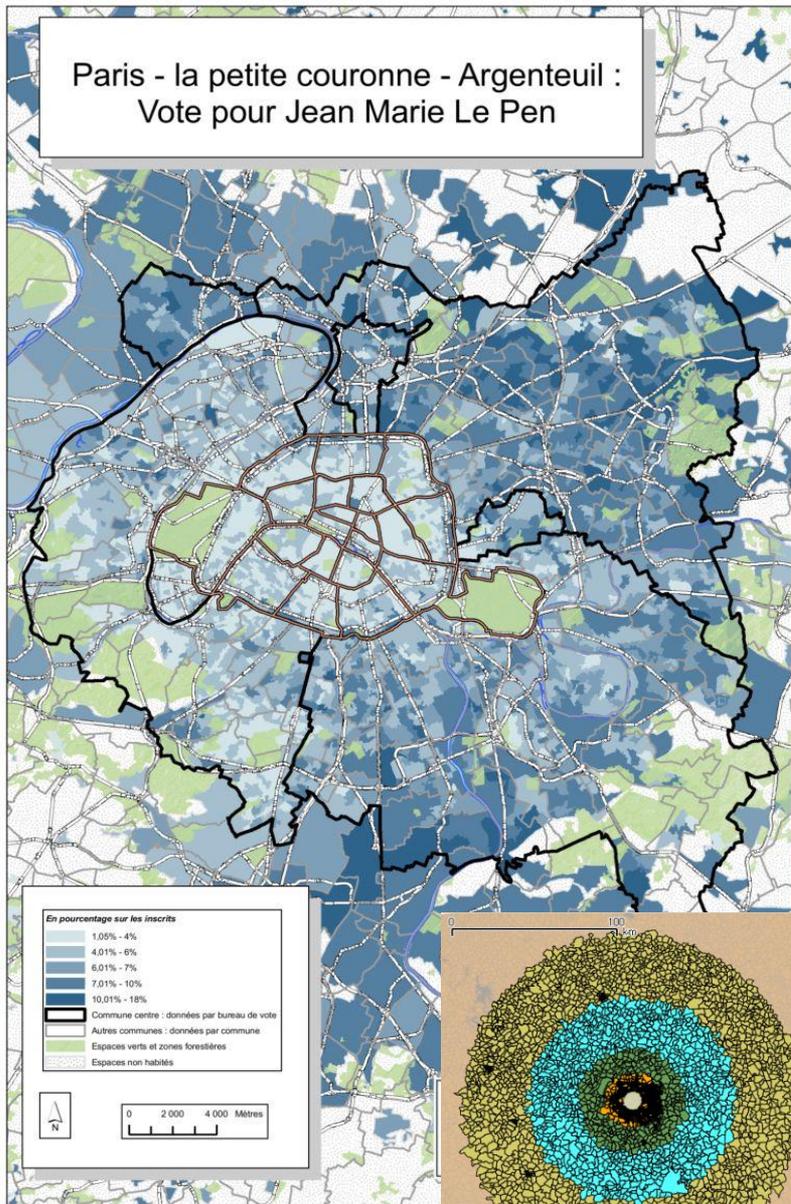
- C1, 0-5 km, 1 231 776 habitants
- C2, 5-15 km, 2 721 656 habitants
- C3, 15-30 km, 1 904 062 habitants
- C4, 30-60 km, 1 169 113 habitants
- C5, 60-100 km, 1 201 264 habitants

Total, 8 224 946 habitants



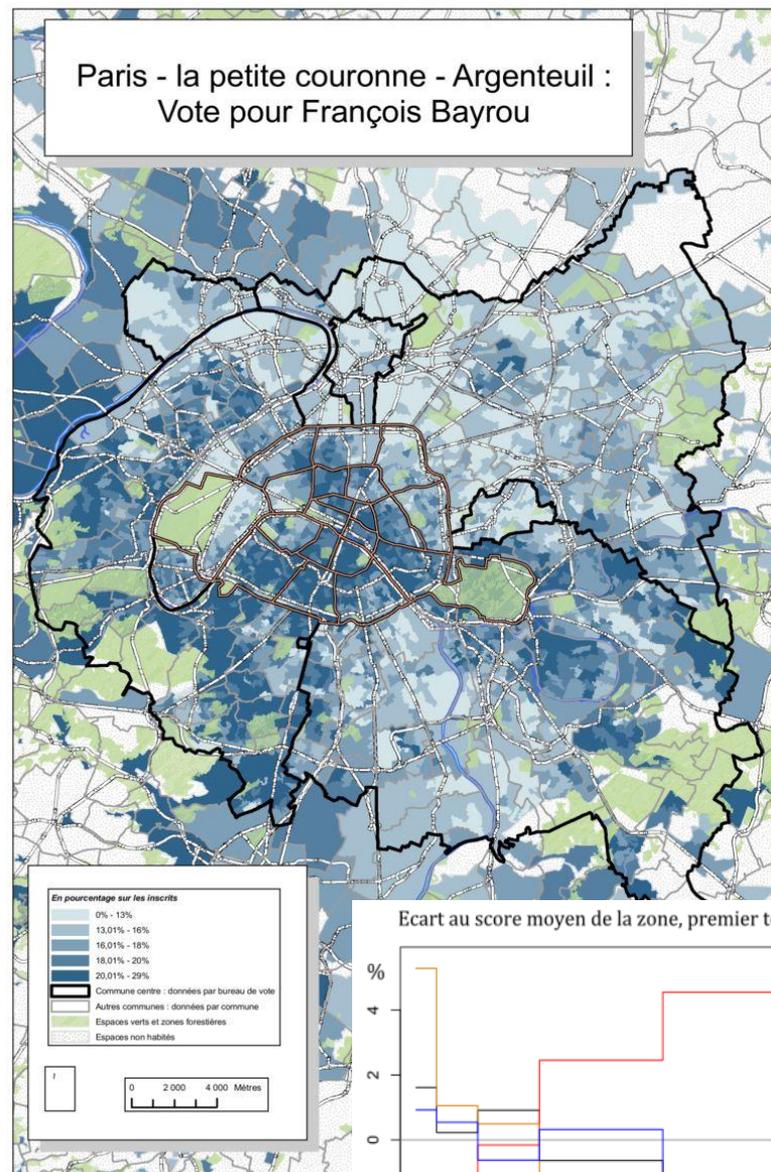
# Un effet centre périphérie très lisible...

Paris - la petite couronne - Argenteuil :  
Vote pour Jean Marie Le Pen



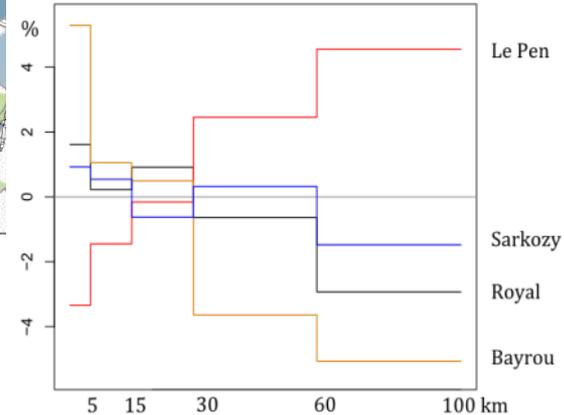
Election présidentielle 2007, premier tour

Paris - la petite couronne - Argenteuil :  
Vote pour François Bayrou

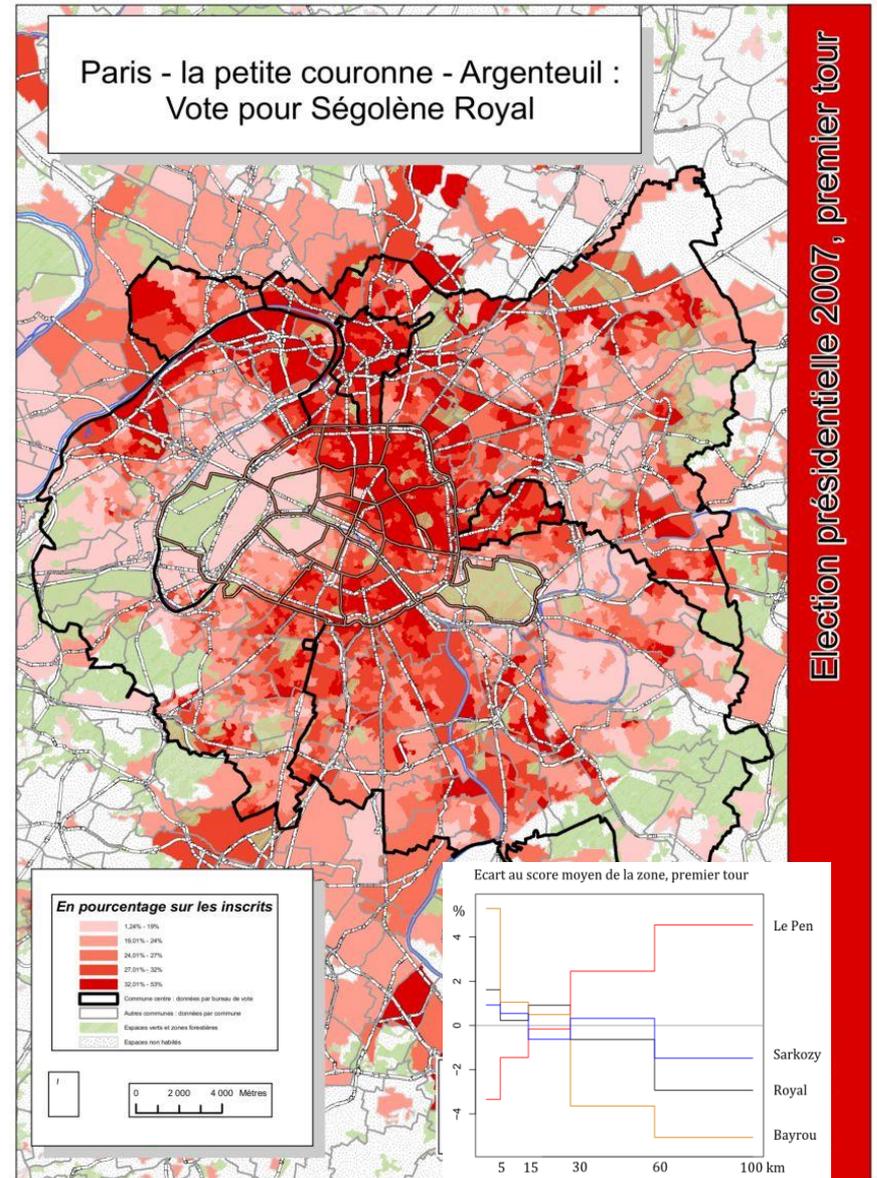
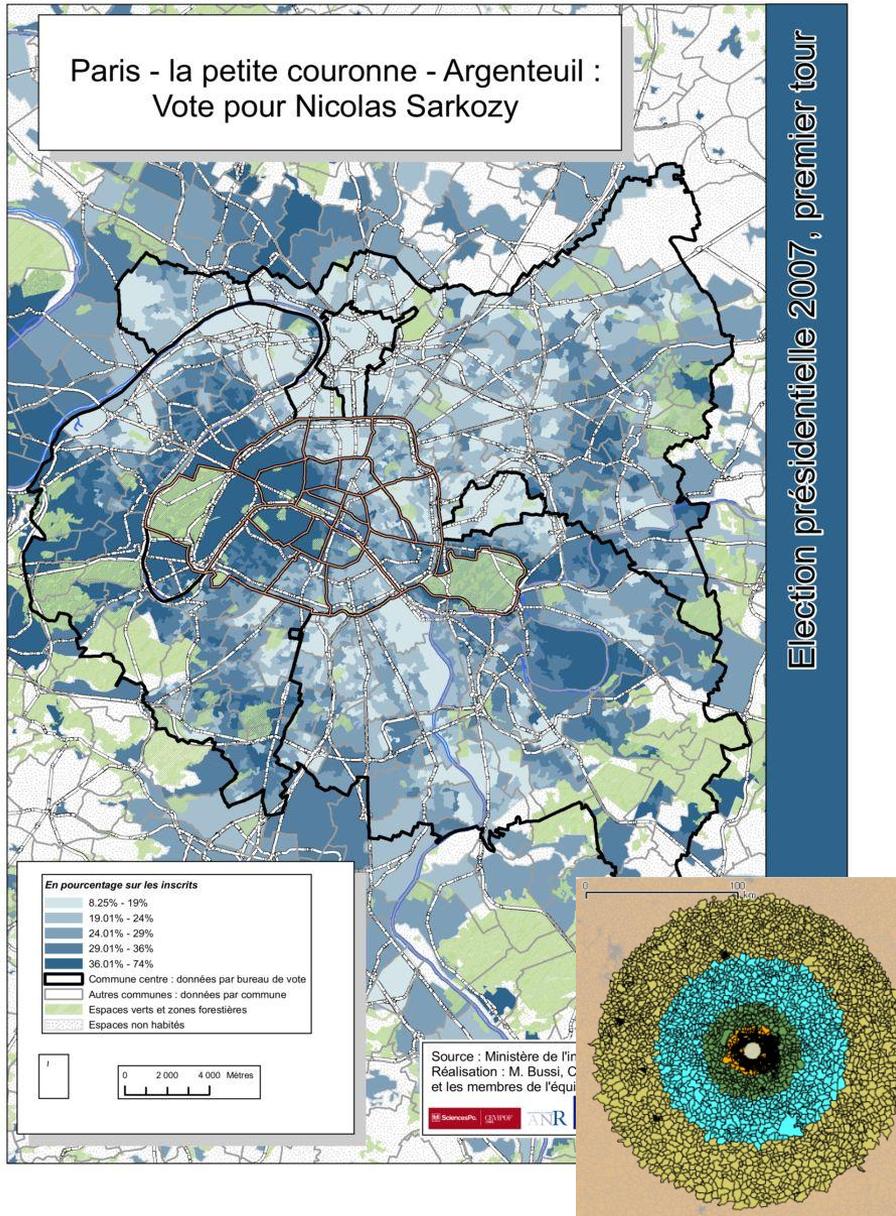


Election présidentielle 2007, premier tour

Ecart au score moyen de la zone, premier tour



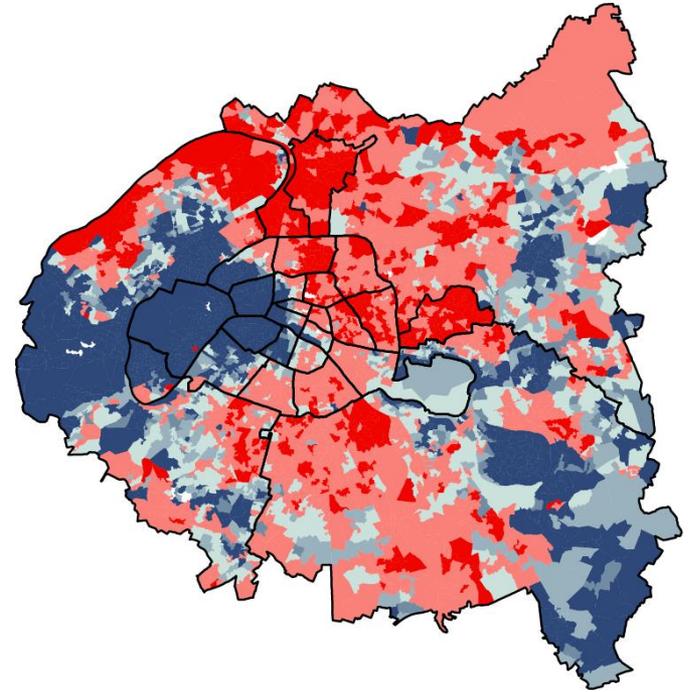
... Mais qui se double d'une forte segmentation de l'espace urbain urbain...



