

Langue de publication et performance en recherche Publier en français a-t-il un impact sur les performances bibliométriques des chercheurs francophones en science politique ?

Louis M. Imbeau et Mathieu Ouimet

Volume 31, numéro 3, 2012

La science politique francophone. Défis et trajectoires

URI : <https://id.erudit.org/iderudit/1014959ar>

DOI : <https://doi.org/10.7202/1014959ar>

[Aller au sommaire du numéro](#)

Éditeur(s)

Société québécoise de science politique

ISSN

1203-9438 (imprimé)

1703-8480 (numérique)

[Découvrir la revue](#)

Citer cet article

Imbeau, L. M. & Ouimet, M. (2012). Langue de publication et performance en recherche : publier en français a-t-il un impact sur les performances bibliométriques des chercheurs francophones en science politique ? *Politique et Sociétés*, 31(3), 39–65. <https://doi.org/10.7202/1014959ar>

Résumé de l'article

La science politique mondiale se communique principalement en anglais, alors qu'elle se pratique dans une multitude de langues. Dans le contexte où l'une d'elles domine dans les communications, il est pertinent de se questionner sur l'articulation entre une langue locale et la langue de communication internationale dans la pratique de la science politique. Nous avons voulu explorer l'impact de la langue de publication des travaux savants des chercheurs francophones de France et du Québec sur leur performance en recherche telle qu'elle est mesurée par Publish or Perish sur la base des travaux répertoriés par Google Scholar. Nous nous sommes demandé si le fait de publier en français diminuait la possibilité de reconnaissance d'un chercheur par ses pairs lorsque cette reconnaissance était fondée sur de tels indices. Nous rapportons ici les résultats de notre enquête. Nous montrons que la langue de publication semble avoir un impact systématique sur les mesures de performance des chercheurs. Ceux qui publient surtout en français publient moins et sont moins cités que les autres. D'autres facteurs sont reliés à la performance : le genre, l'expérience et le milieu.

Langue de publication et performance en recherche : publier en français a-t-il un impact sur les performances bibliométriques des chercheurs francophones en science politique ?

Louis M. Imbeau

Département de science politique, Université Laval
louis.imbeau@pol.ulaval.ca

Mathieu Ouimet

Département de science politique, Université Laval
mathieu.ouimet@pol.ulaval.ca

La science politique mondiale se communique principalement en anglais, alors qu'elle se pratique dans une multitude de langues. Dans le contexte où l'une d'elles domine dans les communications, il est pertinent de se questionner sur l'articulation entre une langue locale et la langue de communication internationale dans la pratique de la science politique. Nous avons voulu explorer l'impact de la langue de publication des travaux savants des chercheurs francophones de France et du Québec sur leur performance en recherche telle qu'elle est mesurée par Publish or Perish sur la base des travaux répertoriés par Google Scholar. Nous nous sommes demandé si le fait de publier en français diminuait la possibilité de reconnaissance d'un chercheur par ses pairs lorsque cette reconnaissance était fondée sur de tels indices. Nous rapportons ici les résultats de notre enquête. Nous montrons que la langue de publication semble avoir un impact systématique sur les mesures de performance des chercheurs. Ceux qui publient surtout en français publient moins et sont moins cités que les autres. D'autres facteurs sont reliés à la performance : le genre, l'expérience et le milieu.

Political science research is mainly published in English but it is produced in a variety of languages. The fact of one language being dominant in communications raises the issue of the articulation between a local language and the dominant communication language in the practice of political science. In this paper, we explore the impact of the publication language of Francophone political scientists of France and Québec on their performance in research as measured by Publish or Perish on the basis of Google Scholar citations. We

ask whether publishing in French decreases one's performance as it is measured by citation indices. We show that the language of publication seems to have a systematic impact on performance measures. When compared with Francophone scholars who publish mostly in English, those who do so mainly in French have less publications and are not cited as often by their peers. Other factors are related to performance: gender, experience, and the department to which a scholar is associated.

La science politique mondiale se communique principalement en anglais. S'il fallait s'en convaincre, il suffirait de consulter l'un des répertoires publiés par l'Association internationale de science politique (IPSA)¹. Par exemple, parmi les 43 sites internet répertoriés par IPSAPORTAL², 26 sont le fait d'instituts de recherche basés aux États-Unis. Les autres sont dispersés dans les pays suivants: Royaume-Uni, Italie, France, Allemagne, Argentine, Bulgarie, Hollande, Suède et Suisse. Tous y publient un résumé de leurs activités en anglais et seulement trois ont un nom dans une autre langue que l'anglais: le Deutsche Aussenpolitik.DE, l'Istituto Cattaneo et l'Unità di Ricerca sulla Governance Europea.

Mais la science politique se pratique dans une multitude de langues. Dans le contexte où l'une d'elles domine dans les communications, il est pertinent de se questionner sur l'articulation entre une langue locale et la langue de communication internationale dans la pratique de la science politique. Ainsi, nous avons voulu explorer l'impact de la langue de publication des travaux savants des chercheurs francophones de France et du Québec sur leur performance en recherche telle qu'elle est mesurée par Publish or Perish sur la base des travaux répertoriés par Google Scholar. Nous nous sommes demandé si le fait de publier en français diminuait la possibilité de reconnaissance d'un chercheur par ses pairs lorsque cette reconnaissance était fondée sur de tels indices. Nous rapportons ici les résultats de notre enquête.

Nous procéderons en trois étapes. Nous commencerons par recenser brièvement les écrits sur les déterminants de la performance des chercheurs en général. Nous détaillerons ensuite notre méthodologie de façon à donner au lecteur l'information utile pour reproduire nos résultats et surtout pour évaluer les menaces à la validité de nos conclusions qu'implique notre démarche. Enfin nous décrirons les résultats de notre analyse et nous les discuterons.

1. L'IPSA est officiellement bilingue anglais-français, mais les échanges qui y ont cours sont principalement en anglais.

2. (<http://www.ipsa.org/fr/ipsaportal-le-portail-de-l%E2%80%99aisp>).

Brève recension des écrits

La performance en recherche

La performance en recherche des professeurs d'université est étudiée depuis plusieurs décennies (Bland et coll., 2005 : 225). Les facteurs susceptibles d'influencer la productivité en recherche des professeurs sont multiples. S'intéressant à la productivité en recherche des professeurs dans les facultés de médecine, Carole J. Bland et ses collègues (2002) sont les auteurs de l'un des rares modèles conceptuels en la matière. La mise à l'épreuve empirique de leur modèle théorique (Bland et coll., 2005 : 233) suggère que la productivité en recherche des professeurs est associée aux facteurs suivants : la motivation pour la recherche, l'accès aux ressources nécessaires, l'octroi de récompenses pour les bons coups en recherche, le dégagement du temps suffisant pour réaliser des activités de recherche, la taille et l'expertise au sein de son unité, des signaux clairs de la part de la direction de l'unité du caractère prioritaire de la production en recherche, l'accès à du mentorat, la culture de recherche au sein de l'unité, la communication entre les membres de l'unité ainsi que le leadership exercé par le professeur dans son unité.

Éric Montpetit, André Blais et Martial Foucault (2008) ont étudié les facteurs susceptibles d'influencer le nombre de citations d'articles produits par 758 politologues répartis dans 37 départements de science politique au Canada. L'unité d'analyse qu'ils ont privilégiée est l'article scientifique plutôt que le professeur. Les données mobilisées dans le cadre de leur étude proviennent de la banque de données Web of Sciences (Social Sciences Citation Index). Leurs résultats montrent que plusieurs variables covarient avec le nombre de citations d'un article, dont le sous-champ disciplinaire, la région géographique où sont localisés les auteurs de l'article, le type de méthodologie employée, la taille du département, le nombre d'auteurs, le genre, l'indice d'impact de la revue et l'âge de l'article.

Par ailleurs, au mois d'avril 2011, le périodique spécialisé américain *PS: Political Science & Politics* publiait les résultats d'une étude empirique financée par l'Association américaine de science politique (APSA) portant sur les facteurs associés à la baisse ou à l'augmentation du nombre de publications chez les professeurs de science politique aux États-Unis. Cette étude empirique de Vicki Hesli et Jay Mook Lee (2011), de nature quantitative, s'appuie sur une recension assez exhaustive³ des facteurs explicatifs documentés dans les études sur le sujet. En combinant le nombre d'articles, de livres, de livres

3. Comme nous, ces auteurs n'ont pas mobilisé l'approche dite de la revue systématique de littérature (*systematic review*) qui permet de répertorier toutes les études publiées et non publiées sur une question précise, d'en faire la synthèse et d'en évaluer la qualité. La réalisation d'une revue systématique de littérature coûte au minimum 15 000 \$ en ressources humaines et nous ne disposons pas de ces fonds pour réaliser la présente étude.

