

Analyse statistique
des voix obtenus par les candidats, par urne-date de dépouillement,
pour les votes au CA d'ATTAC (juin 2006)

1°) Pourquoi une nouvelle analyse

Le rapport de R. Passet et N. Albala concluant à l'existence de la fraude se fondait, entre autres éléments connus en septembre, sur les travaux de G. Duménil et D. Lévy.

Ceux-ci établissaient l'existence de la fraude sur l'urne B3 et B4 à l'aide de deux analyses conjointes (du moins dans le rapport en notre possession): une analyse discriminante et une analyse des styles de croix sur les bulletins (styles photocopiés dans le document).

Réalisant de nombreuses enquêtes dans divers domaines depuis plus de 20 ans pour mes propres recherches et plus fréquemment pour financer celles-ci, ces styles ne m'étaient pas inconnus.

Apprenant par ailleurs l'imminence d'une plainte déposée par le CA actuel d'ATTAC auprès d'un tribunal, j'ai pris l'initiative de vérifier l'existence de styles de croix sur une enquête menée en 2003. Les conclusions que nous commenterons dans la section 2°) ci-dessous ont induit une analyse approfondie des statistiques des votes par urne (résumé dans la section 3).

L'ensemble permet d'avoir un doute sérieux sur l'existence même de la fraude.

C'est la raison pour laquelle je demande à avoir accès, avec l'aide d'une équipe d'adhérents volontaires, à l'ensemble des bacs (dénommés urnes ci-dessous)¹, afin de vérifier la réalité des conclusions.

Les conclusions ci-dessous se fondent, pour la partie concernant les votes au CA d'ATTAC, sur le rapport de A. Barral et D. Goussot², le rapport statistique de J. Chiche - H. Rouanet, celui de G. Duménil - D. Lévy, et l'analyse du fichier fourni par l'huissier. Ce sont les seules informations en ma possession. Si d'autres informations existent, non divulguées au moment où ces lignes sont écrites, et qui établiraient l'existence de la fraude de manière indubitable, il va de soi que la demande ci-dessus deviendrait caduque.

2°) Chaque personne porte sa croix, fréquemment semblable à celles d'autres humains

Afin de reconnaître des styles éventuels de croix, deux études ont été réalisées, l'une rétrospective, l'autre prospective.

La première a consisté à reprendre 272 bordereaux d'une enquête manuelle effectuée sur cinq semaines entre la mi-septembre et la mi-octobre 2003, auprès des usagers d'un restaurant d'une grande école rennaise. La clientèle est formée du personnel de l'institution et en plus grand nombre de professionnels de niveau cadre en formation continue pour quelques jours (il s'agit donc de couches moyennes ... comme essentiellement à ATTAC).

Avec l'aide d'une collègue statisticienne (membre d'ATTAC), j'ai repéré aisément 9 styles bien distincts de croix (voir annexe 1), dont trois prédominent largement :

Style 1 : 29,4 % ; style 2 : 29,8 % ; style 3 : 21,3 %.

¹ Et pas seulement les urnes B3 et B4 qui risquent de conduire à des raisonnements tautologiques.

² Rapport sur les témoignages d'adhérents qui ont participé au dépouillement du scrutin pour l'élection du CA d'ATTAC - 17 sept. 2006

Les styles de croix repérés et scannés dans le rapport de G. Duménil et D. Lévy (notés a, b, c, d,e) s'inscrivent exactement dans des catégories de croix ci-dessus, en particulier les plus fréquentes .

Les fraudeurs étaient-ils présents dans ce restaurant en septembre-octobre 2003 ?
Curieux ...

Nous avons alors choisi un style (style 1 : croix entière dans le cadre), puis repéré deux bordereaux cochés avec des stylos de même épaisseur et de même couleur : impossible de savoir à l'œil nu si c'était d'une même personne ou non. A la fin du bordereau, des questions socio-démographiques étaient posées : il s'agissait bien de deux personnes différentes, l'une membre du personnel, l'autre stagiaire, de sexe et d'âge différents.