## ***Autres thématiques socio-politique à étudier :***

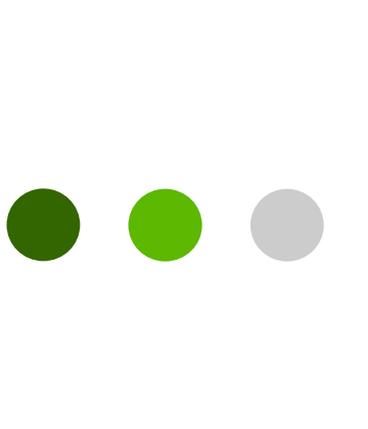
- *L'effet « klatzmann »*
- *Les théories de l'électeur habitant stratège*
- *L'effet « d'amitié locale »*
- *L'effet de « halo »*

**Merci**  
**de votre attention**



E  IPSS

**DYNAMOB**



Le dispositif pilote du  
panel ELIPSS  
et l'enquête DYNAMOB

Anne Cornilleau - Florent Gougou (Sciences Po)

Séminaire SFDS-AFSP

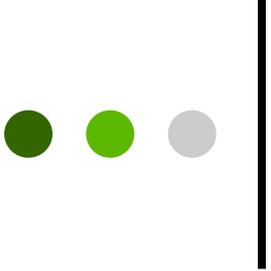
21 novembre 2013



**SciencesPo.**

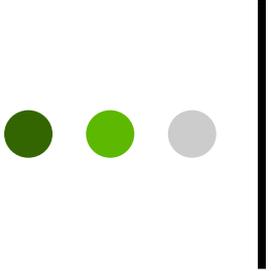


**ined**  
INSTITUT  
NATIONAL  
D'ÉTUDES  
DÉMOGRA  
PHIQUES



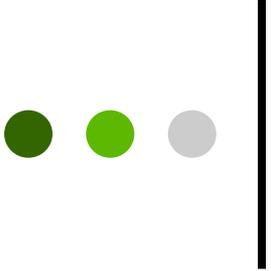
# Plan

- Présentation générale du dispositif ELIPSS
- Le recrutement du panel
- Les enquêtes
- L'enquête Dynamob



# Eléments de contexte

- EquipEx DIME-SHS
  - Dotation de 10.4 M€ pour 9 ans (2011-2020)
  - Partenaires : Sciences Po, INED, GENES, Univ. Paris Descartes, EDF R&D, Telecom ParisTech, Réseau Quetelet
  - Auto-financement à terme
- Infrastructure pour les données en sciences humaines et sociales (quanti, quali, web)
- Support pour la recherche méthodologique



## ELIPSS en quelques mots

- Panel internet en population générale (inspiré du LISS panel)
- Echantillon : tirage aléatoire dans le recensement 2011 par l'INSEE
- Accès internet 3G via tablette tactile pour chaque panéliste
- Finalité scientifique des enquêtes
- Interrogations de 30 minutes maximum par mois

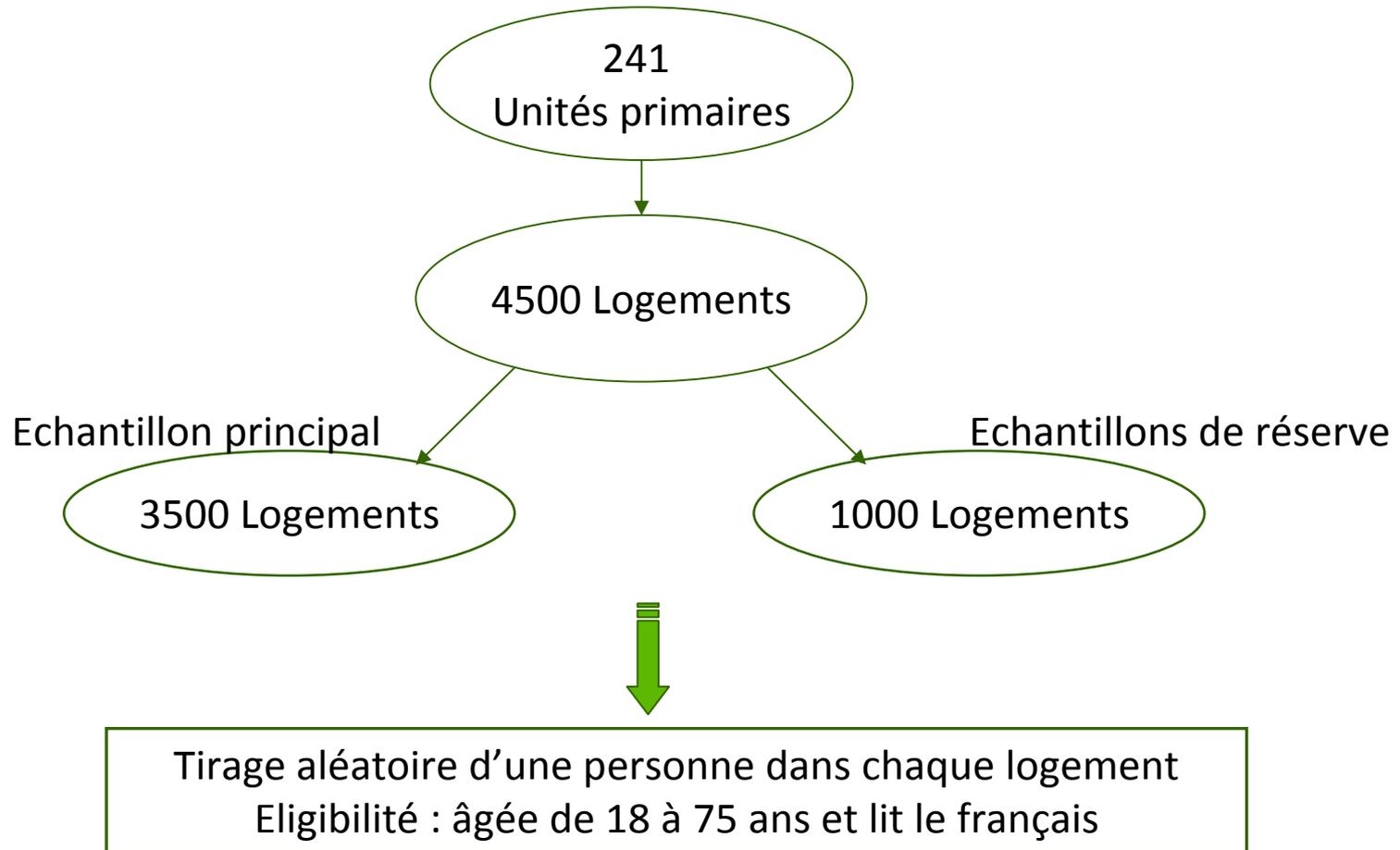
# Déroulement du recrutement du pilote

○ Pilote de 2012 à 2014 : 1 000 individus

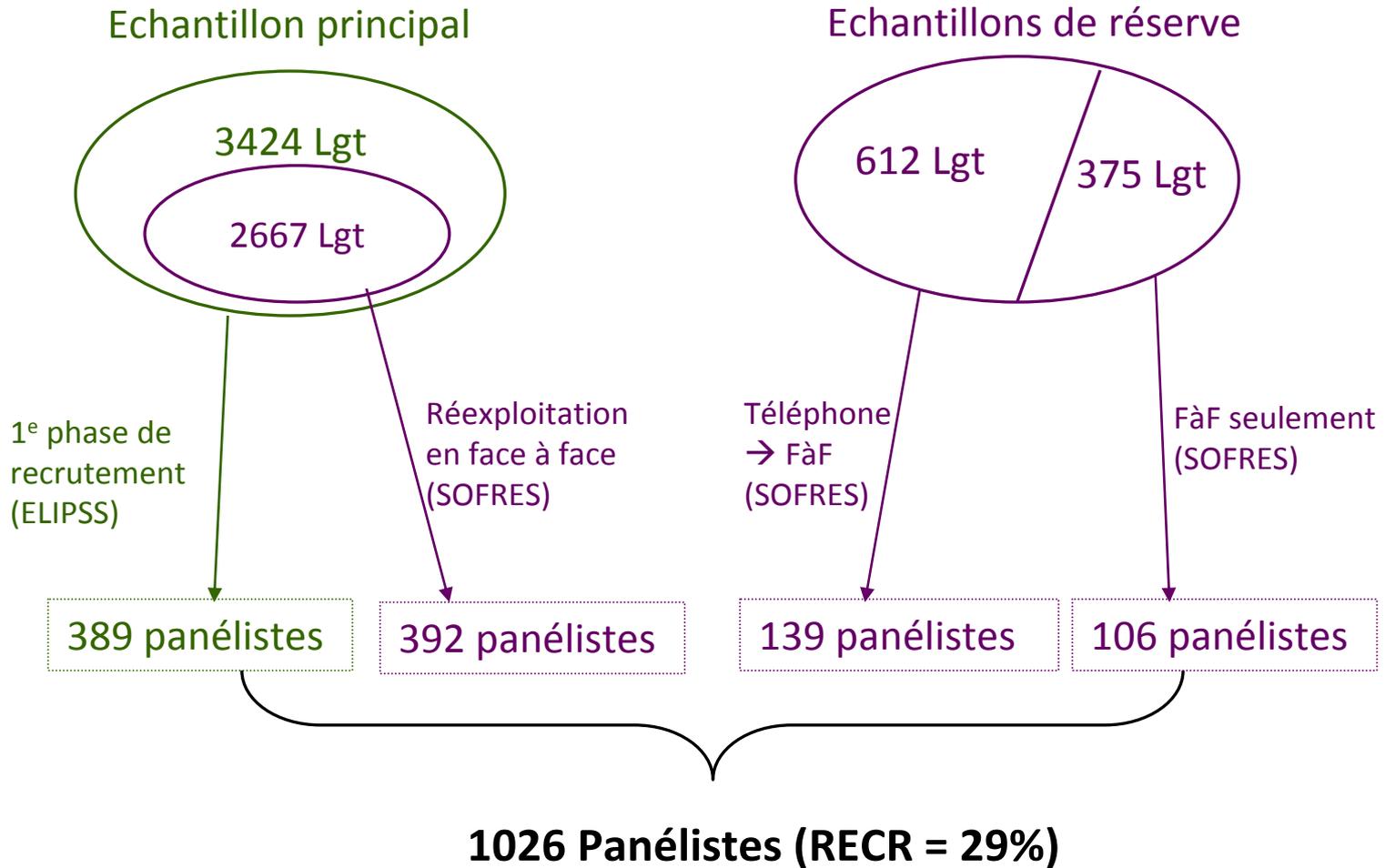
	2012							2013				
	Juin	Juillet	Août	Sept.	Octobre	Nov.	Déc.	Janvier	Février	Mars	Avril	Mai
<b>Recrutement équipe ELIPSS</b>	Invitation, relances courrier	Relances téléphoniques		Lettre annonce enquêteur								
<b>Recrutement Institut de sondage</b>					Réexploitation en face-à-face (éch. principal)							
								Recrutement éch. de réserve				
<b>Opérateur téléphonique</b>		Contractualisation			Formation téléphonique des panélistes							
<b>Enquêtes</b>							Enquête nouveaux entrants				Enquête annuelle ELIPSS	

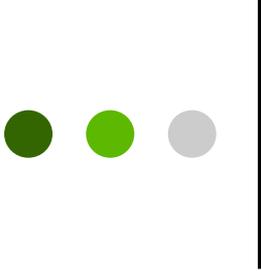
# Procédure d'échantillonnage

Stratification par région \* urbanisation



# Résultats du recrutement





# Résultats du recrutement

	Echantillon principal (N = 3424)	Echantillons de réserve (N = 987)		<b>Total (N = 4411)</b>
	ELIPSS (N = 3424) Réexp. FàF (N = 2667)	Tél >> FàF (N = 612)	FàF (N = 375)	
a. Accord ménages	30.0%	28.4%	33.1%	<b>30.1%</b>
b. Refus ménages	33.4%	38.6%	27.7%	<b>33.6%</b>
c. Non-contacts	18.8%	4.1%	10.1%	<b>17.1%</b>
d. Ménages inéligibles	17.8%	28.9%	29.1%	<b>19.2%</b>
<b>RR3 ménages a/(a+b+c)</b>	<b>36.5%</b>	<b>40.0%</b>	<b>46.6%</b>	<b>37.2%</b>
Nombre de panélistes (Accord individus/a)	781 (76.1%)	139 (80.3%)	106 (85.5%)	<b>1026 (77.6%)</b>

# Comparaison aux données INSEE

## Age

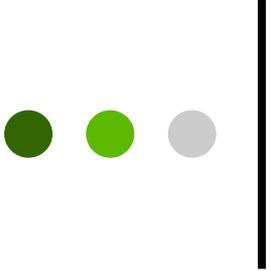
	Recensement (Jan. 2013)	ELIPSS
<b>18-24</b>	<b>12.1</b>	<b>8.8</b>
25-34	17.6	19.0
<b>35-44</b>	<b>19.1</b>	<b>24.6</b>
45-54	19.5	21.9
55-64	18.2	17.0
<b>65-75</b>	<b>13.4</b>	<b>8.7</b>

## Sexe

	Recensement (Jan. 2013)	ELIPSS
<b>Femme</b>	51.0	51.3
<b>Homme</b>	49.0	48.7

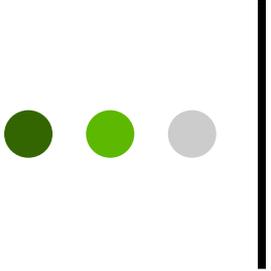
## Accès internet à la maison

	Enquête TIC 2012	ELIPSS
<b>Oui</b>	83.2	92.3
<b>Non</b>	<b>16.8</b>	<b>7.7</b>



# Sélection des enquêtes

- Projets d'équipes de recherche en SHS
  - Jusqu'en 2014 >> Accès gratuit aux partenaires de DIME-SHS
  - A partir de 2015 >> Ouvert à tous les acteurs de la recherche, participation aux frais de fonctionnement
- Appels à projets, sélection par un comité scientifique
- Longitudinales ou transversales (60 min par an max.)



# Diffusion des données

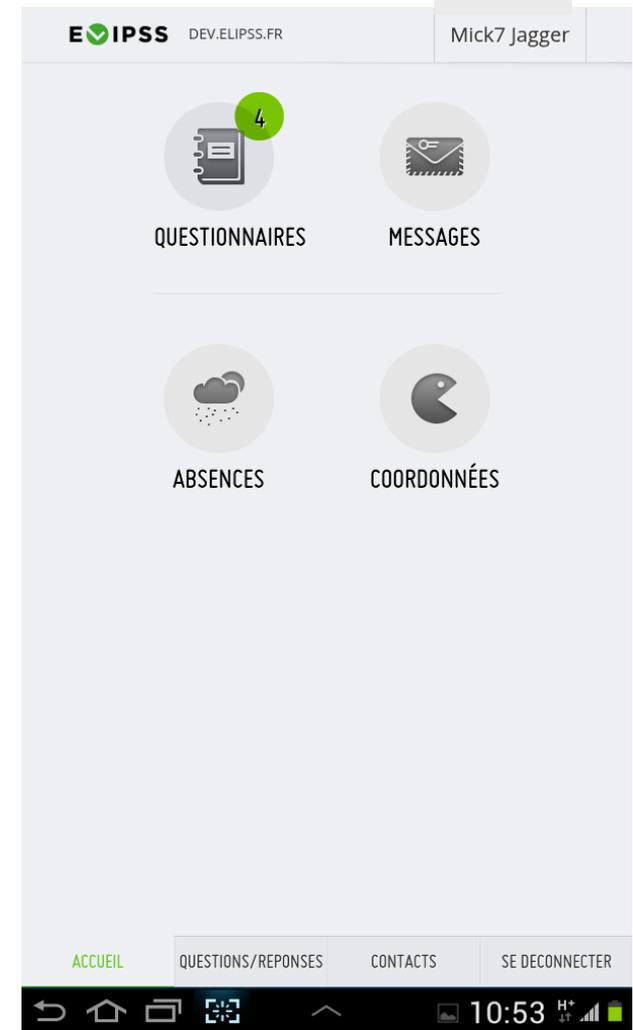
- Mise à disposition 12 mois après la clôture de l'enquête
- Appariement systématique à l'enquête annuelle  
(signalétique, travail et formation, description du ménage, liens sociaux, logement et quartier, revenus et patrimoine)
- Appariements supplémentaires validés par le comité scientifique d'ELIPSS
- Modalités de diffusion du CDSP (Portail du Réseau Quetelet)

# Equipement

## Page d'accueil de la tablette



## Application ELIPSS



# Design des questions

EVIPSS HTTP://DEV.ELIPSS.FR

PMTI

En moyenne combien de temps regardez vous la télévision au cours de la semaine, c'est à dire du lundi au vendredi ?