édités et de chapitres en une seule variable dépendante, les auteurs de cette étude découvrent que les variables indépendantes suivantes covarient de façon significative avec le nombre total de publications : la durée des études doctorales (association négative), le nombre de cours enseignés (association négative), le nombre d'étudiants supervisés (association positive), le niveau de collégialité au sein de l'unité (association négative), l'accès à des ressources – la réduction de la charge d'enseignement, le temps d'auxiliaire de recherche, les fonds discrétionnaires, les fonds pour les déplacements – (association positive), le rang universitaire (association positive) et la participation à des conférences (association positive).

La principale force des études de Hesli et Lee (2011) et de Bland et coll. (2005) est d'avoir considéré une liste assez exhaustive de facteurs explicatifs. La prise en compte d'un grand nombre de facteurs explicatifs a été rendue possible en recourant au sondage. Le recours au sondage dans l'étude de la productivité en recherche de professeurs peut en revanche être une source de limites. En effet, l'ampleur de l'erreur de mesure causée par la capacité de mémorisation limitée des répondants ou encore par la volonté de certains d'entre eux d'enjoliver la situation (phénomène bien connu de désirabilité sociale) pourrait être très importante dans ces deux études. De plus, le recours exclusif au sondage rend difficile, voire impossible, la collecte des données sur le nombre total de citations des professeurs. L'étude de Montpetit, Blais et Foucault (2008) ne comporte pas ces limites car elle s'appuie sur l'analyse de données objectives.

Les méthodes d'analyse de données mobilisées constituent une autre source de limites potentielles de ces études. Bland et coll. (2005) ont utilisé la régression logistique binaire après avoir dichotomisé la variable dépendante reflétant le nombre d'articles publiés par les professeurs, de telle sorte qu'elle prend la valeur 1 lorsque le professeur a publié plus de cinq articles au cours des deux dernières années et 0 pour celui ou celle ayant publié cinq articles ou moins. Cette procédure prive les chercheurs d'une partie substantielle des informations contenues dans la variable dépendante originale. Pour leur part, Hesli et Lee (2011) ont utilisé la méthode de régression des moindres carrés après avoir tenté de linéariser, par transformation logarithmique, la distribution asymétrique de la variable reflétant le nombre de publications des professeurs, et ce, sans avoir considéré l'utilisation de la méthode de régression binomiale négative qui est pourtant la méthode appropriée lorsque la variable dépendante est mesurée à partir de données de dénombrement et que les valeurs de sa variance et sa moyenne sont significativement différentes (Green, 2005 : 723).

Il en est de même des résultats de l'étude de Montpetit, Blais et Foucault (2008) qui reposent sur la méthode de régression des moindres carrés et la méthode de régression Tobit censurée. Il est important de noter que la dis-

tribution de la variable dépendante dans cette étude comporte une proportion élevée de valeurs 0 (37,7%). Il serait intéressant de reproduire leurs analyses en utilisant la méthode adéquate dans ce type de situation, c'est-à-dire l'estimation d'un modèle de régression de type binomial négatif.

La langue de publication

La langue de publication des travaux scientifiques fait l'objet de débats depuis longtemps et la littérature sur cette question est phénoménale. On y décèle deux points de vue, celui des chercheurs anglophones et celui des autres.

Dans les années 1960, on s'inquiétait de l'accessibilité du savoir produit dans des langues autres que l'anglais. Dans un ouvrage publié en 1962, Christopher W. Hanson (cité dans Anderson, 1974) s'inquiétait de la possibilité que les résultats scientifiques publiés en russe, en chinois ou en japonais aillent croissant. À cette époque, D.N. Wood (1967 : 129) estimait que 50 % des écrits scientifiques et technologiques étaient publiés dans une autre langue que l'anglais. Wood montrait aussi qu'à l'exception des travaux publiés en français et, dans une moindre mesure en allemand, la majorité des scientifiques britanniques étaient incapables d'aborder les travaux en langue étrangère. Chez les spécialistes des sciences sociales, plus de 95 % des chercheurs américains et britanniques ne citaient jamais de travaux dans une autre langue que l'anglais (pour une revue, voir Yitzhaki, 1998).

Alors que les conclusions de Hanson, Wood et Anderson portaient principalement sur les sciences naturelles du point de vue des chercheurs anglophones, plusieurs travaux se sont concentrés sur les sciences sociales du point de vue des chercheurs non anglophones. Moshe Yitzhaki, par exemple, montre à partir de neuf revues de sociologie publiées aux États-Unis, au Royaume-Uni, en Allemagne et en France entre 1985 et 1994 que 99 % des références dans les revues américaines et britanniques sont en anglais alors que le pourcentage de références en allemand dans les revues allemandes est de 75 % et en français dans les revues françaises de 66 % (1998 : 243). Mais en utilisant une mesure plus nuancée prenant en compte l'importance de la production scientifique en anglais, en allemand ou en français (le *relative own-language preference indicator* [ROLPI]), Yitzhaki montre que les sociologues allemands ont le plus fort biais d'utilisation de la langue maternelle, suivis des sociologues français et, loin derrière, des sociologues américains et britanniques. Autrement dit, ces derniers citent les travaux en anglais parce que la plupart des travaux sont en anglais. Dans un article récent, Yves Gingras et Sébastien Mosbah-Natanson remarquent que les revues françaises dans le champ de la sociologie publient de plus en plus en anglais. D'une inquiétude des anglophones relative à l'accès au savoir produit et publié dans

une autre langue que l'anglais, nous sommes graduellement passés au désir des francophones de voir leurs travaux lus par les anglophones. Avec le développement des outils automatiques de traduction, il est permis de croire qu'à terme le problème de la langue de communication des travaux scientifiques n'existera plus.

Du point de vue des écrits savants sur l'impact de la langue dans la production scientifique, notre enquête se situe dans la perspective des chercheurs non anglophones, plus spécifiquement des politologues francophones de la France et du Québec. Quant aux écrits sur la performance en recherche, notre étude s'en démarque sur plusieurs aspects. Premièrement, les données analysées proviennent de Google Scholar, considéré comme étant la meilleure source de données bibliométriques pour les sciences humaines et sociales francophones (Kosmopoulos et Pumain, 2007 ; Jacso, 2008 ; Ouimet et coll., 2011). Contrairement à ISI Web of Science, Google Scholar répertorie des ouvrages et des chapitres d'ouvrages rédigés en français. Deuxièmement, les méthodes d'analyse mobilisées dans le cadre de l'étude correspondent aux standards économétriques en vigueur. En effet, le nombre de publications et le nombre de citations ont été modélisés en estimant un modèle binomial négatif plutôt qu'un modèle des moindres carrés ordinaires. Troisièmement, à notre connaissance, il s'agit de la première étude à explorer l'impact de la langue de publication des travaux savants des chercheurs francophones de France et du Québec sur leur performance en recherche telle qu'elle est mesurée par Publish or Perish sur la base des travaux répertoriés par Google Scholar. En revanche, nous admettons que, pour des raisons budgétaires, le nombre de facteurs explicatifs considérés dans notre étude est inférieur à celui des études recensées.

Méthodologie

Collecte des données

Les données mobilisées dans l'étude ont été collectées entre les mois de juillet et octobre 2011. La première phase de la collecte visait à constituer la liste des professeurs réguliers des départements de science politique de l'Université de Montréal (UdM), de l'Université du Québec à Montréal (UQAM), de l'Université McGill, de l'Université Concordia, de l'Université de Sherbrooke, de l'Université Bishop, de l'Université Laval, de l'Université d'Ottawa et de l'École nationale d'administration publique (ENAP). Cette liste a été constituée en consultant le site Internet de ces unités. Nous avons également tenté, malheureusement en vain, d'obtenir la liste des maîtres de conférences et des professeurs œuvrant dans les instituts d'études politiques et les universités françaises (CNU 04). Nous avons cependant réussi à obtenir la liste des chargés de recherche et des directeurs de recherche du Centre

national de recherche scientifique (CNRS – Section 40)⁴. Une banque de données Excel a été constituée; celle-ci répertorie le nom, le prénom ainsi que l’affiliation de tous les individus identifiés.