L'étude prospective a consisté, le jeudi 9 novembre après-midi, à demander à 14 étudiants volontaires (futurs développeurs dans les R&D des entreprises agro-alimentaires) de faire l'expérience suivante : sur le recto d'une feuille A4, faire 20 carrés semblables à ceux des enquêtes habituelles, puis de cocher 10 cases arbitraires à l'aide d'un stylo à mine fine. J'ai demandé de refaire l'expérience sur le verso, mais cette fois ci avec un stylo à mine épaisse (feutre).

Résultat : on trouve cinq styles qui correspondent exactement à cinq des neufs styles repérés précédemment. L'épaisseur de la mine du stylo ne change guère la forme de la croix, sinon qu'elle entraîne naturellement une utilisation plus importante de l'espace des cases.

Conclusions :

- 1- le repérage de croix identiques dans les bulletins n'induit en rien qu'elles sont le fait d'une même personne.**
- 2- Le fait que trois styles dominant largement entraîne que leur apparition dans un paquet de 100 bordereaux est quasi sûre.**
- 3- Si par hypothèse, la distribution des fréquences des styles de croix de l'enquête citée est applicable à l'ensemble des bulletins de vote, alors la probabilité de les retrouver dans une urne de 100 est grande (exemples : urnes B3 et B4), tout comme de trouver une suite de bulletins de style identique³.**

3°) Considérations statistiques à partir du fichier candidats / Urnes

a – Considérations élémentaires générales

a1- Chaque bulletin comporte 62 candidats (une variable à 62 modalités possibles). Dès lors, pour qu'un échantillon ait une probabilité de comporter toutes les modalités, surtout quand il y a des effets de liste, il faut raisonnablement qu'il soit important. Pour ma part, par expérience, je commencerais à analyser a priori un tel échantillon que s'il dépasse au bas mot 250 fiches afin d'avoir une bonne probabilité d'avoir une distribution relative à peu près assurée. Seules 4 urnes approchent ce seuil (10/06 : L2 avec 200 ; 11/06 : R1 avec 250; 12/06 : C3 avec 223; 14/06 : B2 total avec 398).

³ Je laisse le soin aux personnes douées de saines curiosités de trouver la loi statistique adéquate pour nous dire quelle est la probabilité de trouver k bulletins de même style à la suite (loi agrégative – contagious distribution).

a2- Pour une urne contenant 100 bulletins, il suffit d'un glissement de 10 % des votants (soit 10 votants) du choix d'un candidat vers un autre pour créer un écart de 20 points. On voit bien ici qu'il suffit de quelques déplacements de voix dans les petits échantillons (ici urnes à petits effectifs) pour créer de fortes variations des écarts entre les candidats. Les variations, même importantes, des pourcentages entre urnes à petits effectifs sont donc « naturelles » pour une variable à 62 modalités.

a3- On pourrait être tenté de descendre la barre de l'effectif de l'échantillon à, disons, 70, soit un nombre légèrement supérieure au nombre de modalités-candidats. Faisons le comptage des urnes à effectif inférieur, par jour de dépouillement :

- le 10 juin : Une urne sur deux dépouillées⁴;
- le 11 juin : 9 sur 19 dépouillées ;
- le 12 juin : 2 sur les 3 ;
- le 13 juin : 4 sur 7 ;
- le 14 juin : 6 sur 10 ;
- le 15 juin : 15 sur 21.

Il n'est donc pas étonnant qu'à partir du 12 juin et surtout au delà et tout particulièrement les 14 et 15 juin, les variations des pourcentages des voix par candidat ou blancs+nuls entre les urnes peuvent être importantes : ce que l'on peut aisément vérifier globalement en faisant un graphique avec en abscisse les urnes par date de dépouillement et en ordonnées les pourcentages obtenus par candidats⁵. Parmi les urnes contenant plus de 70 bulletins, les plus grands écarts, importants, sont le fait des urnes B3 et B4 (effectif total : 100 bulletins chacune, dépouillées le 14 juin) : écart moyen d'environ 35 points entre les deux listes rivales pour B3, soit l'équivalent de 17 à 18 électeurs (donc 17 à 18 %), et de 55 points environ pour B4, soit 27 à 28 électeurs en faveur de la liste JN. L'urne T3 (115 votants, dépouillée le 15 juin), non suspect car les étapes de dépouillement ont été effectuées à suivre dans le même journée, montre un écart d'environ 36 points, soit un glissement de 15 à 16 votants en faveur de la liste JN. Celle-ci domine aussi dans les résultats de l'urne R2 (75 votants, 15 juin), résultats différents de l'urne R1 (fort effectif : 250, 11 juin). R2 ne montre aucune anomalie après vérification (voire rapport Barral-Goussot).

