

Influence d'une chaîne de TV conservatrice sur le vote aux élections présidentielles américaines

Jérémy L'Hour
jeremy.l.hour@ensae.fr

March 26, 2018

Ce projet tente de quantifier l'effet de la diffusion d'une chaîne de télévision conservatrice entre 1996 et 2000 sur le vote conservateur aux élections présidentielles américaines correspondantes. MAGA News est une chaîne d'information en continu créée en septembre 1996. Les spécificités du marché américain de la télévision câblée rendent le déploiement d'une nouvelle chaîne de TV nécessairement long et coûteux : étant donnés les coûts fixes, chaque ville américaine constitue un monopole local opéré par une des quelques grandes compagnies de télévision câblée. En outre, ces compagnies font face à des contraintes sur le nombre de chaînes qu'elles peuvent diffuser. MAGA News a donc dû négocier sa diffusion auprès des compagnies câblodistributrices, souvent au détriment d'autres chaînes. MAGA News est considérée comme une chaîne à droite sur le spectre politique américain. Son orientation est également considérée comme bien plus conservatrice que les chaînes de télévision concurrentes. Dans la suite de l'énoncé, on appellera "groupe traité" l'ensemble des villes ayant accès à MAGA News en 2000 et "groupe contrôle" l'ensemble des autres villes.

Données La base de données `MAGANews.dta` regroupe des informations concernant la diffusion des chaînes de télé, les caractéristiques sociodémographiques et le vote, mesurées pour 9,265 villes américaines sur plusieurs années. Elle est disponible à l'adresse www.github.com/jlhourENSAE/MAGA-Econometrics.

Consignes importantes Pour toute réponse demandant une estimation, vous devez fournir un écart-type correspondant, dont le mode de calcul sera justifié. Une grande attention sera

portée à la façon d'exposer vos résultats. Inspirez-vous notamment de la façon dont les articles de recherche reportent les résultats de régression.

Ressources utiles Concernant les aspects statistiques de ce projet, “Large Sample Estimation and Hypothesis Testing” de Newey et McFadden (*Handbook of Econometrics*, 1994); à propos du score de propension, la partie 3 du livre *Causal Inference for Statistics, Social and Biomedical Sciences* de Imbens et Rubin (2015).

1. Statistiques Descriptives et Double-Différences

1. Quelle est la proportion des villes dans lesquelles MAGA News est diffusée en 1998, 2000 et 2003 ? En quoi cela offre-t-il la possibilité de mesurer l'impact de MAGA News sur le vote ?
2. La variable `reppresfv2p1996` (resp. `reppresfv2p2000`) mesure la part du vote conservateur aux élections présidentielles de 1996 (resp. 2000). Calculer l'estimateur des différences-de-différences de l'effet de MAGA News.
3. L'estimateur précédent est-il crédible ? Effectuer un ou plusieurs tests pour justifier votre réponse.
4. Produire des statistiques descriptives sur les caractéristiques des villes permettant d'éclairer la réponse à la question précédente. Pour chacune des variables choisies, effectuer un test de différence des moyennes entre le groupe traité et le groupe contrôle.

2. Sélection et Score de Propension

Le but de cette section est d'étudier les facteurs déterminants de la présence de MAGA News dans une ville (*i.e.* de la sélection dans le traitement). Pour cette partie, vous devrez considérer des variables de trois natures différentes : (1) économiques, liées au marché local de la télévision câblée, (2) liées au vote et à la couleur politique des citoyens d'une ville, (3) sociodémographiques. Plus précisément, le but de cette partie est de montrer que l'implémentation de MAGA News dans une ville n'est pas liée à deux facteurs que sont

(i) la part du vote conservateur à l'élection de 1996 (`reppresfv2p1996`) et (ii) le taux de participation à l'élection de 1996 (`totpreslvpop1996`).

1. En quoi serait-il problématique que les deux variables citées précédemment expliquent la présence de MAGA News en 2000 ?
2. Proposer quatre spécifications différentes, avec un nombre croissant de variables, d'un modèle Logit où la variable expliquée est la présence de MAGA News dans une ville en 2000. La première spécification ne fera intervenir que la part du vote conservateur à l'élection de 1996 et le taux de participation à l'élection de 1996. Vous justifierez les variables de contrôle que vous ajoutez, mais pour ces variables uniquement, on ne demande pas de reporter les résultats de l'estimation. On prêtera attention, via des tests statistiques pertinents, à la qualité des modèles proposés.
3. D'après votre réponse à la question précédente, la part du vote conservateur à l'élection de 1996 et le taux de participation à l'élection de 1996 expliquent-ils l'implémentation de MAGA News ? Faire les tests correspondant.
4. Soit D_i la variable aléatoire qui vaut un si MAGA News est disponible dans la ville i en 2000 et zéro sinon, et X_i un vecteur aléatoire de dimension p mesurant des variables de contrôle.