Deux étudiants de deuxième cycle (niveau maîtrise) ont été recrutés et formés pour exécuter les procédures nécessaires à la collecte des données bibliométriques de chaque professeur/chercheur. Chaque étudiant était doté d’un ordinateur fonctionnant sous le système d’exploitation Windows sur lequel avait été installé le logiciel gratuit Publish or Perish (PoP) développé par Anne-Wil Harzing⁵. Le prénom et le nom de chaque professeur/chercheur étaient saisis dans le moteur de recherche du logiciel PoP et la recherche était ensuite lancée. Le logiciel PoP utilise Google Scholar comme source de données. Après la recherche, l’étudiant procédait à un premier nettoyage manuel des résultats en éliminant les références jugées non pertinentes pour l’une des raisons suivantes: i) forte probabilité que le professeur/chercheur n’en soit pas l’auteur (ex.: produit par un homonyme); ii) publication non universitaire (ex.: article dans un quotidien, œuvre littéraire, etc.). Il est à noter que nous n’avons appliqué aucune restriction quant à l’année de la publication, au type de publication universitaire (article, livre, chapitre, rapport de recherche, article présenté lors d’une conférence) ou à l’ordre dans lequel les auteurs et coauteurs sont rapportés. Une fois ce premier nettoyage manuel terminé, l’étudiant procédait à la sauvegarde des résultats obtenus pour le professeur/chercheur sous le format «.csv». Le fichier «.csv» était ensuite transféré dans l’outil CLEANPoP, disponible gratuitement sur Internet⁶. Cet outil, développé par Audrey Baneyx pour le compte de l’Institut Francilien Recherche Innovation et Société (Baneyx, 2008), permet de nettoyer, de manière quasi automatique, les résultats de Publish or Perish. CLEANPoP présente à son utilisateur une liste de publications qui pourraient être des doublons et celui-ci est libre de choisir de définir ces publications comme étant des doublons. Cette procédure permet de ne pas comptabiliser comme distincts deux titres qui réfèrent à la même publication. Une fois le nettoyage terminé, CLEANPoP produit plusieurs indicateurs bibliométriques ainsi que la liste des titres qui ont servi de base aux calculs. Dans la prochaine étape, l’étudiant était chargé de saisir dans un fichier Excel les informations suivantes: le nombre total de publications, le nombre total de citations, l’indice *h*, l’année de publication du titre le plus ancien (en soustrayant ce nombre de 2011, on obtient l’âge scientifique, information nécessaire au calcul de plusieurs des indicateurs bibliométriques mobilisés dans l’étude), le nombre de titres en français, le nombre de titres en anglais

4. Il est à noter que la section 40 du CNRS est une section pluridisciplinaire regroupant des politologues (approximativement 2/3) et des sociologues du travail et des organisations (approximativement 1/3).

5. (<http://www.harzing.com/pop.htm>).

6. (<http://cleanpop.ifris.net/guide.html>).

et ceux dans une autre langue, ainsi que des renseignements portant sur les collaborateurs du professeur/chercheur (information que nous n'avons pas mobilisée aux fins de cet article).

Cette procédure nous a permis de constituer une banque de données de 515 individus répartis dans les organisations/unités suivantes: CNRS (276), Université de Montréal (34), Université du Québec à Montréal (37), Université McGill (30), Université Concordia (36), Université Laval (23), Université de Sherbrooke (11), Université Bishop (4), École nationale d'administration publique (16) Université d'Ottawa (48). Nous avons dû retrancher 11 professeurs/chercheurs pour lesquels il n'a malheureusement pas été possible de collecter des données fiables. Aux fins de cette publication portant sur le lien entre publier en français et performer sur le plan bibliométrique, les professeurs des universités McGill, Concordia et Bishop ont été exclus des analyses car ils sont anglophones pour la plupart. Les professeurs de l'Université de Sherbrooke ont également été exclus des analyses, car le nombre d'observations dans cette unité est plutôt limité. La taille de l'échantillon finalement mobilisé pour les analyses est de 434 observations.

Analyse des données

Afin de ne pas introduire de biais liés à la présentation sélective de résultats (Ioannidis, 2005 et 2008), la distribution et la variation de tous les indicateurs bibliométriques pour lesquels nous détenons de l'information sont présentés dans cette publication. Ces indicateurs bibliométriques sont les suivants: nombre de publications, taux annuel de publications, nombre de citations, taux annuel de citations, taux de citations par publication, indice h et taux annuel de l'indice h (ou quotient m). La définition opérationnelle de chaque indicateur est présentée au tableau 1.

Trois principaux calculs ont été réalisés dans l'étude, soit la prédiction de la valeur moyenne de chaque indicateur bibliométrique selon le genre, selon l'organisation/unité et selon différents points de pourcentage des publications rédigées en français. Ces prédictions ont été calculées en résolvant des équations de régression multiple dans lesquelles un indicateur bibliométrique était entré comme variable dépendante. Un modèle de régression de type binomial négatif a été estimé pour les variables dépendantes de dénombrement (nombre de publications, nombre de citations et indice h), tandis qu'un modèle de régression Tobit a été estimé pour les variables dépendantes de ratio. Les variables indépendantes suivantes ont été entrées dans les modèles de régression afin d'ajuster le mieux possible le calcul des prédictions: pourcentage des publications rédigées en français, âge scientifique, genre, ainsi que cinq variables dichotomiques reflétant l'organisation/unité du professeur/chercheur. Le nombre de publications a été ajouté à la liste des variables indépendantes dans une deuxième version du modèle

TABLEAU 1

Définitions opérationnelles des indicateurs bibliométriques

Indicateur	Définition
Nombre de publications	Dénombrement du nombre total de publications du professeur. Le nombre de publications a été calculé à l'aide de CLEANPoP ^a (source des données : Google Scholar ^b ; collecte des données : Publish or Perish – PoP ^c). Les publications retenues sont des productions universitaires (articles, livres, chapitres, articles de conférence, etc.). Dans la mesure du possible, les œuvres littéraires (ex. : romans) et les articles dans les médias ont été exclus. Cette variable de dénombrement comptabilise les publications à titre de premier auteur et de coauteur.
Taux annuel de publications	Nombre total de publications du professeur divisé par son âge scientifique. L'âge scientifique a été calculé en soustrayant l'année de publication de la première publication répertoriée à l'année de la collecte des données bibliométriques (2011).
Nombre de citations	Dénombrement du nombre total de citations des publications produites ou coproduites par le professeur. Le nombre de citations a été calculé à l'aide de CLEANPoP (source des données : Google Scholar ; collecte des données : Publish or Perish – PoP).
Taux annuel de citations	Nombre total de citations divisé par l'âge scientifique.
Taux de citations par publication	Nombre total de citations d'un professeur divisé par le nombre total de ses publications.
Indice <i>h</i>	Un professeur a un indice <i>h</i> de <i>i</i> si <i>i</i> de ses publications ont été citées au moins <i>i</i> fois chacune et que ses autres publications (s'il en a d'autres) ont été citées moins de <i>i</i> fois chacune (Hirsch, 2005). Par exemple, pour avoir un indice <i>h</i> de 4, il faut que le professeur ait au moins 4 publications citées au moins 4 fois chacune, et que ses autres publications (s'il en a d'autres) aient été citées moins de 4 fois chacune. L'originalité de cet indicateur réside dans le fait qu'il combine en un seul indice l'information quant aux publications et aux citations. L'indice <i>h</i> a été calculé avec CLEANPoP (source des données : Google Scholar ; collecte des données : Publish or Perish – PoP).
Taux annuel de l'indice <i>h</i>	Mieux connu sous l'appellation de « quotient <i>m</i> », cet indicateur a été inventé par Hirsch (2005). Il s'agit simplement de l'indice <i>h</i> divisé par l'âge scientifique.

^a L'outil CLEANPoP a été développé par Audrey Baneyx pour le compte de l'Institut Francilien Recherche Innovation et Société. Cet outil a pour objectif de nettoyer, de manière automatique, les résultats de l'outil Publish or Perish (PoP) développé par Anne-Wil Harzing. Lien vers CLEANPoP : (<http://cleanpop.ifris.net/guide.html>).

^b Google Scholar est une base de données bibliométriques en concurrence avec d'autres bases bien connues comme ISI Web of Knowledge. Google Scholar est la meilleure source de données pour les sciences humaines et sociales francophones. Lien vers Google Scholar : (<http://scholar.google.com/>).