Moins de 2 heures par jour

De 2 à 4 heures par jour

Plus de 4 heures par jour

Je ne regarde pas la télévision

PAGE PRÉCÉDENTE PAGE SUIVANTE

EVIPSS HTTP://DEV.ELIPSS.FR

ENQUÊTES ET INTERNET

De la même manière, en ordonnant vos réponses.

**Une fois la tablette allumée, veuillez décrire, dans l'ordre, ce que vous avez fait : (trois réponses maximum)**

2  vous avez testé la connexion de la tablette en allant sur Internet

1  vous avez regardé la vidéo ELIPSS disponible sur la page d'accueil de la tablette

vous avez consulté le guide détaillé SAMSUNG disponible sur la page d'accueil de la tablette

3  vous avez essayé de vous connecter à l'application ELIPSS

vous avez regardé les applications

EVIPSS HTTP://DEV.ELIPSS.FR

ENQUÊTES ET INTERNET

Pour certaines questions, il vous sera proposé de vous positionner sur une échelle; ici, l'échelle va de 1 à 8.

Pour vous placer sur cette échelle, appuyez sur le chiffre correspondant; un curseur vert marque votre réponse.

**Avez-vous hésité avant d'accepter de participer au projet ELIPSS ?**

*Le 1 signifie ici que vous n'avez pas du tout hésité, le 8 que vous avez fortement hésité.*

1 2 3 4 5 6 7 8

← Vous n'avez pas du tout hésité à participer → Vous avez fortement hésité à participer

PAGE PRÉCÉDENTE PAGE SUIVANTE

# Exemple d'adaptation Dynamob

 IPSS HTTP://DEV.ELIPSS.FR

LES GRANDS SOJETS DE SOCIÉTÉ ET LA VIE CITOYENNE  
(79 / 195)

En France, nous avons un certain nombre de partis politiques et chacun voudrait obtenir votre vote. **Quelle est la probabilité que vous votiez un jour pour les partis suivants ?**  
Veuillez préciser votre opinion sur une échelle de 0 à 10, 0 voulant dire "pas du tout probable" et 10 voulant dire "très probable".

- Lutte Ouvrière

0	1	2	3	4	5	6	7	8	9	10
---	---	---	---	---	---	---	---	---	---	----

← →

[PAGE PRÉCÉDENTE](#) [PAGE SUIVANTE](#)

10:43 3G

# Non-réponse partielle

## Exemple 1

EIPSS HTTP://DEV.ELIPSS.FR

ENQUÊTE ANNUELLE ELIPSS

En plus de votre logement habituel, disposez-vous d'une résidence secondaire ?

Oui

Non

PAGE PRÉCÉDENTE PAGE SUIVANTE



EIPSS HTTP://DEV.ELIPSS.FR

ENQUÊTE ANNUELLE ELIPSS

La réponse à cette question est importante pour notre étude. Si les propositions de réponse ne vous conviennent pas, vous avez la possibilité de choisir "Je refuse de répondre".

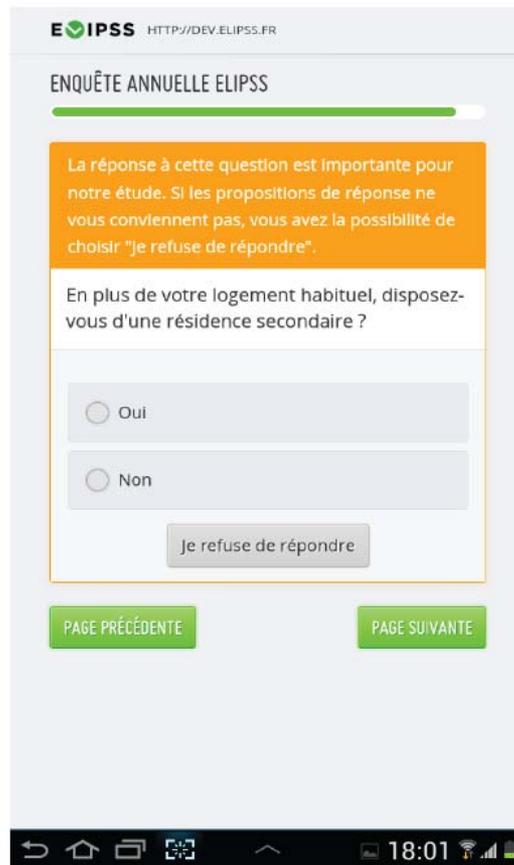
En plus de votre logement habituel, disposez-vous d'une résidence secondaire ?

Oui

Non

Je refuse de répondre

PAGE PRÉCÉDENTE PAGE SUIVANTE



## Exemple 2

EIPSS HTTP://DEV.ELIPSS.FR

ENVIRONNEMENT ET ENERGIE

(34 / 118)

Votre réponse est importante. Si vous n'avez pas d'opinion sur ce sujet, vous pouvez passer à la page suivante.

Pensez-vous que le progrès scientifique et technique résoudra, d'ici 20 ans, les problèmes d'environnement qui se posent aujourd'hui en Europe ?

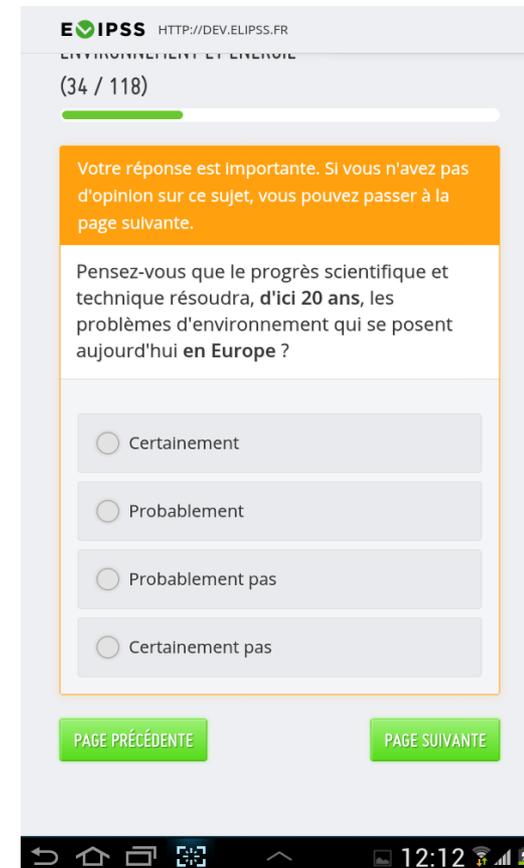
Certainement

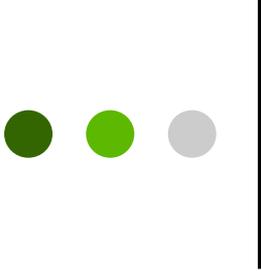
Probablement

Probablement pas

Certainement pas

PAGE PRÉCÉDENTE PAGE SUIVANTE



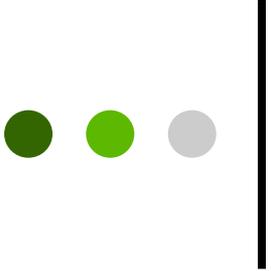


# La réponse aux enquêtes

Date	Titre de l'enquête	Tx de rép.
12/2012	Didacticiel. Internet et enquêtes	93%
02/2013	Votre quartier. Enjeux politiques	86%
04/2013	Enquête annuelle ELIPSS 1 (Vos conditions de vie)	93%
05/2013	Enquête annuelle ELIPSS 2 (Pratiques, croyances et opinions)	88%
06/2013	Pratiques culturelles, Médias et Technologies de l'Information	89%
07/2013	Fécondité, Contraception, Dysfonctions sexuelles	88%
09/2013	Dynamique de Mobilisation : Comprendre la formation des choix électoraux (V1)	86%
10/2013	Enquête sur les valeurs, l'environnement et l'énergie	83%

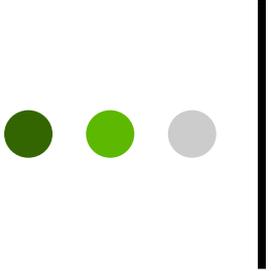
## ○ Administration des enquêtes :

- Ouverture – Clôture : 1<sup>er</sup> jeudi du mois
- Relances les vendredis en semaines 2 et 3 (voire 4)



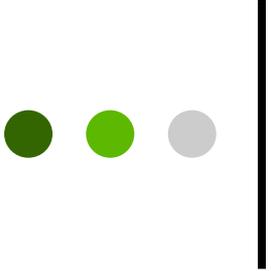
# Les spécificités de Dynamob

- Un dispositif dans le dispositif
  - Un panel politique de long terme
  - Un suivi longitudinal de septembre 2013 à juin 2017
  - Des réinterrogations régulières, au minimum 4 par an
- Un atout exceptionnel pour les études électorales
  - Un dispositif inédit en France, où les panels se sont toujours focalisés sur une seule séquence électorale
  - Un dispositif ambitieux dans une perspective internationale, où les panels électoraux de long terme n'ont jamais excédé deux vagues d'enquête par an



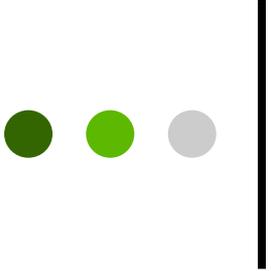
# Les axes structurants du projet

- Les logiques de formation du choix électoral
  - Le constat : la plupart des enquêtes électorales portent sur les scrutins de premier ordre, et en particulier sur la présidentielle
  - L'objectif : comprendre les différences de comportement des électeurs entre les divers types d'élections
- La politique des temps ordinaires
  - Le constat : la plupart des mesures quantitatives sur le rapport au politique sont produites à partir d'enquêtes électorales
  - L'objectif : comprendre le rapport au politique des électeurs en dehors des périodes d'élections



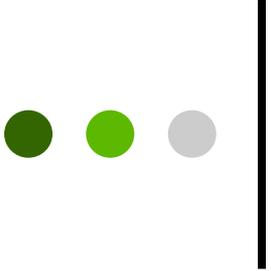
# L'architecture du projet

- Le contrat de base
  - Un volume de 50 minutes par an
  - Un questionnaire initial de 30 minutes en septembre 2013
  - Le principe directeur : l'alternance des temporalités
- Les vagues de politique ordinaire
  - L'administration du questionnaire noyau
  - Un module libre : actualité politique ou thème spécifique
- Les dispositifs électoraux
  - Une vague pré-électorale
  - Une vague post-électorale



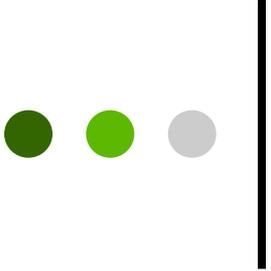
# Dynamob face à la tradition des enquêtes électorales française

- Une rationalisation des questionnaires
  - La sanctuarisation de séries longues
  - L'équilibre entre les dimensions d'analyse
  - La mise en place d'un questionnaire noyau
- Un véritable ancrage comparatif
  - La reprise d'indicateurs issus des enquêtes internationales
  - La confrontation avec les autres grandes enquêtes électorales conduites au cours de la période d'investigation
  - Le cas des élections européennes de 2014



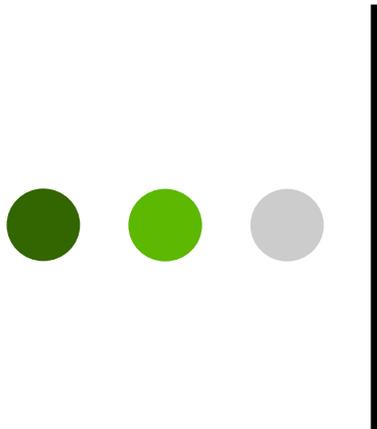
# Les défis du projet

- La date des enquêtes
  - L'adaptation au calendrier électoral
  - Les enjeux du timing de passation du questionnaire
- La répétition des indicateurs
  - Le risque d'effets de lassitude
  - Le recours à des équivalents fonctionnels
- Les synergies avec les autres enquêtes
  - Une logique d'efficacité collective
  - Un enrichissement mutuel avec l'enquête annuelle et les autres enquêtes thématiques



# Dynamob au quotidien

- Un échange permanent avec l'équipe ELIPSS
  - La discussion sur les questions
  - L'invention de nouveaux formats de questions
  - Le test des questionnaires
- Une équipe large
  - La coordination du projet : Vincent Tiberj / Florent Gougou
  - Les membres du projet : des chercheurs issus de cinq unités de recherche (CED, CEE, CEPEL, CEVIPOF, PACTE)
  - Une volonté d'ouverture



# Annexes

- [Exemples de design des questions](#)
- [Détails sur la procédure de recrutement](#)
- [Participation au panel, aux enquêtes](#)
- [Devenir membre du panel](#)

# Design de questions

**EIPSS** HTTP://DEV.ELIPSS.FR

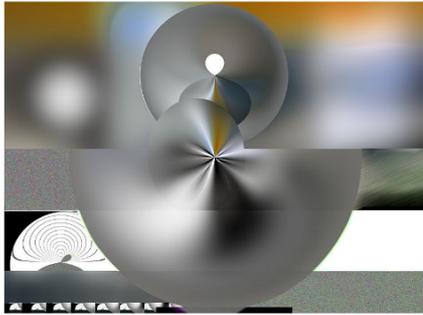
## VOTRE QUARTIER / ENJEUX POLITIQUES

Quelle importance accordez-vous aux grands thèmes suivants ?

	Beaucoup	Assez	Peu	Pas du tout
Pouvoir d'achat et revenus	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Emploi et travail	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Social	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Crise	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Environnement	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Immigration et intégration	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Justice et institutions	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Questions de société	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Union Européenne	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

[PAGE PRÉCÉDENTE](#) [PAGE SUIVANTE](#)

**EIPSS** HTTP://DEV.ELIPSS.FR



Diriez-vous que l'image ci-dessus vous plaît ...

- Beaucoup
- Plutôt
- Plutôt pas
- Pas du tout
- Je n'ai pas d'opinion

# Design de questions



## Les info-bulles

The screenshot shows a mobile app interface for 'EIPSS' with the URL 'HTTP://DEV.ELIPSS.FR'. The page title is 'PMTI'. The main text asks: 'Au cours des 12 derniers mois, combien de fois environ avez-vous regardé des films en VOD, streaming ou téléchargés en cours des 12 derniers mois'. Below the text are five radio button options: 'Jamais', '1 à 3 fois', '4 à 6 fois', '7 à 12 fois', and 'Plus de 12 fois'. A green 'PAGE PRÉCÉDENTE' button is at the bottom left. An info-bubble overlay is positioned over the question text. The bubble has a title 'VOD' and contains the text: 'La VOD (Vidéo à la Demande) permet de commander et de regarder de la vidéo en temps réel sur Internet'. At the bottom of the bubble is a green button labeled 'FERMER CETTE FENÊTRE'. The bottom of the screen shows an Android navigation bar with the time 17:30.