(a) On définit la variable aléatoire:

$$W_i := \frac{\mathbb{E}(1 - D_i)}{\mathbb{E}(D_i)} \exp(X_i' \beta_0).$$

Montrez que si le score de propension est donné par un modèle Logit (*i.e.* $\mathbb{P}[D = 1|X] = [1 + \exp(-X' \beta_0)]^{-1}$) alors:

$$\mathbb{E}[X_i | D_i = 1] = \mathbb{E}[W_i X_i | D_i = 0]. \quad (1)$$

(b) Interpréter cette équation.

5. (a) Montrer que l'équation 1 peut s'écrire comme une condition de moment du type $\mathbb{E}[g(D_i, X_i, \beta_0)] = 0$ où vous préciserez la fonction g .

- (b) Proposer un estimateur GMM (Méthode des Moments Généralisée) de β_0 . Le calculer pour les quatre spécifications choisies à la question 2.

3. Estimation d'impact par Régression Linéaire

Le but de cette section est d'estimer l'impact causal de la présence de MAGA News sur le vote conservateur au moyen de régressions linéaires, afin d'obtenir un estimateur meilleur que celui des différences-de-différences obtenu dans la partie 1.

La variable `totpresvotes1996` donne le nombre de votes exprimés pour l'élection présidentielle de 1996. Dans cette section, on pondérera les résultats d'estimation par cette variable, de façon à interpréter les résultats pour l'électeur moyen, plutôt que pour la ville moyenne.

1. En prenant en compte les résultats statistiques de la partie 2, proposez cinq spécifications de la régression linéaire de la différence entre la part du vote conservateur à l'élection de 2000 et celle de 1996 (mesurée par `reppresfv2p1996`) sur la présence de MAGA News en 2000 (mesurée par `maganews2000`). Justifiez :
 - (a) la sélection des variables de contrôle,
 - (b) la présence ou non d'une variable mesurant la différence entre la part du vote conservateur entre deux élections avant 1996,
 - (c) l'utilisation d'effets fixes,
 - (d) le calcul de l'écart-type.

Le report des résultats d'estimation pour les coefficients associés aux variables de contrôle n'est pas demandé.

2. On note $Y_i(1)$ (resp. $Y_i(0)$) la variable aléatoire `reppresfv2p1996` qui mesure la différence entre la part du vote conservateur à l'élection de 2000 et celle de 1996 quand MAGA News est disponible dans la ville i (resp. quand MAGA News n'est pas disponible dans la ville i). On note la variable observée $Y_i = D_i Y_i(1) + (1 - D_i) Y_i(0)$. On définit le paramètre θ_0 :

$$\theta_0 = \mathbb{E}[Y_i | D_i = 1] - \mathbb{E}[Y_i | D_i = 0]. \quad (2)$$

On suppose que $Y \perp\!\!\!\perp D|X$ (*Hypothèse d'Indépendance Conditionnelle*, ou CIA).

- (a) Montrer que $\theta_0 = \mathbb{E}[Y_i(1) - Y_i(0)|D_i = 1]$.
 - (b) Proposer un estimateur de θ_0 basé sur 2, que l'on notera $\hat{\theta}$.
 - (c) Montrer que $\hat{\theta}$ est asymptotiquement normal. Donner son écart-type asymptotique.
 - (d) A partir de l'estimateur de β_0 calculé à la Q5 de la partie 2, calculer $\hat{\theta}$ ainsi que son écart-type. Comparer aux résultats obtenus par régression.
3. La disponibilité de MAGA News sur le réseau de télévision local impacte-t-il le vote conservateur ? Dans quelle mesure ? Proposez deux explications à ce phénomène (ou à son absence).

4. Hétérogénéité de l'effet et Test Placebo

1. On souhaite étudier l'existence d'effets hétérogènes de MAGA News. Pour une spécification choisie dans la partie 3, on fera interagir la présence de MAGA News avec :
- (a) Le nombre de chaîne disponibles sur le réseau de la ville en 2000,
 - (b) La part de la population résidant en zone urbaine,
 - (c) Le niveau d'éducation de la population.

On réalisera une régression pour chacun des trois items, la spécification (variable en continue, par tranches, indicatrices, etc.) est laissée libre.

2. On souhaite conduire un test placebo. Pour cela, on va regarder l'impact de la présence de MAGA News en 2000 sur l'évolution de la part du vote conservateur entre 1992 et 1996, ainsi qu'entre 1988 et 1992. Quel effet devrait-on observer ? Mettre en œuvre ce test pour deux spécifications choisies. Cela vous donne-t-il plus de confiance dans les résultats obtenus ?