^c L'outil Publish or Perish (PoP) a été développé par Anne-Wil Harzing. Ce logiciel gratuit fonctionnant sous Windows permet d'affiner les recherches bibliométriques dans Google Scholar et de nettoyer les résultats. Ce logiciel peut également sauvegarder les données bibliométriques de chaque professeur. Ces données sauvegardées servent ensuite de base de calcul à l'outil CLEANPoP grâce auquel il est possible notamment d'éliminer les doublons et de réduire les risques liés au phénomène de la double publication. Lien vers Publish or Perish : (<http://www.harzing.com/pop.htm>).

estimé pour prédire le nombre de citations, tandis que le taux annuel de publications a été ajouté comme variable indépendante dans une deuxième version du modèle visant à prédire le taux annuel de citations. Ces ajustements supplémentaires ont permis d'améliorer notre compréhension des différences observées entre les organisations/unités quant au nombre moyen de citations et au taux annuel moyen de citations.

Résultats

Analyses descriptives

Les statistiques descriptives des variables dépendantes et indépendantes sont présentées au tableau 2. La distribution de tous les indicateurs bibliométriques s'éloigne de façon significative de la loi normale, car le coefficient d'asymétrie dépasse largement la valeur 0 et le coefficient d'aplatissement

TABLEAU 2

Statistiques descriptives

	Échelle	N	Min	Max	Moyenne ou %	Écart- type	Asymé- trie	Aplatisse- ment
Indicateurs bibliométriques								
Nombre de publications	Dénombrement	434	2	387	37,8	38,1	3,7	27,0
Taux annuel de publications	Ratio	434	0,06	13,5	1,7	1,3	3,0	20,6
Nombre de citations	Dénombrement	434	0	6546	251,4	587,2	6,3	53,9
Taux annuel de citations	Ratio	434	0	145,47	9,7	18,0	4,5	26,71
Taux de citations par publication	Ratio	434	0	39,16	4,5	5,4	3,2	16,5
Indice <i>h</i>	Dénombrement	434	0	38	5,9	4,8	2,4	12,7
Taux annuel moyen de l'indice <i>h</i>	Ratio	434	0	2,89	0,3	0,2	4,4	44,8
Variables d'ajustement (indépendantes)								
Pub – français (%)	Ratio	434	0	100	57,9	28,4	-0,4	2,2
Âge scientifique	Dénombrement	434	2	57	22,1	10,7	0,3	2,3
Genre	Dichotomique	276 (H) 158 (F)	1 (H)	2 (F)	63,6% (H) 36,4% (F)			
U. Montréal	Dichotomique	34	0	1	7,8%			
U. Laval	Dichotomique	23	0	1	5,3%			
UQAM	Dichotomique	37	0	1	8,5%			
ENAP	Dichotomique	16	0	1	3,7%			
U. Ottawa	Dichotomique	48	0	1	11,1%			
CNRS	Dichotomique	276	0	1	63,6%			

Note: Une valeur extrême affichant un taux annuel de publications de 70, un taux annuel de citations de 896 et un taux annuel de l'indice *h* (quotient *m*) de 13 a été exclue du calcul des statistiques descriptives ainsi que des autres analyses réalisées dans le cadre de cette étude.

la valeur 3. Ce résultat n'est pas surprenant, car il est bien documenté que la productivité scientifique n'est pas distribuée normalement parmi les chercheurs d'une même discipline, non seulement dans les sciences physiques et en chimie (Lotka, 1926), mais également dans les sciences humaines (Murphy, 1973). En termes statistiques, ce phénomène s'illustre par l'observation suivante: la distribution de la plupart des indicateurs bibliométriques calculés à partir du nombre de publications et/ou de citations correspond davantage à la loi de puissance (distribution élitiste) qu'à la loi normale.

En moyenne, les chercheurs de notre échantillon ont publié 57,9 % ($\pm 28,4$) de leurs textes en français. Le temps moyen qui s'est écoulé depuis la date de la plus ancienne publication jusqu'en 2011 est de 22,1 ($\pm 10,7$) années (correspondant à l'âge scientifique moyen). Une minorité de professeurs/chercheurs sont des femmes (36,4 %). Les professeurs/chercheurs ont produit en moyenne 37,8 ($\pm 38,1$) publications correspondant à un taux annuel moyen de 1,7 ($\pm 1,3$) publication. Le nombre moyen de citations est de 251,4 ($\pm 587,2$) pour un taux annuel moyen de 9,7 ($\pm 18,0$) citations. Une publication d'un professeur/chercheur est citée en moyenne 4,5 ($\pm 5,4$) fois. L'indice *h* moyen est de 5,9 ($\pm 4,8$) et il progresse annuellement en moyenne de 0,3 ($\pm 0,2$) point.

Le tableau 3 présente les moyennes des indicateurs bibliométriques et du pourcentage de publications en français pour chaque unité de rattachement. Afin de mieux illustrer l'importance de l'effet « étoile », ce tableau présente les moyennes calculées avec *et* sans les vedettes (c.-à-d., les professeurs dont le nombre total de publications ou de citations était de trois écarts-types au-dessus de la moyenne de leur organisation/unité). Le Département de science politique de l'Université de Montréal est celui où l'effet « étoile » est le plus marqué. En effet, en retirant de l'échantillon les professeurs André Blais et Richard Nadeau, ce département voit le nombre moyen des publications de ses professeurs passer de 71,8 à 53,9 et le nombre moyen de citations passer de 606,7 à 279,6! En ce qui concerne l'École nationale d'administration publique, l'arrivée récente du professeur Jean-Louis Denis a fait passer le nombre moyen de citations par professeur de 95,4 à 233! Il est également intéressant de noter que le Département de science politique de l'Université Laval est le seul à ne pas avoir d'« étoile » au sens où ce concept a été opérationnalisé dans la présente étude. En retirant les deux vedettes du Département de science politique de l'Université de Montréal, les professeurs du Département de science politique de l'Université Laval affichent des performances moyennes comparables à celles de leurs collègues de l'Université de Montréal.

Le tableau 3 révèle également le rang agrégé de chaque unité. Le rang agrégé a le mérite de classer les organisations/unités en tenant compte des

sept indicateurs bibliométriques⁷. En considérant l'ensemble des indicateurs et en éliminant l'effet « étoile », le Département de science politique de l'Université de Montréal se situe au premier rang, suivi, dans l'ordre, du Département de science politique de l'Université Laval (2^e), du CNRS (3^e), de l'École d'études politiques de l'Université d'Ottawa (4^e), de l'École nationale d'administration publique (5^e) et du Département de science politique de l'UQAM qui occupe le dernier rang⁸. La surprise réside du côté du CNRS. Nous nous attendions à ce que les chercheurs de cette unité performant beaucoup mieux que tous les autres puisqu'ils devraient avoir une charge d'enseignement et d'administration en moyenne plus légère que leurs collègues canadiens. Or ils arrivent au troisième rang dans le classement général selon le rang agrégé. C'est loin de ce que devraient réaliser des chercheurs qui peuvent consacrer la majeure partie de leur temps de travail à la recherche. Nous y reviendrons plus loin.

Par ailleurs, dans la dernière rangée du tableau 3, nous constatons une grande variation entre les unités eu égard au pourcentage de publications en français⁹. Sur le plan statistique¹⁰, les chercheurs du CNRS publient plus en français que les professeurs de l'Université de Montréal et de l'Université Laval. Les professeurs de l'université d'Ottawa, une université bilingue, publient en moyenne peu en français. Dans les deux autres unités, l'ENAP et l'UQAM, on publie en français dans une proportion semblable à celle du CNRS. La question centrale de notre étude fait le lien entre la langue de publication et la performance bibliométrique. Nous y répondons à l'aide d'un ensemble d'analyses de régression.

7. Il est possible de créer un classement général en additionnant les rangs de chaque unité de rattachement (minimum théorique = 7 ; maximum théorique = 42). À partir de cette somme des rangs, on peut établir le rang agrégé selon lequel l'organisation/unité qui occupe le premier rang est celle dont la somme des rangs est la plus petite. À l'inverse, l'organisation/unité occupant le dernier rang est celle dont la somme des rangs est la plus grande. Il est à noter que le rang agrégé accorde une pondération égale à tous les indicateurs bibliométriques.