## Les enregistrements vocaux

The screenshot shows a mobile app interface for 'EIPSS' with the URL 'HTTP://DEV.ELIPSS.FR'. The page title is 'ENVIRONNEMENT ET ENERGIE (36 / 118)'. The main text asks: 'Pourquoi pensez-vous que le progrès scientifique et technique résoudra, dans les prochaines années, les problèmes d'environnement qui posent aujourd'hui en Europe? Merci d'enregistrer votre réponse.' Below the text is an orange 'M'ENREGISTRER' button and a green 'PAGE PRÉCÉDENTE' button. A voice recording overlay is positioned over the question text. The overlay has a title 'Enregistrez votre message' and a large digital timer showing '00:00'. Below the timer is a progress bar and a set of playback controls (stop, play, pause). The bottom of the screen shows an Android navigation bar with the time 12:16.

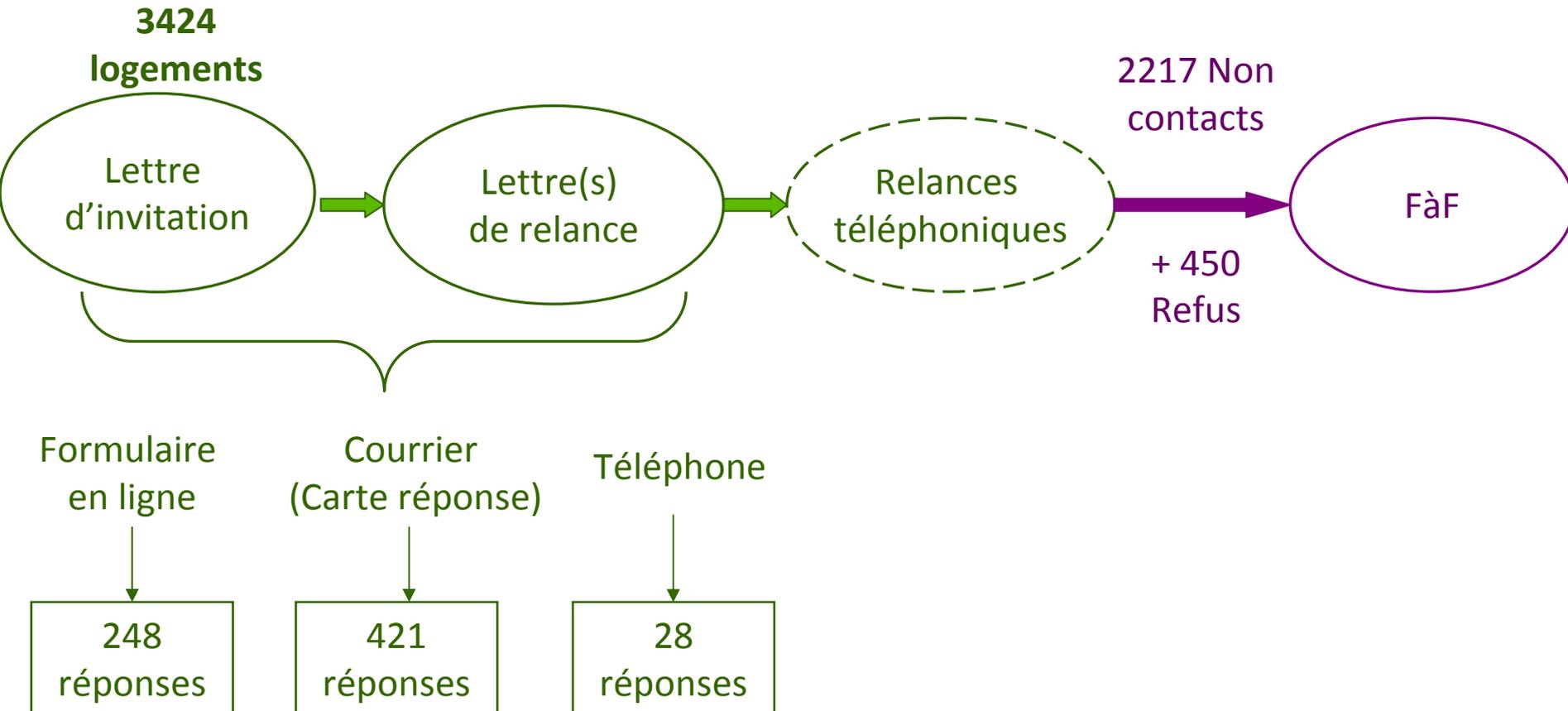
# Recrutement ech. principal

Recrutement équipe ELIPSS

Juin 2012 – Août 2012

Institut de sondage

Oct. 2012 – Fév. 2013

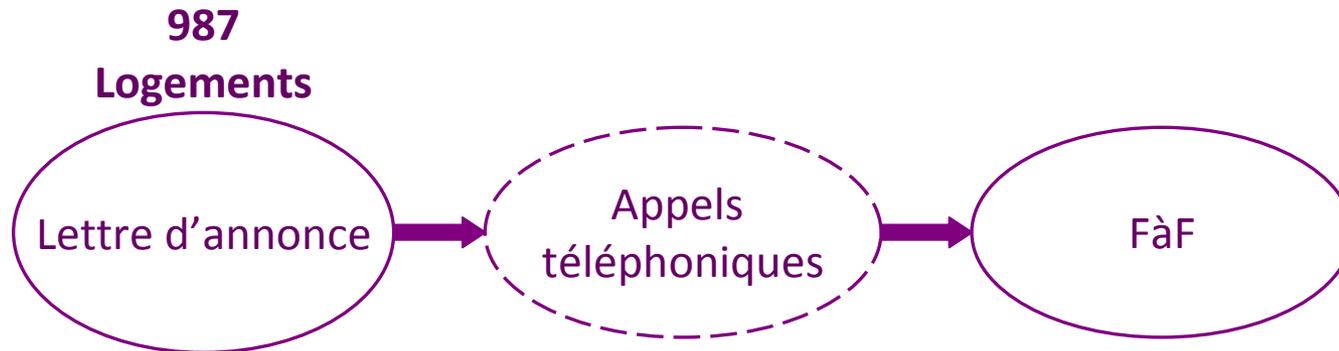


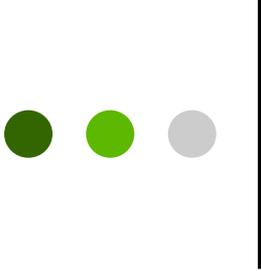
# Recrutement ech. réserve



Institut de sondage

Janvier – Avril 2013





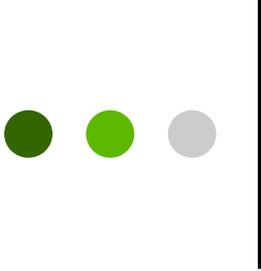
# La participation à ELIPSS

## Participation à une enquête avant l'entrée dans ELIPSS

Non, n'a jamais été sollicité (n = 302)	32.7 %
Non, a toujours refusé (n = 125)	<b>13.5 %</b>
Oui (n = 497)	53.8 %
Total (n = 924)	100.0

## Principales raisons de participation à ELIPSS (2 réponses poss.)

Prêt de la tablette (n = 571)	<b>61.8 %</b>
Confiance dans les institutions ELIPSS (n = 428)	46.3 %
Originalité du projet (n = 337)	36.5 %
Intérêt pour la recherche (n = 297)	32.1 %
Accès internet gratuit (n = 123)	13.3 %



# Devenir membre du panel



## ○ Convention de participation

	Panéliste	ELIPSS
Droits	<ul style="list-style-type: none"><li>○ Utilisation personnelle de la tablette dans le respect de la loi</li><li>○ Libre de quitter le panel n'importe quand</li><li>○ Droit d'accès et de rectification des données personnelles</li></ul>	<ul style="list-style-type: none"><li>○ Exclure en cas de non-réponse répétée ou d'utilisation illégale de la tablette (de l'internet)</li></ul>
Obligations	<ul style="list-style-type: none"><li>○ Participer régulièrement aux enquêtes</li><li>○ Prendre soin de la tablette et informer ELIPSS en cas de vol ou de casse</li><li>○ Rendre la tablette lors de la sortie du panel</li></ul>	<ul style="list-style-type: none"><li>○ Exclure les enquêtes à fins commerciales</li><li>○ Fournir la tablette et l'accès à internet et prendre en charge les frais de livraison de la tablette</li><li>○ Garantir la confidentialité et l'anonymat des panélistes</li></ul>



**ined**  
INSTITUT  
NATIONAL  
D'ÉTUDES  
DÉMOGRA  
PHIQUES



# L'apport de l'analyse multiniveau pour les données électorales

Valérie Golaz

INED

UMR 196 CEPED



**ined**  
INSTITUT  
NATIONAL  
D'ÉTUDES  
DÉMOGRA  
PHIQUES



# Plan de la présentation

- Qu'est-ce que l'analyse multiniveau ?
- Application aux données électorales





**ined**  
INSTITUT  
NATIONAL  
D'ÉTUDES  
DÉMOGRA  
PHIQUES



## Usage

---

### Traiter une information hiérarchique :

---

Les unités statistiques au sein d'un groupe présentent des similarités, il y a une corrélation interne au groupe.

- L'hypothèse d'indépendance des observations est violée
- La variance de l'information n'est pas identique à chaque degré d'agrégation



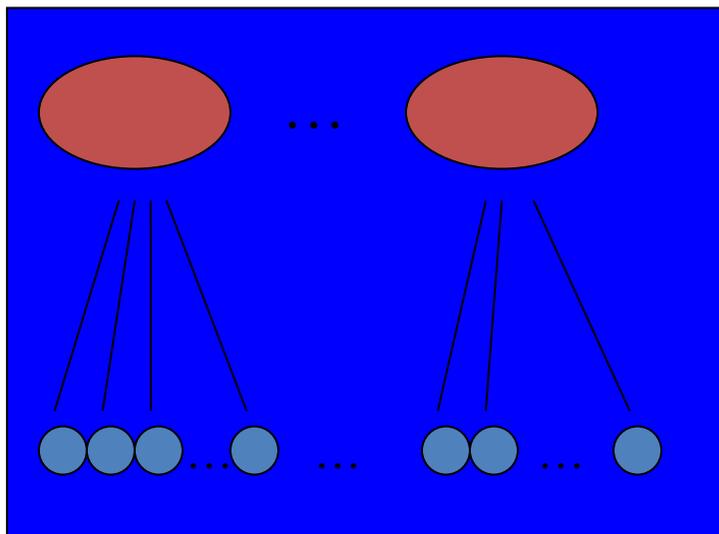
**ined**  
INSTITUT  
NATIONAL  
D'ÉTUDES  
DÉMOGRA  
PHIQUES

# Définir le(s) niveau(x) d'agrégation à utiliser...

## Quelques exemples courants



Zones administratives



Individus

- Micro / Macro**
- Individu / Famille
- Événement / Trajectoire
- ...
- Patients / Hôpital
- Enfants / École
- Employé/Entreprise
- Électeurs / Circonscription
- ...

