

L’empreinte indélébile de la fraude

Thomas Coutrot & Michel Husson, 20 août 2006

Le débat sur le vote au CA s’est intensifié pendant l’été, dans l’attente des rapports des experts indépendants. Après une première vague de documents concluant à la fraude, on a vu se multiplier des contributions visant à montrer, soit qu’il n’existait pas d’anomalies, soit qu’elles pouvaient s’expliquer autrement que par la fraude, et notamment par les modalités mêmes du dépouillement. Les principales contributions sont en ligne à l’adresse suivante : <http://hussonet.free.fr/attac0.htm>.

Ce document vise à résumer arguments et contre-arguments en utilisant une présentation simplifiée qui renvoie aux contributions plus détaillées. Il résume la version mise à jour de [L’empreinte de la fraude](#). Il est organisé autour de deux questions qui structurent le débat : 1) existe-t-il vraiment des anomalies ? 2) s’il existe des anomalies, peuvent-elles s’expliquer autrement que par la fraude ? La conclusion de cette discussion est claire : il existe des anomalies, et celles-ci ne peuvent s’expliquer que par la fraude.

1. Le scrutin présente des anomalies

De nombreux documents (Coutrot, Husson, Duménil-Lévi) ainsi que l’expertise Chiche-Rouanet concluent à l’existence de fortes anomalies. Le repérage des lots atypiques repose toujours sur une méthode qui consiste : 1) à traiter le problème des « petits » lots ; 2) à calculer une « distance » entre chaque lot et un résultat de référence ; 3) à fixer des seuils à partir desquels on peut parler d’anomalie.

Le calcul de la « distance » revient à travailler sur l’écart entre le pourcentage obtenu par chaque candidat dans un lot donné et son pourcentage d’ensemble. Comme ces écarts sont par définition positifs pour certains candidats et négatifs pour d’autres, et que leur somme s’annule, il n’existe que deux solutions : soit prendre la valeur absolue des écarts (c’est la solution d’Alain Gély), soit prendre la somme des carrés des écarts, ce que font des indicateurs comme le coefficient de corrélation ou le Khi-deux. La théorie statistique s’est développée à partir du second choix, parce qu’il permet d’établir des lois statistiques précises et tabulées, ce que la valeur absolue ne permet pas. La solution d’Alain Gély est donc pré-scientifique et, en tout état de cause, ne devrait pas conduire à des classements différents.

Le seuil définissant l’anomalie pour des lots suffisamment importants est arbitraire. Ces choix ont une part d’arbitraire qui permet de comprendre, par exemple, que le travail de Duménil et Lévy ne conduit pas à la même liste de lots atypiques que celui de Husson. Mais ce degré d’arbitraire peut être contrôlé après coup.

En dépit d’une liste de lots atypiques différente, Duménil-Lévy et Husson arrivent exactement au même résultat quant à la composition du CA qui aurait résulté de la seule prise en compte des lots « conformes ». C’est un premier test de robustesse qui montre qu’il existe un « noyau dur atypique » commun à presque toutes les études, malgré le « halo » qui peut exister sur son périmètre exact.

Un autre contrôle très probant concerne les petits lots de moins de 50 bulletins qui apparaissent comme atypiques (Husson). Si on les cumule, on obtient un profil très voisin des lots conformes. En revanche, les lots atypiques confirment qu’ils le sont effectivement, puisque le fait de les regrouper ne fait pas disparaître ce caractère atypique.

On ne peut faire disparaître le caractère atypique d’une dizaine de lots qu’en se fixant des seuils extrêmement laxistes (Gély) ou exigeants (Fenayon) de telle sorte que presque tous les lots apparaissent alors respectivement comme conformes ou atypiques. Mais ce ne sont que des procédés pour gommer la très faible probabilité de tirer au hasard des lots si atypiques (voir les calculs de Grenouillet et Tzanos, en attendant ceux des experts indépendants). Il faut donc conclure, comme le font d’ailleurs la majorité des études disponibles, favorables ou non à la thèse de la fraude, à la réalité des anomalies.

2. Les anomalies ne peuvent s'expliquer par un effet chronologique

résumé de [A nouveau sur l'hypothèse chronologique](#)

Plusieurs contributions (notamment celles de Lasserre, Fenayon et Franz) admettent la présence d'anomalies mais cherchent à montrer qu'elles s'expliquent en fait par le mode de constitution des « lots ». Le tri alphabétique n'aurait pas suffi à éliminer un double effet (de liste et chronologique) qu'on peut résumer ainsi : les premiers votants n'ont pas rempli leur bulletin en fonction des listes annoncées (celle de Susan George, puis celle « des 32 »). L'influence de ces quasi-listes n'a joué que sur la seconde vague des votants. Or, les lots n'ont pas suffisamment « brassé » ces deux vagues, de telle sorte que les lots « atypiques » n'apparaîtraient comme tels que parce qu'ils regrouperaient des votes intervenus plus tardivement, beaucoup moins « panachés », et marqués par une volonté de sanctionner la liste de Susan George.