8. Une analyse de variance (test *F*) a été réalisée pour chaque indicateur bibliométrique afin de détecter de possibles différences entre les moyennes de chaque unité/organisation. Tous les tests se sont avérés statistiquement significatifs au seuil de 1, 5 ou 10 %, à l'exception de celui effectué pour le « taux annuel de citations » ($p = 0,12$). Le test « pairé » de Bonferroni a permis de détecter peu de différences significatives entre des paires d'organisations/unités. Par exemple, à aucun moment les moyennes de l'Université de Montréal et de l'Université Laval ne se sont avérées significativement différentes. Ces résultats sont probablement attribuables au nombre limité d'observations dans toutes les organisations/unités à l'exception du CNRS. En d'autres termes, nous ne disposons pas de la puissance statistique nécessaire pour détecter des différences significatives. En revanche, cela ne signifie pas que certaines différences observées ne sont pas importantes sur le plan pratique.

9. Test *F* (ANOVA) statistiquement significatif au seuil de 1 %.

10. Test pairé de Bonferroni.

TABLEAU 3

Moyennes des indicateurs bibliométriques et du pourcentage de publications en français selon l'unité de rattachement

	CNRS		UdM		Laval		UQAM		ENAP		Ottawa	
	Étoiles ^a	Étoiles ^b	Étoiles	Étoiles	Étoiles ^c	Étoiles	Étoiles	Étoiles	Étoiles	Étoiles	Étoiles	Étoiles
Nombre de publications	34,4	31,6	71,8	53,9	55,6	55,6	39,5	36,1	29,5	22,9	26,4	22,6
Taux annuel de publications	1,6	1,6	2,7	2,7	2,3	2,3	1,5	1,4	1,5	1,5	1,8	1,8
Nombre de citations	244,1	191,7	606,7	279,6	252,1	252,1	151,8	129,0	233,0	95,4	124,2	101,4
Taux annuel de citations	9,7	7,8	17,7	10,3	9,3	9,3	4,4	3,8	11,6	6,7	7,4	6,2
Taux de citations par publication	5,0	4,8	4,7	4,0	3,9	3,9	2,8	2,7	5,3	4,4	3,3	3,1
Indice <i>h</i>	5,7	5,4	9,0	7,2	6,7	6,7	5,0	4,8	4,9	3,8	5,1	4,6
Taux annuel moyen de l'indice <i>h</i> (quotient <i>m</i>)	0,27	0,26	0,35	0,32	0,30	0,30	0,19	0,19	0,27	0,23	0,38	0,36
Somme des rangs des 7 indicateurs ^d	22	22	10	11	19	16	37	35	26	33	31	30
Rang agrégé ^e	3	3	1	1	2	2	6	6	4	5	5	4
Pourcentage de publications en français	64,0	64,3	44,2	46,0	54,5	54,5	64,6	63,8	62,9	65,2	28,7	28,9

Note: Une valeur extrême affichant un taux annuel de publications de 70, un taux annuel de citations de 896 et un taux annuel de l'indice *h* (quotient *m*) de 13 a été exclue des analyses.

^a Résultats avant élimination de l'effet « étoile ».

^b Afin d'éliminer l'effet « étoile », les professeurs dont le nombre total de publications ou de citations était de 3 écarts-types au-dessus de la moyenne de leur organisation/unité ont été exclus de l'analyse.

^c Aucune « étoile » n'a été identifiée au Département de science politique de l'Université Laval. En d'autres termes, aucun professeur n'affichait un nombre total de publications, de citations ou un score sur l'indice *h* supérieur à 3 écarts-types au-dessus de la moyenne départementale.

^d Dans un premier temps, le rang (de 1 à 6) de chaque organisation/unité a été établi pour chaque indicateur bibliométrique. Dans un deuxième temps, les sept chiffres (correspondant au rang de chaque organisation/unité sur chaque indicateur) ont été additionnés. Le résultat de cette addition constitue la somme des rangs.

^e Le rang agrégé a été calculé à partir de la somme des rangs. L'organisation/unité qui occupe le premier rang est celle dont la somme des rangs est la plus petite. À l'inverse, l'organisation/unité occupant le dernier rang est celle dont la somme des rangs est la plus grande. Il est à noter que le rang agrégé accorde une pondération égale à tous les indicateurs bibliométriques.

TABLEAU 4

Résultats des régressions (coefficients de régression et niveaux de signification)

Méthode d'estimation→	Modèles de régression ^a								
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Variable dépendante→	<i>Binomial négatif</i>	<i>Tobit</i>	<i>Binomial négatif</i>	<i>Binomial négatif</i>	<i>Tobit</i>	<i>Tobit</i>	<i>Tobit</i>	<i>Binomial négatif</i>	<i>Tobit</i>
	N ^{bre} de publications	Taux annuel de publications	N ^{bre} de citations	N ^{bre} de citations	Taux annuel de citations	Taux annuel de citations	Taux de citations par publication	Indice <i>h</i>	Taux annuel de l'indice <i>h</i>
Pub – français (%)	-0,01***	-0,01***	-0,02***	-0,01***	-0,19***	-0,08***	-0,05***	-0,01***	-0,002***
Âge scientifique	0,05***		0,09***	0,03***			0,16***	0,04***	
Genre (H=1 ; F=2)	-0,16***	-0,29**	-0,20*	-0,10	-2,56	-0,17	-0,28	-0,07	-0,01
U. Montréal ^b	0,30***	0,76***	-0,11	-0,48***	3,83	-2,37	-1,89**	0,05	0,03
U. Laval ^b	0,25*	0,52**	-0,37	-0,50**	-2,33	-6,59**	-2,14**	-0,06	0,01
UQAM ^b	-0,16	-0,19	-1,03***	-0,73***	-5,84*	-4,40*	-3,18***	-0,35***	-0,08**
ENAP ^b	-0,20	-0,20	0,10	0,23	0,91	2,56	0,16	-0,16	-0,01
U. Ottawa ^b	-0,23**	-0,28	-0,90***	-0,52**	-9,67***	-7,35***	-2,60***	-0,16	0,003
N ^{bre} de publications				0,03***					
Taux annuel de pub.						8,13***			
Constante	3,09***	2,90***	4,85***	3,89***	25,28***	1,59	5,24***	1,52***	0,47***
N ^{bre} de cas	434	434	434	434	434	434	434	434	434
Test du ratio de vraisemblance - Khi-deux (Degrés de liberté)	292,91 (8)***	68,04 (7)***	275,52 (8)***	457,52 (9)***	46,95 (7)***	225,70 (8)***	71,39 (8)***	215,12 (8)***	66,63 (7)***

*** = Significatif au seuil de 1 % , ** = au seuil de 5 % , * = au seuil de 10 % , test bilatéral.

^a Une valeur extrême affichant un taux annuel de publications de 70, un taux annuel de citations de 896 et un taux annuel de l'indice *h* (quotient *m*) de 13 a été exclue des analyses.

^b Le CNRS (France) est la catégorie de référence.

Analyses de régression

Les résultats des analyses de régression sont présentés au tableau 4. Le résultat du test du ratio de vraisemblance de chaque modèle de régression est significatif au seuil de 1 %, ce qui signifie que tous les modèles estimés sont statistiquement pertinents pour expliquer/prédire les valeurs de la variable dépendante¹¹. Les résultats présentés au tableau 4 révèlent une association négative et statistiquement significative au seuil de 1 % entre le pourcentage de publications rédigées en français et chaque indicateur bibliométrique, et ce, en ajustant les résultats selon l'âge scientifique, le genre et l'organisation/unité.