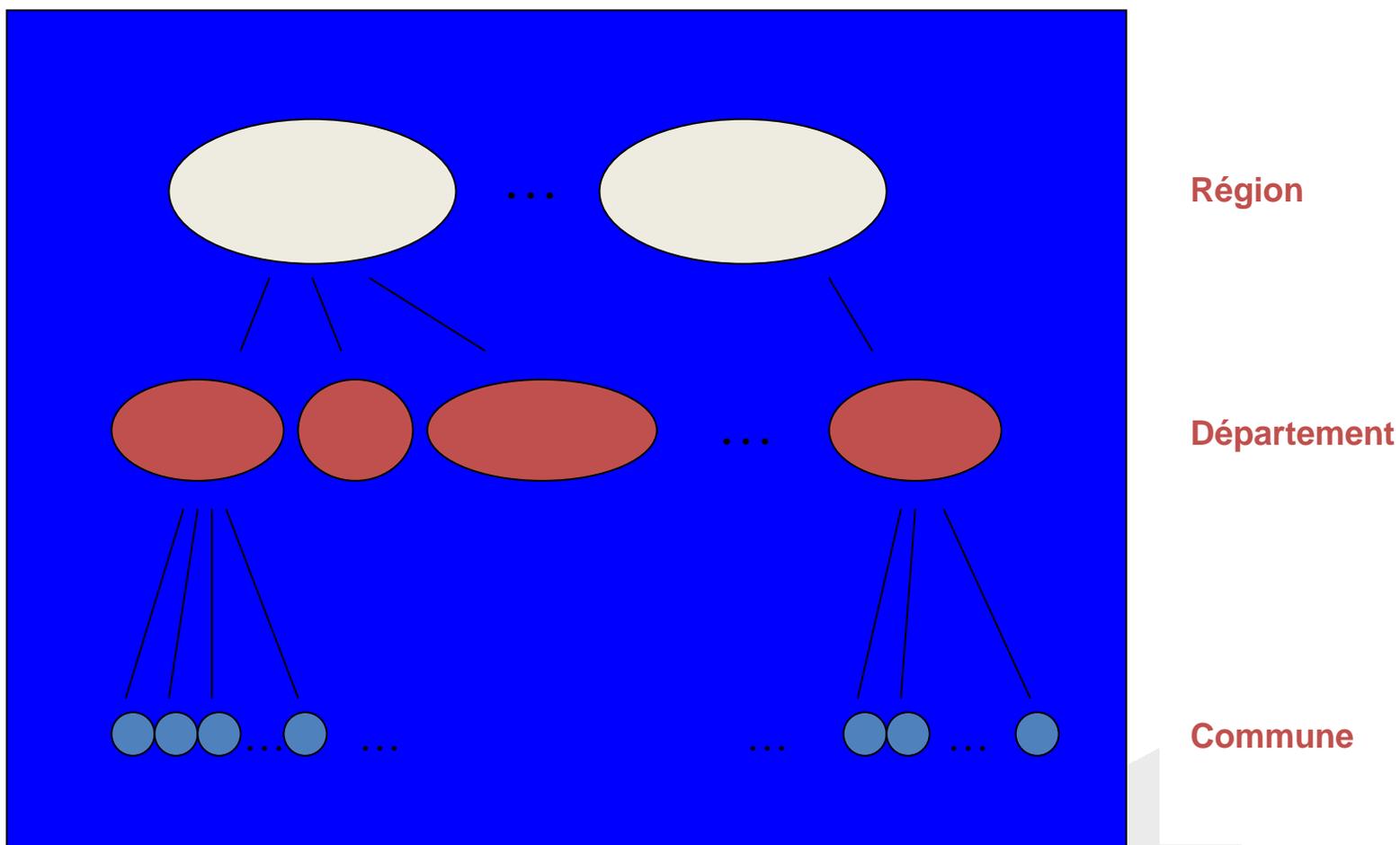


**ined**  
INSTITUT  
NATIONAL  
D'ÉTUDES  
DÉMOGRA  
PHIQUES

**CePeD**

Centre  
Population  
& Développement

# Données électorales



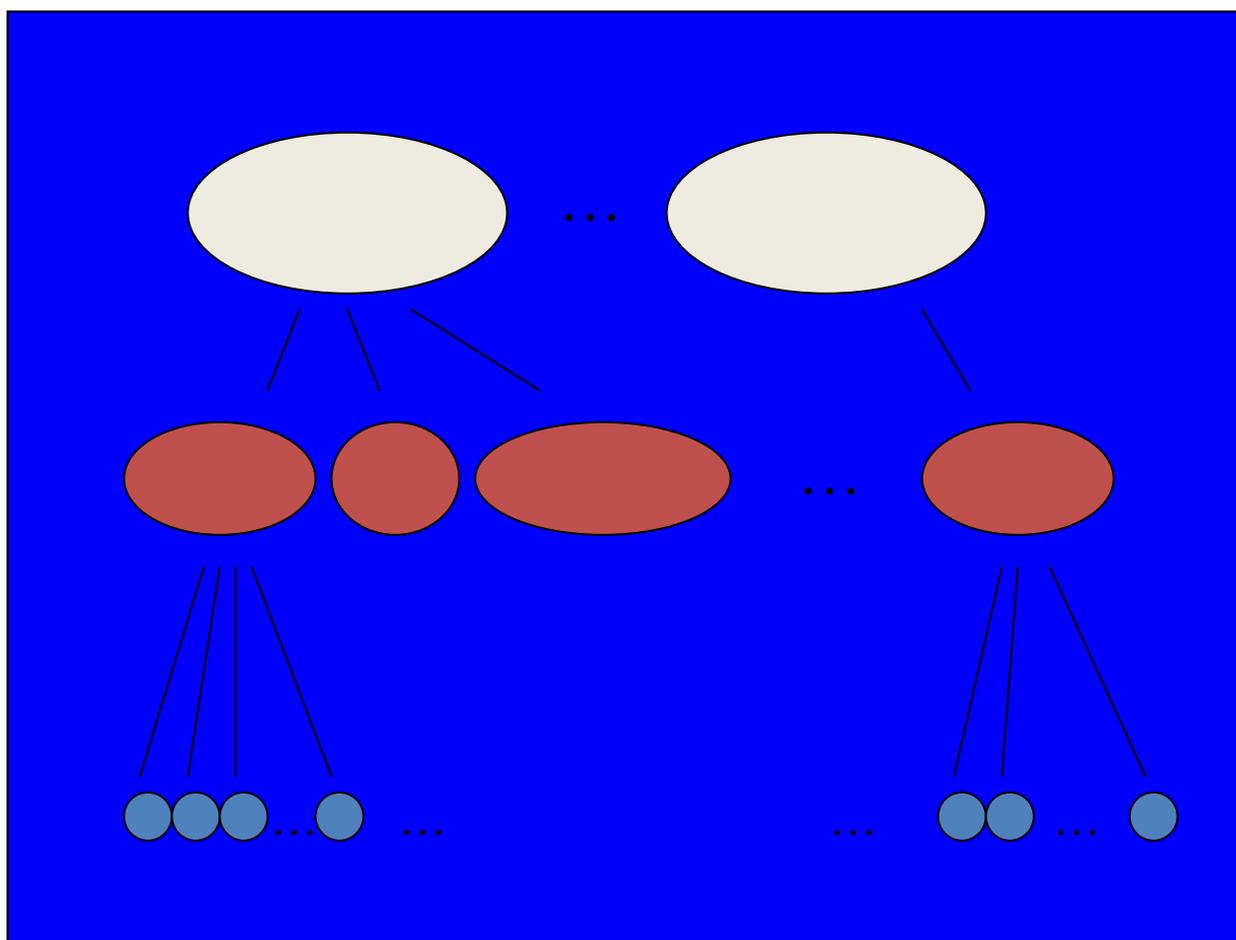


**ined**  
INSTITUT  
NATIONAL  
D'ÉTUDES  
DÉMOGRA  
PHIQUES

**CePeD**

Centre  
Population  
& Développement

# Données électorales



Département

Commune

Bureau de vote

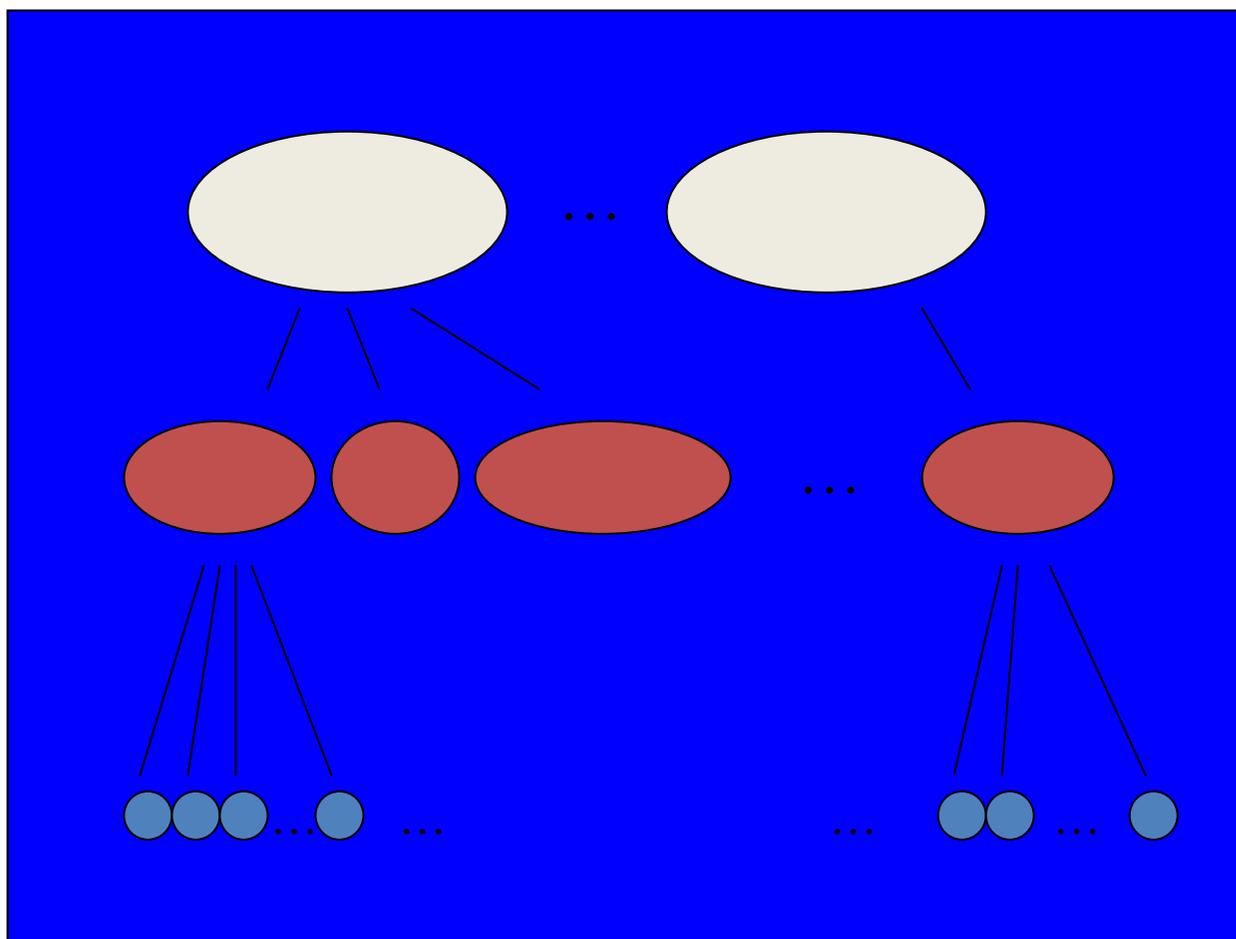


**ined**  
INSTITUT  
NATIONAL  
D'ÉTUDES  
DÉMOGRA  
PHIQUES

**CePeD**

Centre  
Population  
& Développement

# Données électorales



Commune

Bureau de vote

Individu

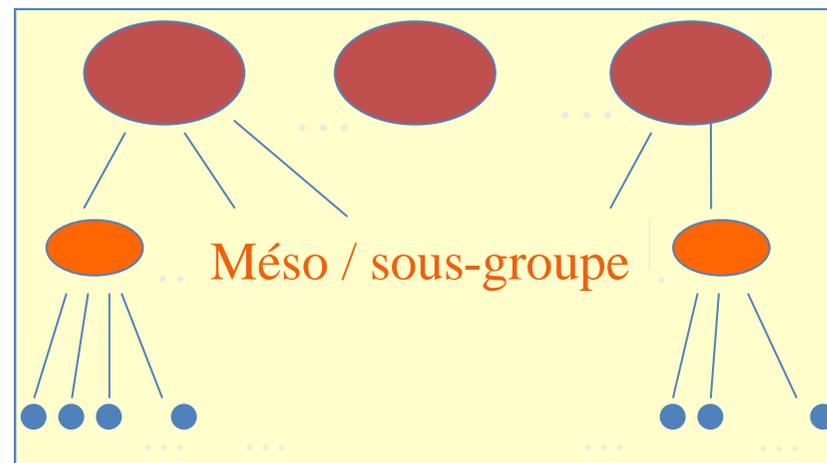
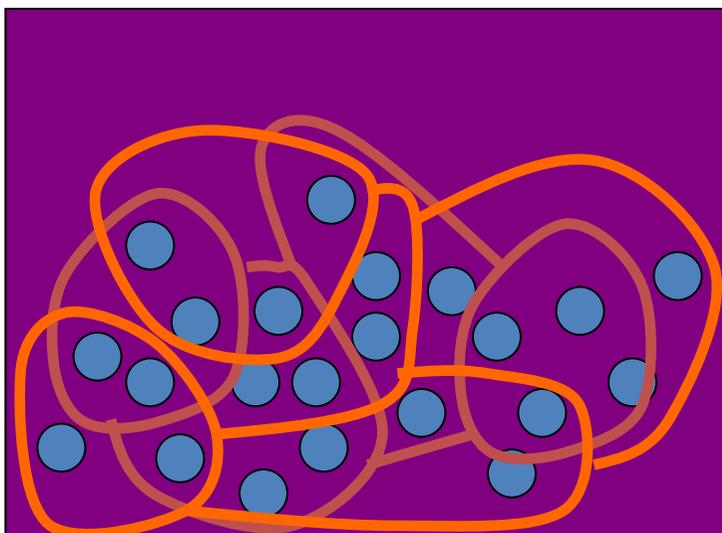


**ined**  
INSTITUT  
NATIONAL  
D'ÉTUDES  
DÉMOGRA  
PHIQUES

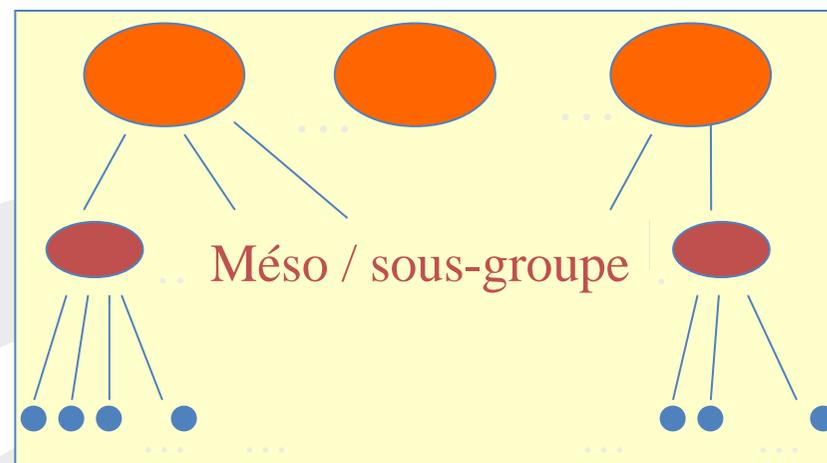


# Définir les niveaux d'agrégation à utiliser...

Différents types de contextes  
peuvent être importants



Entrelacement...





**ined**  
INSTITUT  
NATIONAL  
D'ÉTUDES  
DÉMOGRA  
PHIQUES



# Principes

Mener de front les analyses macro et micro, au sein d'un modèle statistique unique.

=> Mesurer les effets relatifs des caractéristiques individuelles et contextuelles

L'analyse contextuelle (non-multiniveau) est 'optimiste' (tendance à la surévaluation de l'effet des variables contextuelles)

=> Evaluer l'hétérogénéité intra-groupe et inter-groupe des données

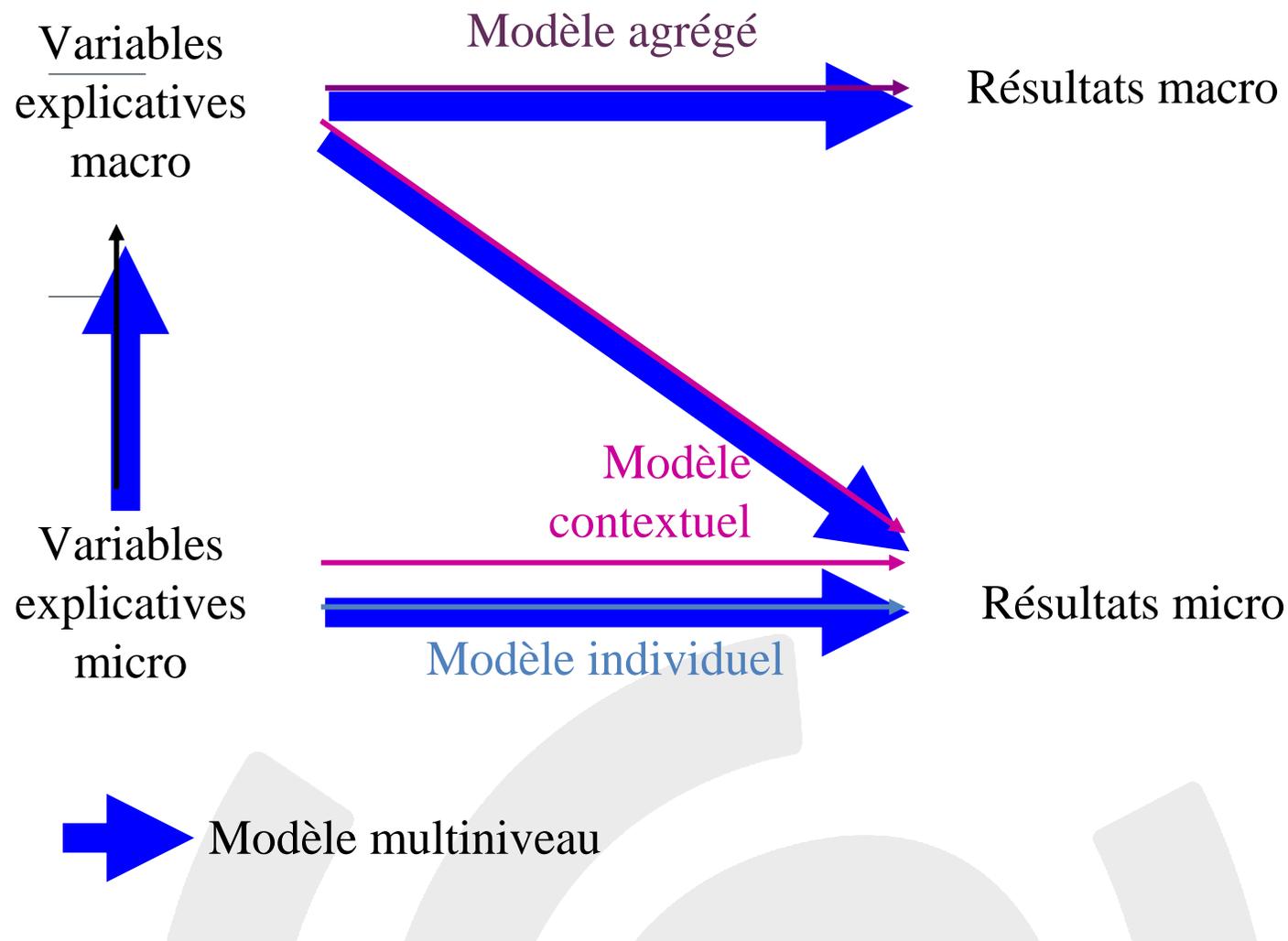
=> Evaluer la part non expliquée du modèle.