a4- Pour toute enquête, outre le protocole de remplissage (ici choisir **au plus** 24 noms sur les 62), celui du dépouillement est, en fonction de l'objectif affiché de mélange à effets aléatoires pour chaque lettre, primordial. Le rapport de A. Barral et D. Goussot montre que le protocole établi n'a pas pu être suivi. On peut affirmer que le contenu des urnes n'est pas aléatoire. Cette dernière remarque, croisée avec celles du paragraphe précédent a3, montre désormais qu'il serait incorrect de prétendre qu'une ou des urnes auraient une distribution « juste », « normale » (dans le sens courant du terme), « typique ». Les bulletins des urnes ne représentent donc pas des échantillons aléatoires sur lesquels pourraient être fait des tests de comparaisons inférentiels pertinents. Par ailleurs, le mode de dépouillement qui a consisté à séparer les bulletins en deux lots Harribey → Azam puis Maurel → Clequin⁶ (afin d'accélérer le dépouillement) n'aurait eu aucun effet sur le résultat des urnes où cette technique a été réalisée si chaque lot avait été terminé le même jour, ce qui n'a pas été le cas, en particulier pour la lettre B (... « d'autres dépouilleurs, prenant le relais le lendemain, ont formé de nouveaux lots avec les bulletins qui restaient à dépouiller » nous apprend le rapport).

⁴ Je n'ai pas inclus à cette date les urnes A4, G5, M3, et M4, leur date respective de dépouillement étant ambiguë.

⁵ Bien étirer le graphique sur au moins deux pages Excell pour s'en apercevoir.

⁶ Premier lot → liste JN : 10 élus ; liste SG : 6 élus . Second lot → liste JN : 5 élus ; liste SG : 3 élus.

L'urne B4, par exemple, comprend 16 noms qui totalisent plus de 63 voix selon les dépouilleurs 7 et 8 (p17 du rapport) : Nikonoff, Jauffret, Jonquet, Tassi, Goussot, Karbowska, Ludi, Coignard, Lambert, Bernier, Laurent, Barral, Gaziello, Weber, Gicquel, Landfried. Sur ces 16 noms, 11 font partie du haut de la liste (lot Harribey → Azam). Fort possible qu'une partie de l'autre lot ait été mélangée aux bulletins d'une autre urne (hypothèse sans doute difficile à vérifier aujourd'hui). On remarquera qu'à l'inverse, la liste de S. George domine pour l'urne B5 (contenant certainement des bulletins issus d'enveloppes retirées de la Poste le 3 juin - et peut-être une partie du deuxième lot ?).

a5- La lecture attentive de ce dernier rapport (à lire ligne par ligne, surtout les commentaires sur chaque lettre), montre qu'il y a un certain effet chronologique des réceptions des votes à la Poste (réceptions du 3 juin, du 10 juin et du 15 juin : pas de mélange pour plusieurs lettres à effectif important : notamment B, D, G, L). Il n'est pas aisé de reconstituer la chronologie par lettre-urne. L'urne B5 (50 bulletins) par exemple, dépouillée seulement le 14 juin, contient certainement des bulletins retirés à la Poste le 3 juin ; l'urne R2, dépouillée le 15 juin, contient 75 enveloppes retirées le 10 juin (suite de R1), globalement favorable à la liste JN, tandis que l'urne R1 (250 bulletins eux avec mélange des dates d'arrivée 3 et 10 juin) ne dégage pas une distinction nette entre les candidats des deux listes (pour les candidats ayant obtenu globalement entre 32 % et 49 %). Rappelons ici que la liste S. George a été diffusée le 20 mai (date Attac-Local), et que celle de J. Nikonoff l'a été, à ma connaissance, plus tardivement, au moins une semaine plus tard. Dans ces conditions, il est fort possible que cela ait eu un certain effet différencié sur les arrivées chronologiques des enveloppes et donc des dépouillements tels qu'ils ont été effectivement réalisés. Cela expliquerait peut-être aussi les dispersions de plus en plus forte des voix à partir du 13 juin.