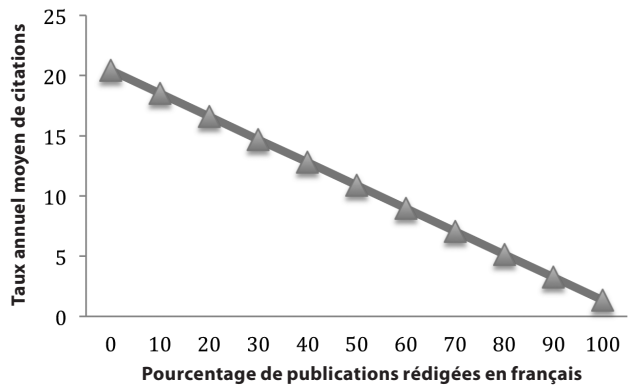
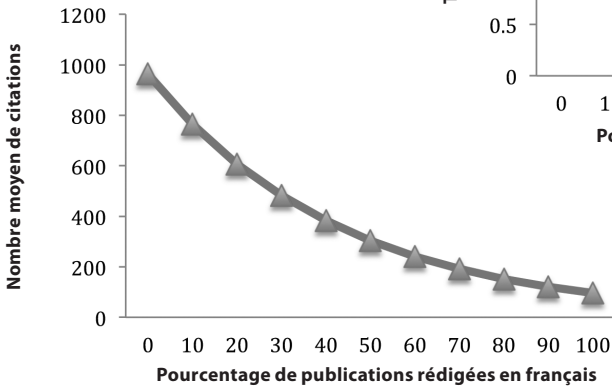
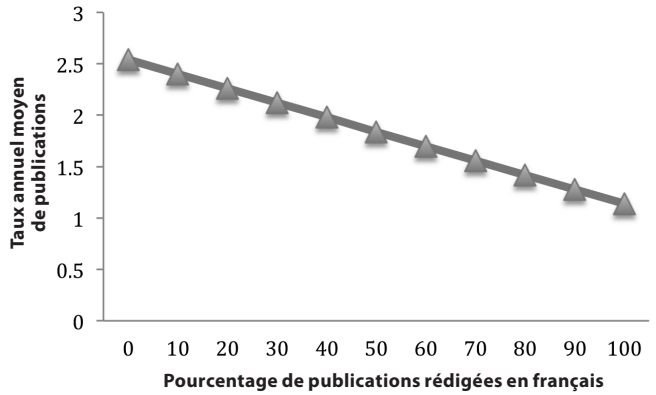
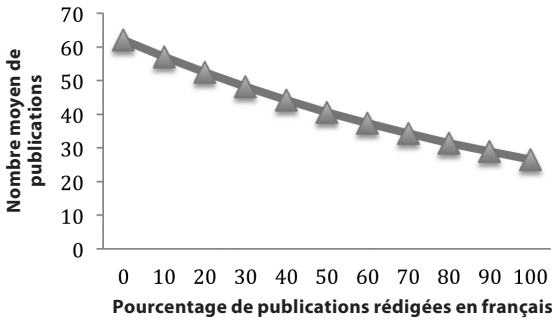
Les courbes de la figure 1 permettent de mieux saisir l'ampleur de cette association statistique. Ces courbes ont été produites en calculant la valeur prédite moyenne de la variable dépendante de chaque modèle de régression selon différentes valeurs de la variable reflétant le pourcentage de publications rédigées en français. La valeur prédite moyenne du nombre de publications d'un professeur/chercheur dont 10 % des publications ont été rédigées en français est de 57, tandis que celle d'un collègue ayant rédigé 90 % de ses publications en français est de seulement 28,9, une différence de plus de 28 publications ! Un professeur/chercheur ayant rédigé 10 % de ses publications en français aurait en moyenne 765,2 citations, ce nombre moyen étant de seulement 120,6 citations pour celui ou celle qui aurait rédigé 90 % de ses publications en français. L'ampleur du phénomène est sensiblement la même pour les autres indicateurs bibliométriques.

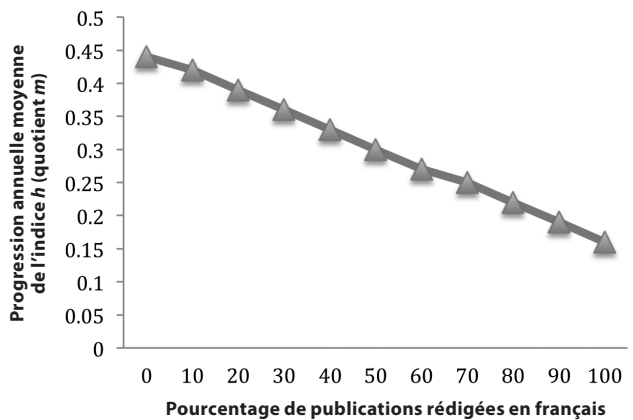
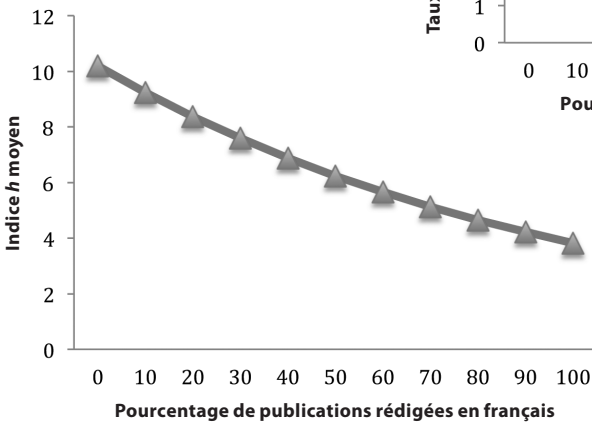
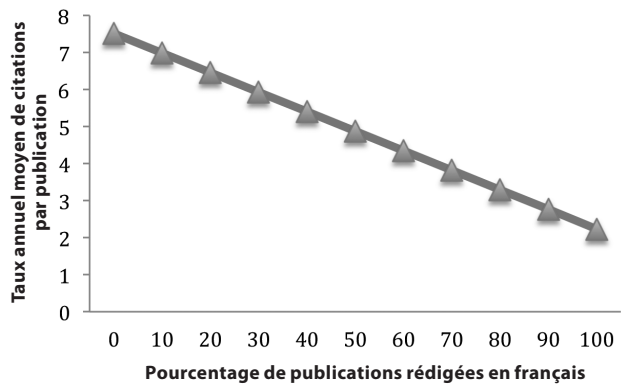
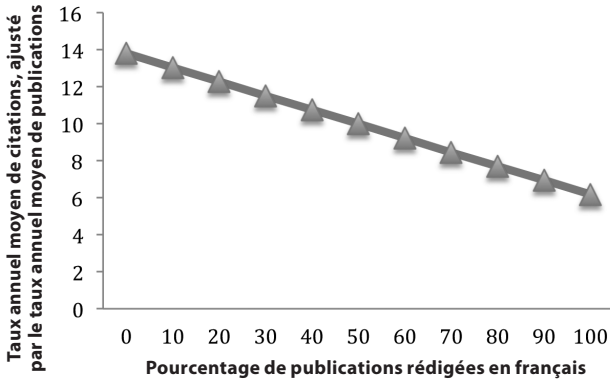
Les résultats des analyses de régression présentés au tableau 4 montrent également (et sans surprise) que l'âge scientifique est positivement et significativement associé au nombre de publications, au nombre de citations, au taux de citations par publication et à l'indice h . Il est à noter que l'âge scientifique ne figure pas dans la liste des variables indépendantes mobilisées pour expliquer/prédire le taux annuel de publications, le taux annuel de citation et le taux annuel de l'indice h (modèles 2, 5, 6 et 9), car l'âge scientifique est le dénominateur de ces indicateurs. Les résultats de régression révèlent en outre qu'en moyenne, et en ajustant pour le pourcentage des publications rédigées en français, l'âge scientifique et l'organisation/unité, les femmes publient moins et à un rythme annuel moins soutenu que les hommes. En ce qui concerne le nombre et le taux annuel de citations, les

11. Le Khi-deux et le niveau de signification sont conçus pour généraliser des résultats d'un échantillon à la population dont il a été tiré. Nous sommes conscients que nos données ne sont pas des données d'échantillon mais des données de population et que, par conséquent, l'utilisation des tests d'inférence statistique n'est pas strictement appropriée. Nous rapportons quand même ces statistiques parce que leur usage est très répandu pour les données de population et parce qu'elles peuvent être utilisées comme un indicateur de l'importance *relative* d'une relation ou d'un modèle dans un contexte où le nombre de cas est stable. Dans cette perspective, nous considérons les coefficients non significatifs comme étant nuls.

FIGURE 1

Prédiction de la valeur moyenne de la variable dépendante de chaque modèle de régression calculée selon différentes valeurs de la variable « pourcentage de publications rédigées en français »





résultats présentés au tableau 4 montrent qu'en ajustant l'analyse selon le nombre et le taux annuel de publications (modèles 4 et 6), la différence liée au genre n'est plus statistiquement significative. Aucune différence statistiquement significative entre les hommes et les femmes n'est observée dans la modélisation du taux de citations par publication, de l'indice h et du taux annuel de l'indice h .