Cette discussion est centrale, parce que cette hypothèse chronologique est la seule alternative rationnelle à celle de la fraude, si on laisse de côté celles qui imaginent une intervention extérieure aux deux listes visant à déstabiliser Attac en délégitimant le scrutin. Cela permet de comprendre l'insistance de Nikonoff à invalider le premier décompte au profit du second réalisé par l'huissier. Ce dernier fait disparaître un phénomène troublant et totalement inexplicable avec l'hypothèse chronologique qui est la séquence en trois segments du premier décompte : un segment S1 (jusqu'au 11 juin) favorable à la liste George ; un segment S2 (13 au 15 juin) qui lui est très défavorable ; et un dernier segment S3 qui ressemble au premier. Or, les regroupements opérés par l'huissier tendent à diluer cette séquence à trois temps au profit d'une séquence à deux temps compatible avec l'hypothèse chronologique (ce point est documenté par Delepouve).

Cette hypothèse a déjà été discutée dans [L'empreinte de la fraude](#) où il était montré qu'elle ne rendait pas compte de trois éléments contrefactuels : absence de lots atypiques parmi les lettres intégralement dépouillées avant le 11 juin ; présence de lettres dont les lots sont conformes de bout en bout ; distribution aléatoire des lots atypiques (ainsi les lots de la lettre D sont conformes le 11 juin, atypique le 12, puis de nouveau conformes les 13, 14 et 15).

La réfutation de l'hypothèse peut être menée de manière plus rigoureuse encore, selon les indications de Theulière et Tosti, en utilisant une ventilation des lettres en trois catégories :

- cinq lettres (A, E, F, H et M) ont été dépouillées en une fois avant le 11 Juin (828 bulletins) ;
- huit lettres (I, J, K, S, W, X, Y, Z) ont été dépouillées en plusieurs fois après le 11 Juin mais ne présentent pas de lots atypiques (493 bulletins) ;
- les autres lettres ont été dépouillées en plusieurs lots, dont un au moins est atypique.

Le premier groupe de lettres constitue donc un ensemble de référence, dépourvu par définition de biais chronologique puisqu'il n'y a pas eu de dépouillement fractionné en plusieurs lots. Les résultats sont alors sans appel : le groupe de lettres sans lots atypiques est bien corrélé avec le groupe de référence « déchronologisé » ; la plus grosse des lettres à lots atypiques (B=732 bulletins) est significativement moins bien corrélé avec ce même groupe de référence. Cette analyse invalide donc les objections (Lasserre, Fenayon) selon lesquelles l'anatomie de la fraude reposerait sur une « analyse insuffisante des conditions de dépouillement » et sur un postulat erroné selon lequel « le tri alphabétique assure le caractère aléatoire des lots ».

3. Les anomalies ne peuvent s'expliquer que par la fraude

Trois faits sont ainsi établis, que les rapports des experts ne pourront que confirmer :

1. Il existe des lots atypiques
2. Ces lots atypiques n'apparaissent qu'après le cumul du 11 Juin.
3. Les distorsions introduites par les lots atypiques favorisent une « liste » au profit d'une autre.

On vient de voir que l'hypothèse chronologique ne pouvait rendre compte de ce premier ensemble de fait. Mais il faut en expliquer en outre un résultat essentiel qui résulte d'une comparaison entre deux informations :

- l'écart entre le nombre de voix que chaque candidat a effectivement obtenues et celui qu'il aurait obtenu en fonction de son pourcentage calculé sur les seuls lots conformes ;
- les résultats intermédiaires du scrutin fournis au soir du 11 Juin.

Cette comparaison montre que la distribution des écarts de voix opérés par les atypiques dépend du rang de chaque candidat à mi-parcours : **les candidats en position éligible bénéficient ou pâtissent de déplacements de voix plus élevés.** Cette dépendance peut être établie sur la base d'un test statistique rigoureux (voir [A nouveau sur l'hypothèse chronologique](#)).

Autrement dit, la distribution des écarts de voix engendrés par les lots atypiques dépend de manière non aléatoire du classement au soir du 11 Juin. C'est un fait essentiel qui valide l'hypothèse de la fraude.

Pour bien comprendre pourquoi, il faut rappeler les éléments suivants :

1. avant le 11 Juin, personne ne disposait a priori d'information sur le classement des 62 candidats et il était impossible de dresser une liste de ceux qui se retrouveraient en position éligible.
2. les écarts ultérieurs dépendent de cette information ;
3. tous les votes avaient déjà été émis à la date du 11 Juin.

Seule la fraude permet d'expliquer que les écarts constatés après le 11 Juin intègrent une information qui n'était pas disponible avant cette date. La fraude est donc établie, parce que les écarts repérés statistiquement dépendent des résultats connus seulement le 11 Juin et n'apparaissent pas avant cette date, ce qui récuse l'hypothèse chronologique. La fraude a donc consisté à remplacer des bulletins authentiques par des bulletins ciblés sur les seuls candidats éligibles. Elle a conduit au résultat visé, mais elle a laissé son empreinte indélébile sur les résultats du scrutin.