4°) Analyses statistiques des bulletins-urnes

Nous avons réalisé une analyse multidimensionnelle sur l'ensemble du fichier candidats-Urnes. La méthode utilisée est l'AFCM couplée à une CAH qui permet de dégager des classes (ici neuf classes significatives homogènes) de candidats et d'urnes de manière rigoureuse⁷. Les classes de candidats montrent un effet liste quels que soient les scores obtenus. Il serait trop long ici de décrire l'ensemble des résultats descriptifs. Si cela paraît utile, nous en ferons un rapport complet.

Ces méthodes nous ont permis de repérer des classes de candidats homogènes (vis à vis des pourcentages obtenus par urne, et vis à vis des listes affichés), puis de les comparer.

Les classes 7+8 (tous candidats « S. George », ayant obtenu entre 32 % et 36% des voix, donc proches ou assez proches du seuil d'élection) ont été comparées avec la classe 6 (tous candidats « Nikoviens », ayant obtenu entre 38 % et 42 %, donc tous élus juste au dessus du seuil, et dont l'élection fait l'objet de soupçons). On retrouve la tendance générale : faible distinction des listes par urne le 10 juin et en grande partie le 11 juin, puis distinction de plus en plus prononcée, les écarts les plus importants se dégageant pour les urnes B3 et B4.

On notera que les pourcentages de voix par urne des candidats les plus mal placés (classe 1 et 2 : de 7 % à 33 %), tous « nikonoviens » ou non alignés à deux exceptions près,

⁷ Le plan 2-3 est de loin celui qui permet l'interprétation la plus aisée, couplée avec les classes. J. Chiche et H. Rouanet ont utilisé une autre méthode, l'ACP, méthode puissante dans le cas de ce fichier mais qui a l'inconvénient de ne pas permettre d'analyser directement urnes et candidats sur un même plan. Celle utilisée par G. Duménil et D. Lévy est l'analyse discriminante, non sur les candidats, mais sur les bulletins, ce qui est la meilleure méthode dès lors que l'on a eu accès à ceux-ci. Malheureusement, cette dernière étude ne porte que sur les urnes dites suspectes B3 et B4, ce qui peut porter à la critique d'un raisonnement tautologique (on trouve ce que l'on a défini à l'avance). Il conviendrait de généraliser cette étude à TOUTES les urnes pour pouvoir tirer des enseignements fiables généralisables.

suivent en général les mêmes mouvements de hausse et de baisse d'urne-date en urne-date que les candidats nikonoviens élus (classe 6 et 5). Ceci est particulièrement visible pour l'urne B4, urne considérée comme particulièrement « atypique » (terme peu pertinent comme nous l'avons montré plus haut). Dans ces conditions, il n'est pas étonnant que la liste de S. George ait subi un grignotage général de leurs voix, grignotage qui se mesure par des différences de pourcentages élevés compte tenu de la faiblesse des effectifs de bulletins dans la plupart des urnes, et ce d'autant que les candidats de S. George sont minoritaires parmi les 62 candidats (42 %).

Nous avons comparé d'autres classes entre elles, mais les commentaires précédents nous paraissent suffisants à ce stade.

CONCLUSION :

- 1- L'étude des croix montre qu'elles sont typées, certaines à fréquence forte et qu'il n'est guère pertinent d'en déduire qu'il y a eu obligatoirement fraude de la part d'une ou plusieurs personnes.
- 2- L'analyse statistique, sauf peut-être une analyse plus poussée telle que nous le demandons en introduction, ne nous certifie rien : les variations les plus fortes constatées peuvent aussi bien être le fait d'une fraude que des effets conjugués des dates décalées de diffusion des listes + des modes de dépouillement + des faibles effectifs de la plupart des urnes relativement au nombre de candidats + du nombre minoritaire des candidats de S. George.

J.P. ESCAFFRE
CL de Rennes
Le 12-11-2006