Les comparaisons homme/femme sont mieux illustrées au tableau 5 qui présente la prédiction de la valeur moyenne de la variable dépendante pour chaque modèle de régression calculée selon les deux valeurs de la variable reflétant le genre. Bien que statistiquement significative, la différence homme/

TABLEAU 5

Prédiction de la valeur moyenne de la variable dépendante pour chaque modèle de régression calculée selon les deux valeurs de la variable « genre »

	Homme	Femme	Seuil de signification statistique
Nombre moyen de publications (modèle 1) ^{d,h}	40,16	34,22	**
Taux annuel moyen de publications (modèle 2) ^{e,i}	1,83	1,54	**
Nombre moyen de citations (modèle 3) ^{d,h}	292,19	237,88	*
Nombre moyen de citations (modèle 4) ^{f,h}	375,64 ^a	338,08 ^b	NS
Taux annuel moyen de citations (modèle 5) ^{e,i}	10,31	7,75	NS
Taux annuel moyen de citations (modèle 6) ^{g,i}	9,45	9,28	NS
Taux moyen de citations par publication (modèle 7) ^{d,i}	4,56	4,28	NS
Indice h moyen (modèle 8) ^{d,h}	6,00	5,62	NS
Taux annuel moyen de l'indice h (modèle 9) ^{e,i}	0,29	0,27	NS

^a Valeur calculée en excluant les chercheurs « étoiles ». En incluant les chercheurs « étoiles », l'ajout du nombre de publications dans la liste des variables indépendantes génère un nombre moyen de citations aberrant (plus de 29 000 citations!).

^b Valeur calculée en excluant les chercheurs « étoiles ». En incluant les chercheurs « étoiles », l'ajout du nombre de publications dans la liste des variables indépendantes génère un nombre moyen de citations aberrant (plus de 26 000 citations!).

^c Une valeur extrême affichant un taux annuel de publications de 70, un taux annuel de citations de 896 et un taux annuel de l'indice h (quotient m) de 13 a été exclue des analyses.

^d Valeur prédite (moyenne) ajustée en fonction de l'âge scientifique, de la proportion de publications de langue française et de l'organisation/unité.

^e Valeur prédite (moyenne) ajustée en fonction de la proportion de publications de langue française et de l'organisation/unité.

^f Valeur prédite (moyenne) ajustée en fonction du nombre de publications, de la proportion de publications de langue française, de l'âge scientifique et de l'organisation/unité.

^g Valeur prédite (moyenne) ajustée en fonction du taux moyen annuel de publications, de la proportion de publications de langue française et de l'organisation/unité.

^h La valeur prédite a été calculée après avoir estimé une équation de régression de type binomial négatif.

ⁱ La valeur prédite a été calculée après avoir estimé une équation de régression de type Tobit.

*** = Significatif au seuil de 1 %, ** = au seuil de 5 %, * = au seuil de 10 %, test bilatéral. Seuil associé à la valeur p du coefficient de la variable « genre » dans le modèle de régression correspondant.

femme quant au nombre et au taux annuel de publications n'est pas très grande (une moyenne de 40,2 publications pour les hommes contre 34,2 pour les femmes; un taux annuel de publications de 1,8 pour les hommes contre 1,5 pour les femmes). Nous observons également que les écarts homme-femme du nombre moyen de citations (ajusté par le nombre de publications), du taux annuel de citations, du taux moyen de citations par publication, de l'indice h et du taux annuel moyen de l'indice h sont non significatifs.

Le tableau 6 présente la prédiction de la valeur moyenne de la variable dépendante de chaque modèle de régression calculée selon les différentes organisations/unités. Il est important de noter que le CNRS est la catégorie de comparaison. Sur le plan des publications, ce sont les départements de science politique de l'Université de Montréal et de l'Université Laval qui affichent le nombre moyen de publications et le taux annuel moyen de publications les plus élevés, et ce, après avoir ajusté les analyses selon l'âge scientifique, le genre et le pourcentage des publications rédigées en français (modèles 1 et 2).

En ce qui concerne le nombre de citations non ajusté selon le nombre de publications (modèle 3), deux organisations/unités affichent un score moyen statistiquement plus faible que le CNRS. En effet, les politologues de l'Université du Québec à Montréal et de l'Université d'Ottawa ont en moyenne été beaucoup moins cités que les chercheurs de la section 40 du CNRS, de l'Université de Montréal, de l'Université Laval et de l'École nationale d'administration publique. En ajustant l'analyse selon le nombre de publications (modèle 4), l'Université de Montréal et l'Université Laval deviennent elles aussi des unités dont les professeurs sont significativement moins cités que leurs collègues du CNRS et de l'École nationale d'administration publique¹². En effet, le nombre moyen de citations pour les professeurs de l'ENAP et du CNRS est respectivement de 672,9 et 505,3, tandis que les politologues de l'Université de Montréal (320,1), de l'Université Laval (294,8), de l'Université du Québec à Montréal (215,7) et de l'Université d'Ottawa (305,6) sont en moyenne nettement moins cités une fois qu'on a purgé l'effet du nombre de publications. Quant au taux annuel moyen de citations, que l'on ajuste ou non l'analyse en fonction du taux annuel de publications, les résultats montrent une infériorité moyenne significative des politologues de l'Université du Québec à Montréal et de l'Université d'Ottawa. De plus, lorsqu'on ajuste l'analyse en fonction du taux annuel de publications, les résultats font ressortir une infériorité moyenne significative des politologues de l'Université Laval en ce qui concerne le taux annuel moyen de citations.

12. Valeur calculée en excluant les chercheurs « étoiles ». En incluant les chercheurs « étoiles », l'ajout du nombre de publications dans la liste des variables indépendantes génère un nombre moyen de citations aberrant.

TABLEAU 6

Prédiction de la valeur moyenne de la variable dépendante de chaque modèle de régression calculée selon différentes organisations/unités

Modèle de régression utilisé →	N ^{bre} moyen- publications ^{c,g}	Taux annuel moyen- publications ^{d,h}	N ^{bre} moyen- citations ^{c,g}	N ^{bre} moyen- citations ^{a,e,g}	Taux annuel moyen de citations ^{c,h}	Taux annuel moyen de citations ^{f,h}	Taux moyen de citations/ publication ^{c,h}	Indice <i>h</i> moyen ^{c,g}	Taux annuel moyen de l'indice <i>h</i> ^{d,h}
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Université de Montréal	50,82***	2,46***	301,17	320,15***	14,57	8,64	3,37**	6,47	0,31
Université Laval	48,47*	2,22**	231,98	294,84***	8,40	4,42**	3,12**	5,82	0,29
Université du Québec à Montréal	32,06	1,50	120,02***	215,75***	4,90*	6,62*	2,09***	4,33***	0,20**
ENAP	30,65	1,49	372,56	672,95	11,64	13,57	5,43	5,26	0,27
Université d'Ottawa	29,84**	1,41	137,65***	305,57***	1,07***	3,66***	2,70***	5,28	0,29
CNRS (Catégorie de référence)	37,64	1,69	337,74	505,32	10,74	11,02	5,27	6,18	0,28

^a Valeur calculée en excluant les chercheurs « étoiles ». En incluant les chercheurs « étoiles », l'ajout du nombre de publications dans la liste des variables indépendantes génère un nombre moyen de citations aberrant.

^c Valeur prédite (moyenne) ajustée en fonction du genre, de l'âge scientifique et de la proportion de publications de langue française.

^d Valeur prédite (moyenne) ajustée en fonction du genre et de la proportion de publications de langue française.

^e Valeur prédite (moyenne) ajustée en fonction du nombre de publications, du genre et de la proportion de publications de langue française.

^f Valeur prédite (moyenne) ajustée en fonction du taux moyen annuel de publications, du genre et de la proportion de publications de langue française.

^g La valeur prédite a été calculée après avoir estimé une équation de régression de type binomial négatif.

^h La valeur prédite a été calculée après avoir estimé une équation de régression de type Tobit.

*** = Significatif au seuil de 1 %, ** = au seuil de 5 %, * = au seuil de 10 %, test bilatéral. Seuil associé à la valeur *p* du coefficient des variables binaires reflétant l'unité/l'organisation dans le modèle de régression correspondant.

Les résultats présentés au tableau 6 révèlent également que les professeurs de l'École nationale d'administration publique et les chercheurs du CNRS sont en moyenne plus efficaces que leurs collègues des autres organisations/unités à générer des citations à partir de leurs publications, car ils affichent un taux moyen de citations par publication plus élevé que leurs collègues des autres organisations. Les politologues les moins efficaces en la matière sont ceux de l'Université d'Ottawa, de l'Université Laval et de l'Université du Québec à Montréal.