**ined**  
INSTITUT  
NATIONAL  
D'ÉTUDES  
DÉMOGRA  
PHIQUES



# Modèles à différents niveaux





**ined**  
INSTITUT  
NATIONAL  
D'ÉTUDES  
DÉMOGRA  
PHIQUES



# L'apport du multiniveau

Structurer les données sur différents niveaux permet une décomposition simple en effets fixes et aléatoires

Dans un modèle statistique, il y a toujours une part d'inobservé, une partie de la réalité qui n'est pas explicitée

En dissociant dans un modèle les caractéristiques à différents niveaux d'observation, on peut percevoir de manière plus fine cette « **hétérogénéité non observée** » : on obtient une mesure de la variance par niveau.



**ined**  
INSTITUT  
NATIONAL  
D'ÉTUDES  
DÉMOGRA  
PHIQUES



## Un objectif et trois modes d'action...

Objectif : (toujours le même) évaluer l'effet des variables explicatives et réduire la part d'hétérogénéité non observée dans le modèle, en particulier celle qui se trouve entre individus, dans les groupes

- En jouant sur la définition des groupes
- En jouant sur la modélisation statistique des effets fixes et aléatoires
- En cherchant les variables explicatives pertinentes





**ined**  
INSTITUT  
NATIONAL  
D'ÉTUDES  
DÉMOGRA  
PHIQUES



# Spécificités de l'analyse multiniveau

## Modèle linéaire

Variables au niveau 1 (V) et au niveau 2 (C)

$$Y_{ij} = \alpha_0 + u_{0j} + e_{0ij} \\ + \alpha_1 V_{1ij} + \alpha_2 V_{2ij} + \dots + \alpha_n V_{nij} \quad (\text{niveau 1}) \\ + \beta_1 C_{1j} + \beta_2 C_{2j} + \dots \quad (\text{niveau 2}) \\ + \gamma_1 V_{1ij} C_{1j} + \dots \quad (\text{termes d'interaction})$$

$$\text{avec } \text{Var}(e_{0ij}) = \sigma^2_{e_0}$$

$$\text{et } \text{Var}(u_{0j}) = \sigma^2_{u_0}$$



**ined**  
INSTITUT  
NATIONAL  
D'ÉTUDES  
DÉMOGRA  
PHIQUES



# Spécificités de l'analyse multiniveau

## Modèle logit

Variables au niveau 1 (V) et au niveau 2 (C)

$$P(Y_{ij} = 1) = (1 + \exp[-[\alpha_0 + u_{0j} + e_{0ij} \\ + \alpha_1 V_{1ij} + \alpha_2 V_{2ij} + \dots + \alpha_n V_{nij} \\ + \beta_1 C_{1j} + \beta_2 C_{2j} + \dots \\ + \gamma_1 V_{1ij} C_{1j} + \dots]])^{-1}$$

avec  $\text{Var}(e_{0ij}) = \sigma^2_{e_0}$

et  $\text{Var}(u_{0j}) = \sigma^2_{u_0}$



**ined**  
INSTITUT  
NATIONAL  
D'ÉTUDES  
DÉMOGRA  
PHIQUES



## A quoi cela sert-il ?

Quelle part de l'hétérogénéité se place au niveau 1 et au niveau 2 ?

- entre individus au sein de chaque groupe
- ou entre groupes ?

⇒ Identifier le niveau auquel la variance doit être réduite

Toutes choses égales par ailleurs, qu'est ce qui va faire la différence entre les **individus**, d'autres caractéristiques individuelles ou bien des caractéristiques partagées avec leurs pairs au sein des groupes définis ?



**ined**  
INSTITUT  
NATIONAL  
D'ÉTUDES  
DÉMOGRA  
PHIQUES



# Et à quoi encore ?

Toutes choses égales par ailleurs, quelles caractéristiques contextuelles vont expliquer les différences entre groupes ?

Quels sont les « contextes » pertinents pour l'analyse ?



**ined**  
INSTITUT  
NATIONAL  
D'ÉTUDES  
DÉMOGRA  
PHIQUES



# Mais quel(s) contexte(s) prendre en compte ?

## Différents types de variables contextuelles :

- **Endogènes** : issues des micro-données, effet de composition de la population étudiée
- **Exogènes** : issues d'une autre source

La prise en charge de sources de données différentes est l'un des intérêts de l'approche multiniveau



**ined**  
INSTITUT  
NATIONAL  
D'ÉTUDES  
DÉMOGRA  
PHIQUES



## Résultats

### L'analyse multiniveau permet :

- De guider la modélisation des effets fixes
- D'éviter des estimations trop optimistes des effets mesurés à des niveaux supérieurs d'agrégation (analyse contextuelle)
- D'apparier plusieurs sources d'information différentes
- De tirer parti d'une information rare au sein des sous-groupes





**ined**  
INSTITUT  
NATIONAL  
D'ÉTUDES  
DÉMOGRA  
PHIQUES



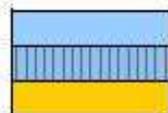
# Application aux données électorales



SciencesPo. | Centre de données socio-politiques

## ELECTIONS PRESIDENTIELLES

Date	Tour	Commune*	Canton	Circonscription	Département	Région
1965	1					
	2					
1969	1					
	2					
1974	1					
	2					
1981	1					
	2					
1988	1					
	2					
1995	1					
	2					
2002	1					
	2					
2007	1					
	2					



Consultation en ligne

<http://cdsp.sciences-po.fr/AfficheElec.php>

[http://cdsp.sciences-po.fr/Vil\\_AfficheElec.php](http://cdsp.sciences-po.fr/Vil_AfficheElec.php)

Demande auprès du CDSP

[via le portail du Réseau Quetelet](#)

\* Les résultats sur demande au CDSP ne sont disponibles que pour les communes de plus de 9 000 habitants.

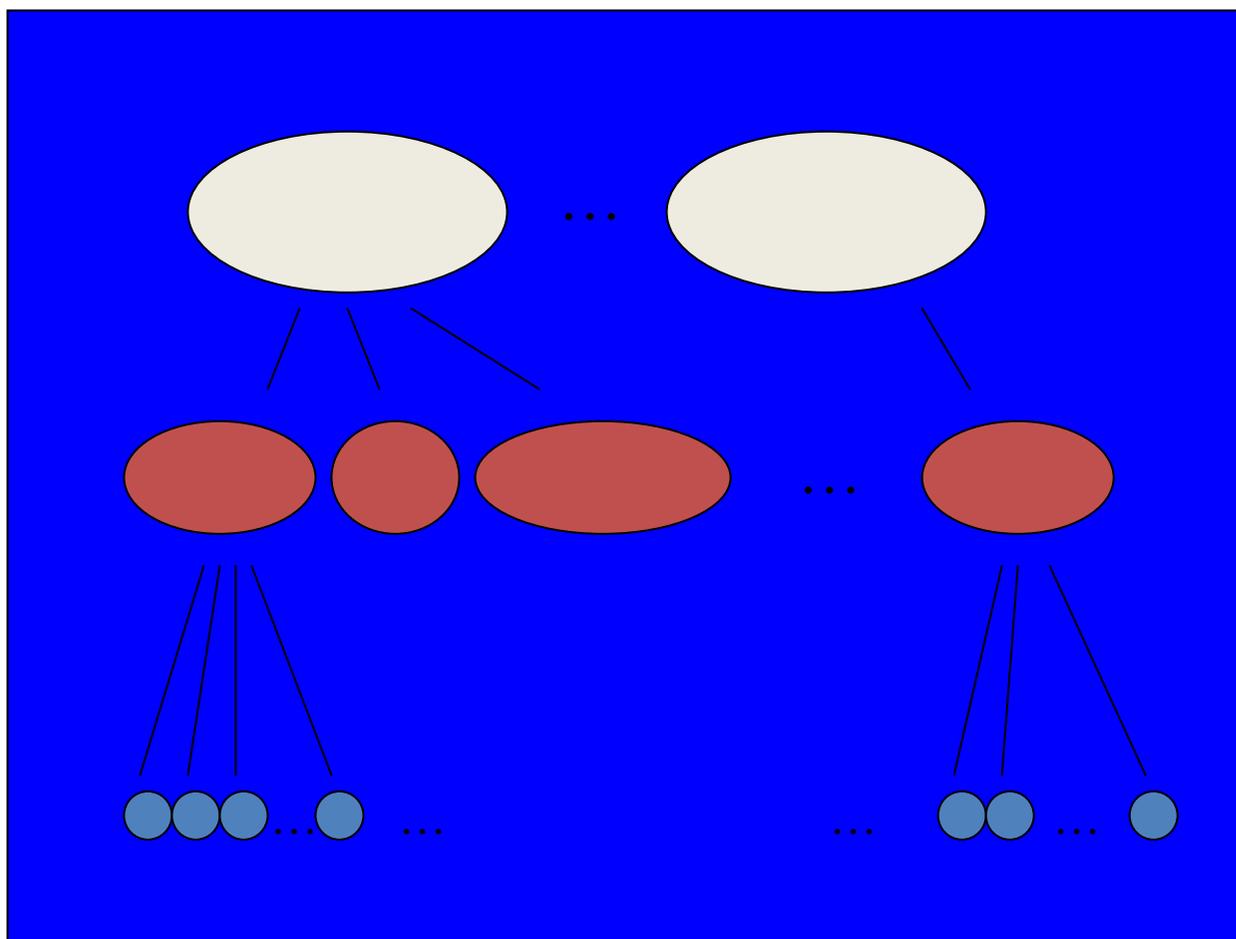


**ined**  
INSTITUT  
NATIONAL  
D'ÉTUDES  
DÉMOGRA  
PHIQUES

**CePeD**

Centre  
Population  
& Développement

# Données électorales publiées



Région

Département

Commune



**ined**  
INSTITUT  
NATIONAL  
D'ÉTUDES  
DÉMOGRA  
PHIQUES



## Précautions spécifique

### Attention à l'erreur écologique:

L'utilisation de variables contextuelles endogènes, liées à la composition de la population des unités 1, peut donner la fausse impression que l'on travaille sur des caractéristiques individuelles

Pour des analyses incluant le niveau individu, nécessité d'utiliser des données d'enquêtes ou de sondages

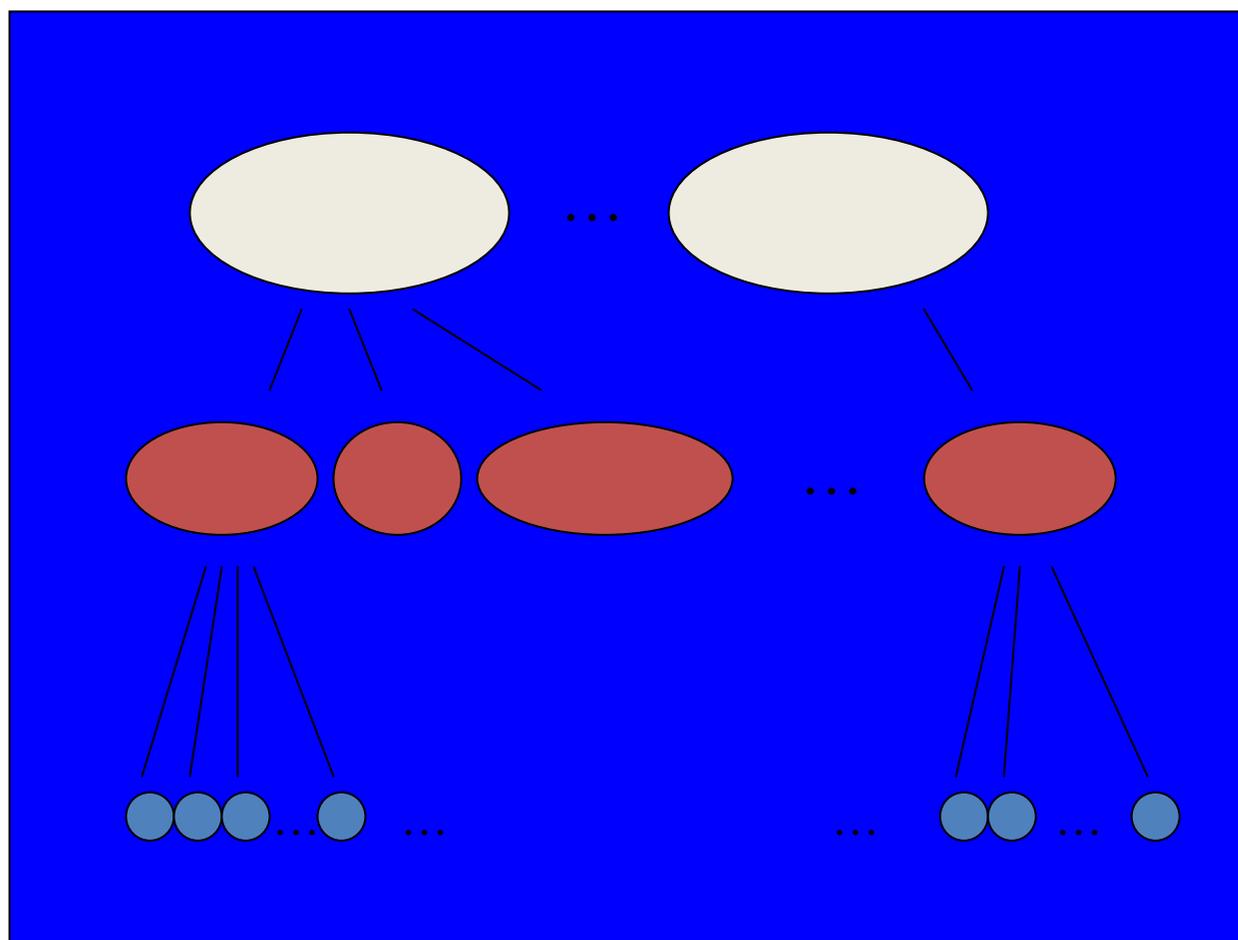




**ined**  
INSTITUT  
NATIONAL  
D'ÉTUDES  
DÉMOGRA  
PHIQUES

# Enquêtes à la sortie des bureaux de vote

Par exemple sur les grandes communes:



**Commune**

**Bureau de  
vote / quartier**

**Individu**

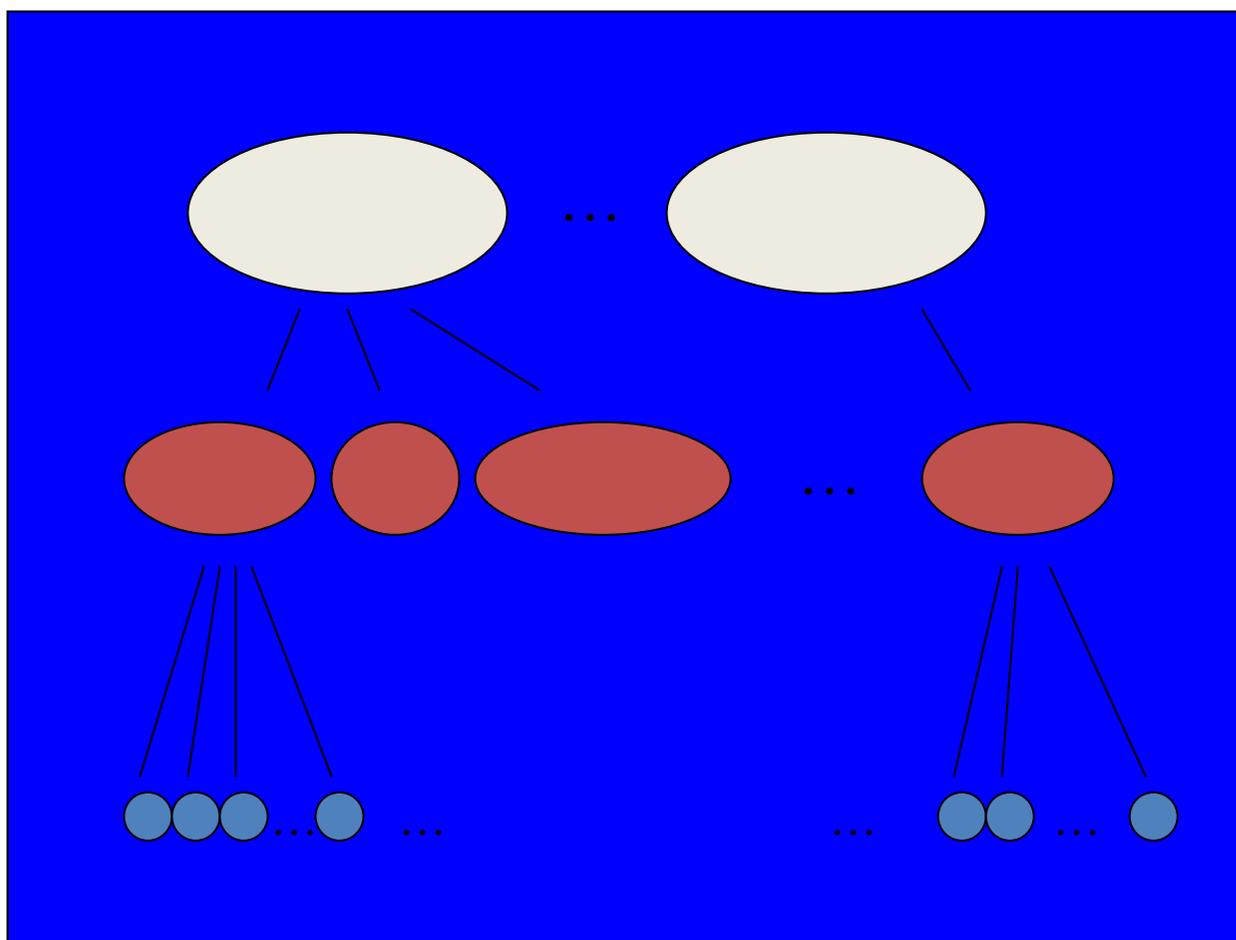


**ined**  
INSTITUT  
NATIONAL  
D'ÉTUDES  
DÉMOGRA  
PHIQUES



# Enquêtes à la sortie des bureaux de vote

Autres types de regroupement:



Département

CSP

Individu



**ined**  
INSTITUT  
NATIONAL  
D'ÉTUDES  
DÉMOGRA  
PHIQUES



# Conclusion : qu'apportent les modèles multiniveau?

Se rapprocher le plus possible de la réalité

Modéliser la complexité

Guider la modélisation

=> Identifier les niveaux d'observations pertinents et les caractéristiques importantes à chaque niveau

=> Travailler sur des données moins nombreuses que si l'on considérait les groupes séparément les uns des autres

Mais... faire face à encore beaucoup d'écueils



**ined**  
INSTITUT  
NATIONAL  
D'ÉTUDES  
DÉMOGRA  
PHIQUES



## Limites

C'est toujours de la modélisation statistique (il y aura toujours de l'hétérogénéité inobservée, heureusement!)

La complexité de certains modèles est encore mal prise en charge

Attention à ne pas confondre faits et perceptions





**ined**  
INSTITUT  
NATIONAL  
D'ÉTUDES  
DÉMOGRA  
PHIQUES

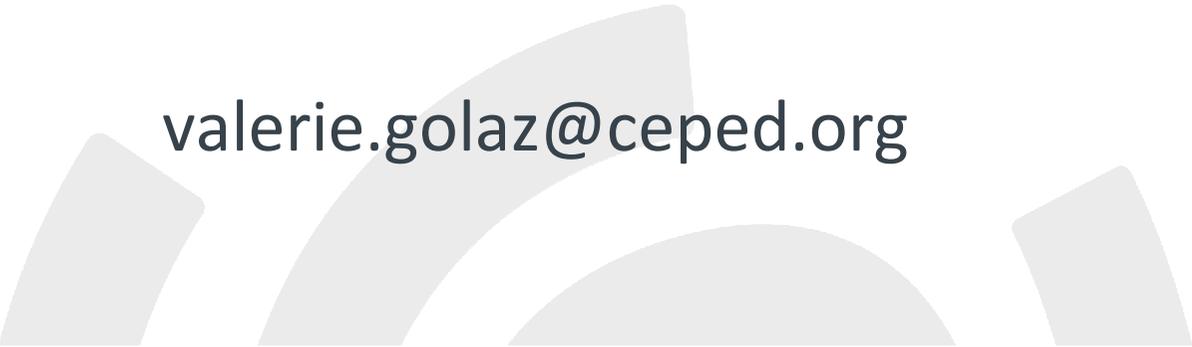


---

Merci pour votre attention!

---

[valerie.golaz@ceped.org](mailto:valerie.golaz@ceped.org)



# ANALYSES CONTEXTUELLES ET MULTINIVEAUX

JOËL GOMBIN (CURAPP-ESS UMR 7319, CNRS, UPJV)

- 1 Introduction
- 2 La modélisation multiniveau
- 3 Rendre contextuelle l'analyse multiniveau
- 4 Données individuelles et agrégées
- 5 Les modèles de croissance

Un intérêt renouvelé  
pour une sociologie électorale  
contextuelle et spatialisée

Utiliser la modélisation multiniveau ?

# Utiliser la modélisation multiniveau ?

- Modélisation

# Utiliser la modélisation multiniveau ?

- Modélisation
- Multiscale

# Utiliser la modélisation multiniveau ?

- Modélisation
- Multiscale
- Variabilité spatiale des relations

# Utiliser la modélisation multiniveau ?

- Modélisation
- Multiscale
- Variabilité spatiale des relations
- Contextes

# La modélisation multiniveau

# La modélisation multiniveau

- Multiniveau

# La modélisation multiniveau

- Multiniveau
- Hiérarchique

# La modélisation multiniveau

- Multiniveau
- Hiérarchique
- Mixtes

# La modélisation multiniveau

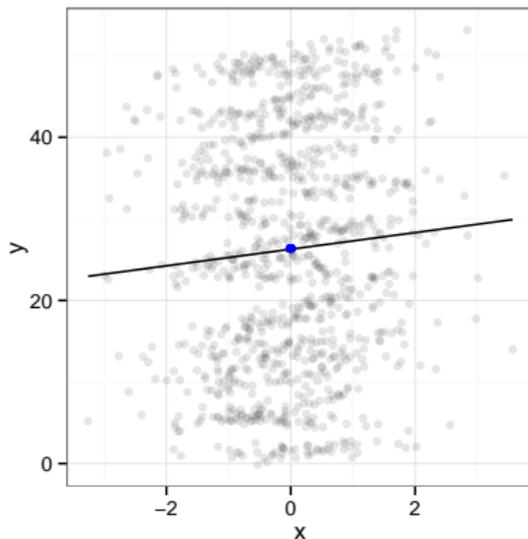
- Multiniveau
- Hiérarchique
- Mixtes
- À coefficients aléatoires

# Une généralisation du modèle linéaire

$$y_i = \alpha + \beta x_i + \epsilon_i$$

avec

$$\epsilon_i \sim N(0, \sigma_\epsilon^2)$$



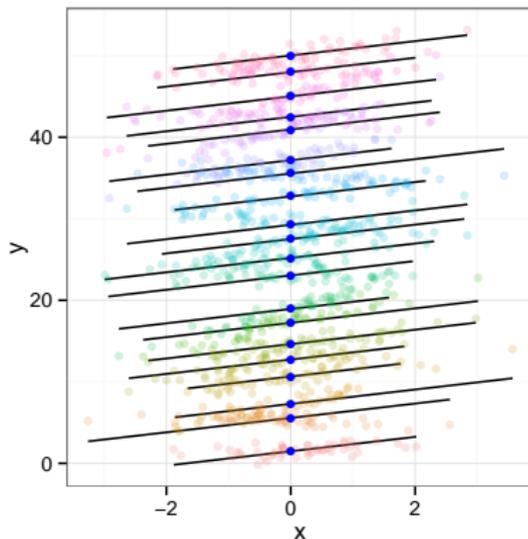
# Une généralisation du modèle linéaire

$$y_i = \alpha_{j[i]} + \beta x_i + \epsilon_i$$

avec

$$\alpha_{j[i]} \sim N(\mu, \sigma_\alpha^2)$$

$$\epsilon_i \sim N(0, \sigma_\epsilon^2)$$



# Une généralisation du modèle linéaire

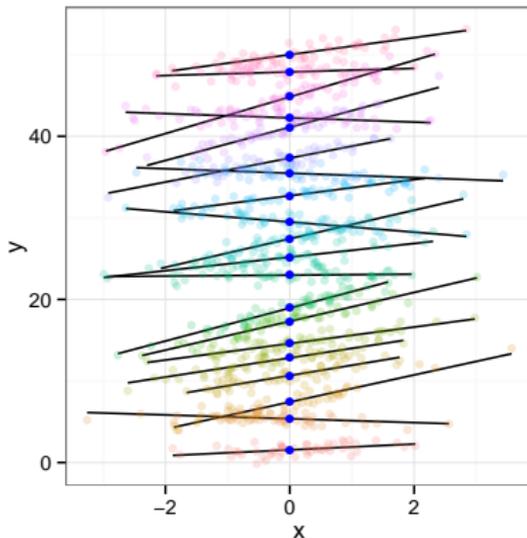
$$y_i = \alpha_{j[i]} + \beta_{j[i]}x_i + \epsilon_i$$

avec

$$\alpha_{j[i]} \sim N(\mu_\alpha, \sigma_\alpha^2)$$

$$\beta_{j[i]} \sim N(\mu_\beta, \sigma_\beta^2)$$

$$\epsilon_i \sim N(0, \sigma_\epsilon^2)$$



# Une généralisation du modèle linéaire

**Modèle linéaire** : hypothèse de constance des coefficients dans toutes les unités

# Une généralisation du modèle linéaire

**Modèle multiniveau** : hypothèse sur la distribution (usuellement normale) de la distribution des coefficients

## Rendre contextuelle l'analyse multiniveau

Derivry et Dogan (1971) :

# Rendre contextuelle l'analyse multiniveau

Derivry et Dogan (1971) :

- unité d'analyse

# Rendre contextuelle l'analyse multiniveau

Derivry et Dogan (1971) :

- unité d'analyse
- espace de référence

# Une généralisation du modèle linéaire

$$y_i = \alpha_{j[i]} + \beta_{j[i]}x_i + \epsilon_i$$

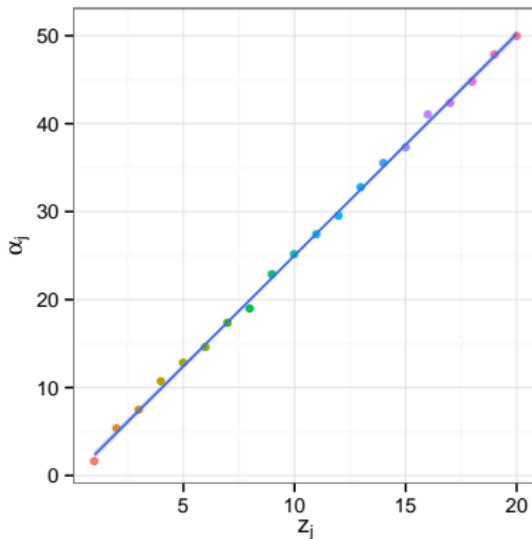
avec

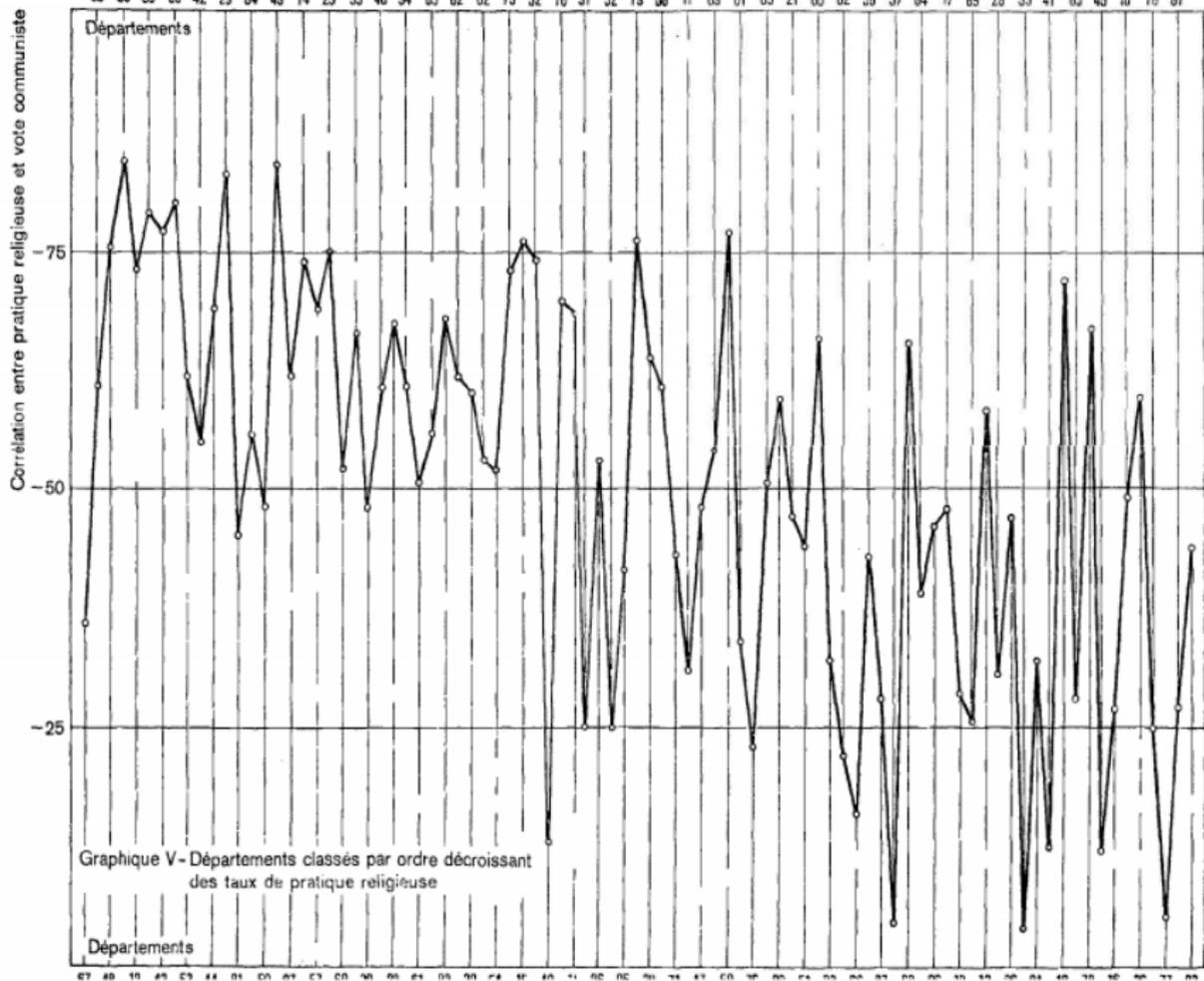
$$\alpha_{j[i]} = \mu_\alpha + \theta z_j + \gamma_j$$

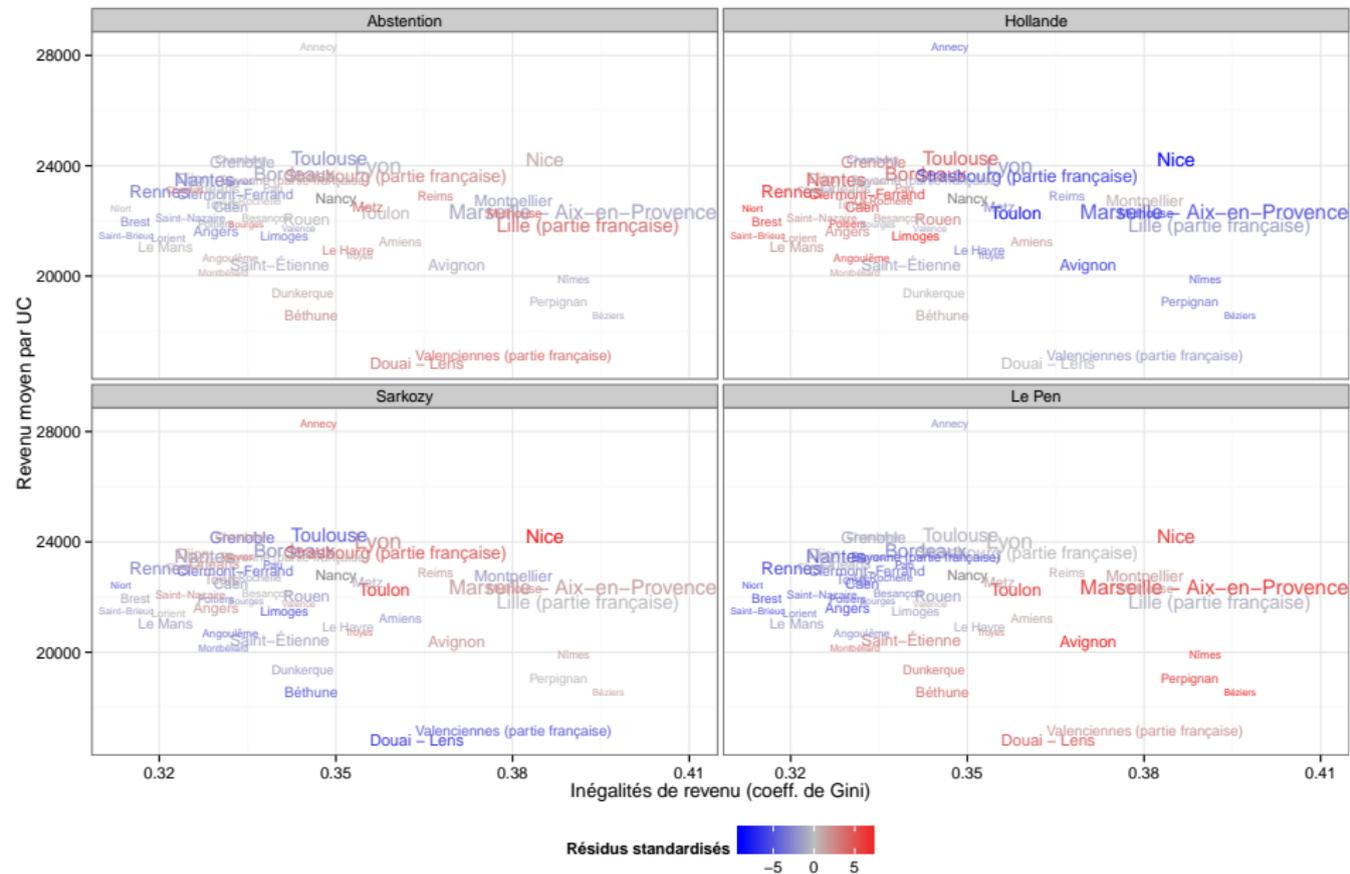
$$\beta_{j[i]} \sim N(\mu_\beta, \sigma_\beta^2)$$

$$\epsilon_i \sim N(0, \sigma_\epsilon^2)$$

$$\gamma_j \sim N(0, \sigma_\gamma^2)$$







# Données individuelles et agrégées

# Données individuelles et agrégées

- Multiniveau usuellement utilisé sur données individuelles et agrégées

# Données individuelles et agrégées

- Multiniveau usuellement utilisé sur données individuelles et agrégées
- Mais rien n'interdit de prendre des données agrégées comme niveau 1

## Données individuelles et agrégées

Boudon (1963), suivant Lazarsfeld et Menzel (1961) :

# Données individuelles et agrégées

Boudon (1963), suivant Lazarsfeld et Menzel (1961) :

- propriétés analytiques

# Données individuelles et agrégées

Boudon (1963), suivant Lazarsfeld et Menzel (1961) :

- propriétés analytiques
- propriétés structurelles

# Données individuelles et agrégées

Boudon (1963), suivant Lazarsfeld et Menzel (1961) :

- propriétés analytiques
- propriétés structurelles
- propriétés globales

## Données individuelles et agrégées

La modélisation multiniveau peut résoudre partiellement la fallace écologique (Robinson 1950) : voir Subramanian (2009).

## Articuler le temps à l'espace : les modèles de croissance

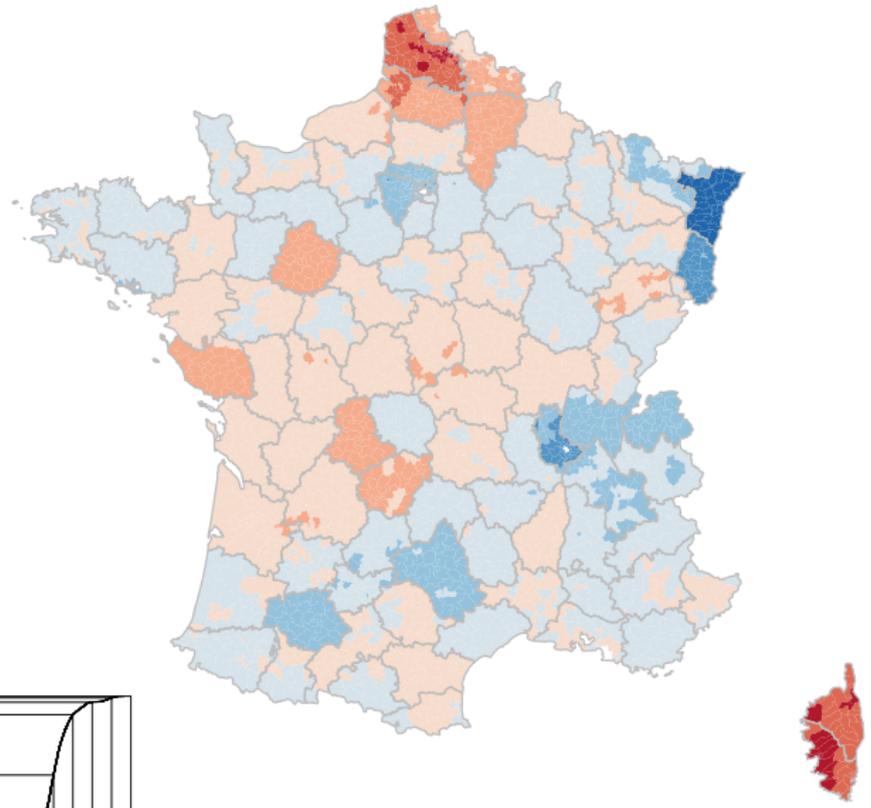
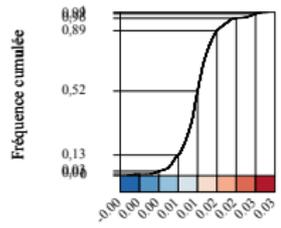
$$y_t = \alpha_i + \beta_i T_t + \epsilon_t$$

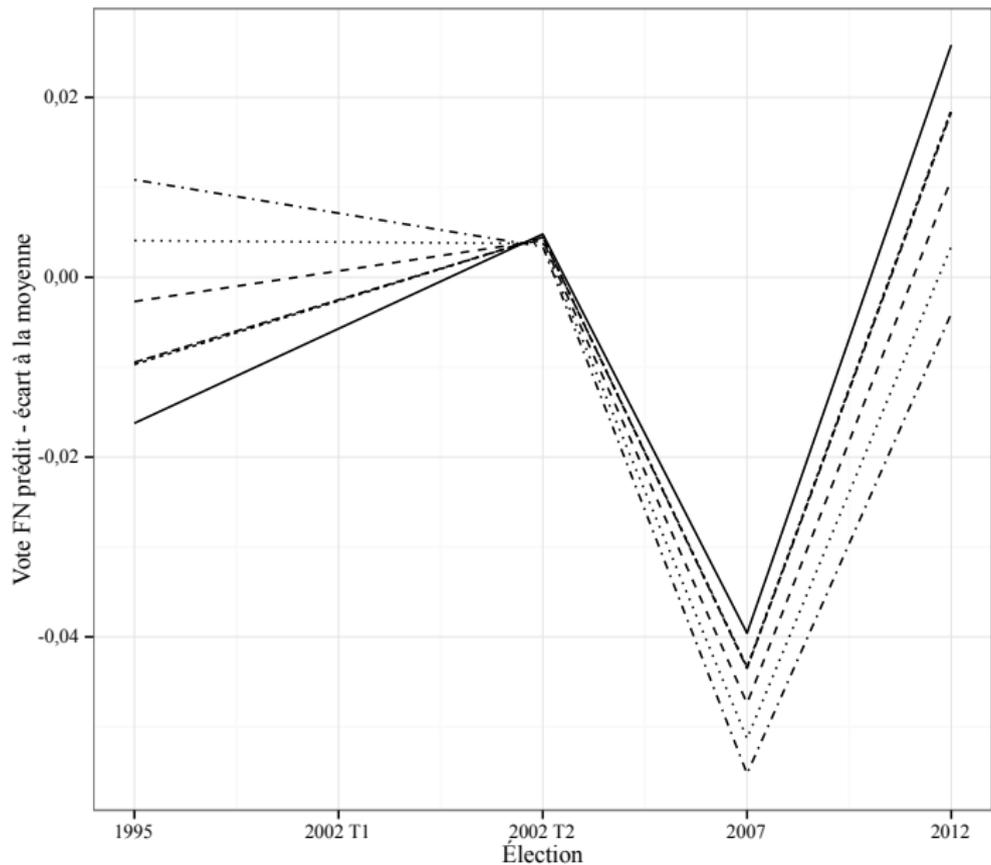
avec

$$\alpha_i \sim N(\gamma, \delta_i^2)$$

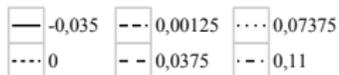
$$\beta_i \sim N(\eta, \theta_i^2)$$

$$\epsilon_t \sim N(0, \sigma_i^2)$$





CS2  
écart à la moyenne



Merci de votre attention !

## Table ronde

Le séminaire SFdS-AFSP s'est achevé par une table ronde organisée par Jean Chiche qui ainsi a fait fructifier son double profil de politologue du CEVIPOF et de statisticien, président du groupe Enquêtes de la SFdS. Elle comprenait deux politologues Nonna Mayer, présidente de l'AFSP et Patrick Lehingue, politologue à l'Université de Picardie, et deux statisticiens aux responsabilités émérites à la SFdS, Avner Bar-Hen et Jean-François Royer.

Première question de Jean Chiche : Quel est le bon niveau territorial de données électorales ?

Cette question de Jean Chiche fait suite aux interventions des géographes et statisticiens qui ont mis en valeur l'apport des méthodes statistiques multi-niveaux aux études politiques, que l'on cherche à mesurer l'impact sur l'électeur du contexte politique local ou, à titre d'exemple, l'impact sur le vote local d'un contexte culturel régional.

Les intervenants sont partis du constat que le fichier électoral ignorait l'unité « bureau de vote », que celle-ci était très difficile à reconstituer auprès des préfetures ou des mairies et de plus évolutif et socialement pas toujours pertinent, que bien entendu les adresses des électeurs ne seront jamais rendues accessibles aux chercheurs sur un fichier national, ce qui rend illusoire l'espoir d'un carroyage des données électorales à rapprocher des nouvelles données démographiques de INSEE carroyées par pavé de 200 mètres de côté. Aussi s'est-il imposé un point de vue pragmatique : espérer des données à un niveau territorial le plus fin possible et au chercheur ensuite de faire les agrégations territoriales au niveau qu'il pourra et qui l'intéressera. Les données locales diffusées par le projet Cartelec y contribuent de façon très positive. Ses cartes richissimes présentées en cours de séance en témoignent pleinement et font espérer un développement pour les communes périurbaines, tant la distance au centre de l'agglomération est devenue politiquement cruciale. Le Centre d'accès sécurisé aux données (CASD) mis en place par l'INSEE permet dès maintenant d'y mener des appariements très intéressants. Il appartient aux chercheurs en science politique de s'y engouffrer.

L'existence de 60000 bureaux de vote et de 36000 communes constitue évidemment une maille d'information très fine que nos voisins européens nous envient. Les récentes publications électroniques du ministère de l'Intérieur au niveau des 36000 communes et 4000 cantons constitue un progrès précieux.

Seconde question : Comment mesurer l'abstention ?

Les intervenants soulignent la grande pertinence de premières enquêtes Participation électorale de l'INSEE échantillonnées à partir de l'Echantillon Démographique Permanent, un extrait longitudinal au 1/100<sup>ème</sup> du recensement et de l'état civil fondé sur 4 jours annuels de naissance. Mais la réforme du recensement, transformé en sondage, a détérioré cet outil puissant en rompant la continuité de la collecte. Pour

son enquête Participation au vote, l'INSEE a donc abandonné la technique du sondage dans l'EDP suivi d'un appariement au fichier électoral pour un sondage direct dans le fichier électoral. On y perd la non inscription sur la liste et les riches variables sociales et démographiques de l'EDP pour finalement ne connaître de l'électeur que son sexe, son âge, sa commune inscription et sa participation aux scrutins successifs au vu des feuilles d'émargement. L'enquête Participation de l'INSEE n'en demeure pas moins un outil de connaissance politologique exceptionnel. Mais surtout, l'appariement entre l'enquête Participation 2012 et le fichier de l'enquête censitaire de 2010 mené à l'INSEE par Xavier Niel est extrêmement prometteur.

Les politologues attendraient donc beaucoup pour 2017 d'un appariement du fichier électoral et de l'EDP.

En plus de cette belle source quantitative, les études qualitatives telles que celle de Céline Braconnier et Jean Yves Dormagen auprès de la Cité des Cosmonautes avec une collecte en porte à porte fournissent un éclairage complémentaire très précieux, sans oublier les méthodes expérimentales de terrain développées avec Vincent Pons sur le modèle des travaux américains de Gerber et Green, et reprises par Céline Braconnier et Jean Yves Dormagen destinées à tester l'impact de la pression sociale et de la réduction du coût matériel d'inscription sur la participation électorale.

Avec de moins en moins de moyens pour la recherche, de plus en plus de données hétérogènes et de qualité incertaine ou d'une autre nature, souvent plus subjective, une nouvelle réflexion s'impose. Il s'agit en particulier des méga-données (big data) issues du web ou des réseaux et des techniques statistiques de l'apprentissage, et non plus seulement celles du sondage.

L'évolution sociale et la montée de la précarité ont fait perdre de sa pertinence à la CSP à un chiffre et rendu d'autant plus utile le recours à la CSP à deux chiffres sur des gros fichiers. C'est un apport précieux attendu de l'EDP et des appariements.

Dernière question : Et pour 2017 ?

La source CARTELEC, l'étude longitudinale 2013-2017 DYNAMOB au sein du panel Internet ELIPSS, les nouvelles méthodologies expérimentales développées à la suite de la monographie de la Cité des Cosmonautes témoignent de l'ampleur de l'innovation en cours. La mise à la disposition des chercheurs du fichier électoral à travers le réseau Quételet devra permettre dès ces élections municipales de relever les carences grossières de données statistiques électorales, par exemple en substituant pour chaque commune à la pyramide des âges des résidents français majeurs celle des inscrits ! Evidemment, dans les villes universitaires, l'écart entre ces distributions est majeur.

Nonna Mayer conclut la journée en soulignant la qualité du dialogue intervenu au cours de ce séminaire entre politologues, géographes et statisticiens, en les invitant à poursuivre cet échange dans le tout nouveau réseau de recherche de l'AFSP « Futur des Etudes Electorales » (FEEL) et à relever pour 2017 le défi des méga-données et surtout la mise sur pied d'une « French National Election Study » digne de ce nom.