En ce qui concerne l'indice h , les résultats montrent que les politologues de l'Université de Montréal, de l'Université Laval, de l'École nationale d'administration publique, de l'Université d'Ottawa et du CNRS performant en moyenne aussi bien, et significativement mieux que leurs collègues de l'Université du Québec à Montréal. Enfin, sur le plan du taux annuel moyen de progression de l'indice h (quotient m), les politologues de l'Université du Québec à Montréal affichent également un taux moyen significativement moins élevé que leurs collègues des autres unités/organisations.

Discussion

Une relation causale ?

Il serait facile d'utiliser les résultats empiriques que nous venons de rapporter pour soutenir l'hypothèse que le fait de publier en français nuise à la performance des chercheurs. La relation négative et significative que nous avons exposée semble en effet relativement importante et robuste. D'ailleurs, la covariation n'est-elle pas une condition nécessaire de la causalité ? Mais une telle interprétation des résultats ne résiste pas à une réflexion sur la relation causale entre les deux variables. La covariation n'est pas *suffisante* pour établir la causalité.

Car la relation langue-performance pourrait bien être fallacieuse. Il est certes possible que langue de publication *et* performance soient l'effet d'une troisième variable, les aptitudes générales du chercheur par exemple. Les chercheurs ayant de meilleures aptitudes auraient plus de facilité à rédiger en anglais et ils seraient plus performants, non pas parce qu'ils publient en anglais mais parce qu'ils ont de meilleures aptitudes en recherche. Si c'était le cas, la véritable relation causale relierait aptitudes et langue d'une part, et aptitudes et performance d'autre part. La corrélation observée entre langue et performance serait fortuite.

Pour vérifier le caractère fallacieux de la relation langue-performance, il suffirait d'introduire une variable « aptitudes » dans nos modèles de régression. Malheureusement nous ne disposons d'aucune donnée semblable. Il nous est donc impossible de trancher le débat sur cette question. Mais aurions-nous montré qu'il n'y a pas de relation entre les aptitudes et la

performance ou la langue de publication que nous ferions toujours face au même problème : la relation langue-performance pourrait être fallacieuse si une autre troisième variable était liée aux deux premières. Et ainsi de suite jusqu'à ce que nous ayons épuisé la liste des troisièmes variables possibles.

Les difficultés d'établir un lien de causalité dans un contexte comme le nôtre ne s'arrêtent pas là. Il faudrait aussi que toutes les variables pertinentes, et seulement elles, soient incluses dans la régression, pour ne rien dire des possibles erreurs de mesure. L'oubli d'une variable pertinente ou l'ajout d'une variable non pertinente peut produire des coefficients biaisés. C'est le problème bien connu de la spécification du modèle. La manière de le résoudre consisterait à recourir à une théorie qui permette d'identifier toutes les variables à inclure, ce que nous n'avons pas fait faute de moyens.

Bref, nous devons reconnaître qu'il nous est impossible d'affirmer que publier en français nuit à la performance des chercheurs francophones. Mais nous pouvons affirmer qu'il y a une relation négative relativement robuste entre le pourcentage de publications en français et la performance d'un chercheur. Les chercheurs qui publient surtout en français publient moins et sont moins cités que les autres. Les chercheurs les plus performants publient plutôt en anglais.

La performance des chercheurs du CNRS

Nous avons souligné plus haut que les performances des chercheurs du CNRS n'étaient pas à la hauteur de ce que nous en attendions. En effet, Bland et ses collaborateurs ont montré que le dégagement du temps suffisant pour les activités de recherche, des signaux clairs de la part de la direction de l'unité sur le caractère prioritaire de la recherche et la culture de recherche au sein de l'unité sont trois facteurs qui sont liés positivement à la performance des chercheurs dans les facultés de médecine. Les mêmes facteurs sont susceptibles d'être à l'œuvre dans notre discipline. Pourquoi alors les chercheurs du CNRS ne sont-ils pas largement plus performants que leurs collègues professeurs ?

Les résultats du tableau 6 fournissent une réponse partielle. Les chercheurs du CNRS ne performant pas autant qu'ils le devraient parce qu'ils publient beaucoup en français. Lorsque nous tenons compte du pourcentage de publications en français et des autres variables de contrôle, nous retrouvons les chercheurs du CNRS dans le peloton de tête de tous les modèles sauf les deux premiers. En effet, en dehors du nombre de publications et du taux annuel de publications, les moyennes du CNRS ne se démarquent pas significativement de celles des unités les plus performantes d'Outre-Atlantique. Elles placent, au total, les chercheurs affiliés à l'organisation française au premier rang avec l'Université de Montréal. La somme des rangs pour chacun des neuf modèles se répartit comme suit : UdM (10), UL (19), CNRS (22),

ENAP (26), UO (31), UQAM (37)¹³. Or, ce dernier classement ne correspond pas à ce que prédirait le modèle proposé par Bland et ses collaborateurs que vous avons cité plus haut. Les performances des chercheurs du CNRS ne sont pas à la hauteur des conditions qui leur sont faites en théorie. Le modèle de Bland et de ses collaborateurs suggère un certain nombre de facteurs qui pourraient expliquer ce fait. Parmi ceux-ci, mentionnons la motivation, l'accès aux ressources (subventions de recherche, assistants), les récompenses pour les bons coups, l'accès au mentorat et la communication entre les membres de l'unité. Une étude approfondie des conditions d'exercice de la recherche au CNRS serait nécessaire pour expliquer la performance des chercheurs de cette organisation.

Publier en français pour se créer une niche

Dans un autre ordre d'idées, nous nous sommes demandé ce qui pouvait motiver le choix de la langue de publication dans le contexte où le fait de publier en français est associé à une performance plus faible. Postulant une forme de rationalité procédurale selon laquelle les chercheurs veulent optimiser leur position de pouvoir préceptoral (Imbeau et Couture, 2010), nous suggérons que la volonté de plusieurs chercheurs de publier surtout ou seulement en français, tout comme le discours à l'effet qu'il existe une science politique « française » qui se distinguerait d'une science politique « anglo-saxonne », pourraient être liés à la volonté de créer une niche francophone dans laquelle ces chercheurs occuperaient une position enviable. Ils seraient alors relativement plus performants par rapport à leurs collègues francophones qu'ils pourraient l'être s'ils se comparaient à l'ensemble des politologues du monde dont la langue principale de communication est l'anglais.

Ce phénomène rappelle le concept de « fermeture sociale » (*social closure*) développé par Max Weber (1922) ou ce que Charles Tilly (1998) appelle l'*opportunity hoarding*, un mécanisme par lequel les membres d'un réseau s'accaparent des ressources renouvelables et créent des croyances et des pratiques visant à maintenir le contrôle de ces ressources¹⁴. Les manifestations de ce mécanisme sont variées. Elles vont des pratiques protectionnistes de plusieurs professions (les comptables, les médecins, etc.) à l'exploitation du statut minoritaire par les élites amérindiennes, afro-américaines ou francophones du Canada (Murphy, 1988). Dans le cas qui nous occupe, la fermeture sociale permet à des politologues francophones de s'accaparer les ressources associées à la pratique de leur profession, notamment les postes

13. Ce résultat est rapporté dans le tableau 3. Le minimum théorique de cette somme est de 9 (9 x 1), le maximum de 54 (9 x 6).

14. Nous remercions James McGuire d'avoir porté à notre attention la parenté entre notre concept de niche et celui d'*opportunity hoarding*.

dans les universités et les centres de recherche, les subventions de recherche, la visibilité dans les médias, la notoriété, etc., en contribuant à la création d'une frontière autour des « parlant français » par un discours à l'effet qu'il existe une science politique typiquement française et par des pratiques de publication en français ainsi que de traduction et d'appropriation des travaux publiés en anglais.

Pour illustrer ce phénomène, comparons la distribution des chercheurs francophones dont la proportion des travaux publiés en français dépasse les 2/3 à la distribution de l'ensemble des chercheurs de notre base de données. Les deux distributions présentent une asymétrie positive avec un indice h maximal de 16 pour les francophones publiant surtout en français et de 38 pour l'ensemble des chercheurs de notre base de données¹⁵. Nous offrons une représentation visuelle de ces distributions à la figure 2. Nous y voyons bien l'effet de niche. En limitant l'univers de comparaison aux francophones publiant surtout en français, le chercheur ayant un indice h de 16 se retrouve à l'extrémité positive de la distribution : il est le plus performant. Mais lorsque nous le comparons à l'ensemble des chercheurs du CNRS et du Québec y compris ceux qui publient surtout en anglais, son indice h est inférieur de 12 points à l'indice h maximal (38). Il reste performant, mais il n'est plus le meilleur, loin s'en faut.

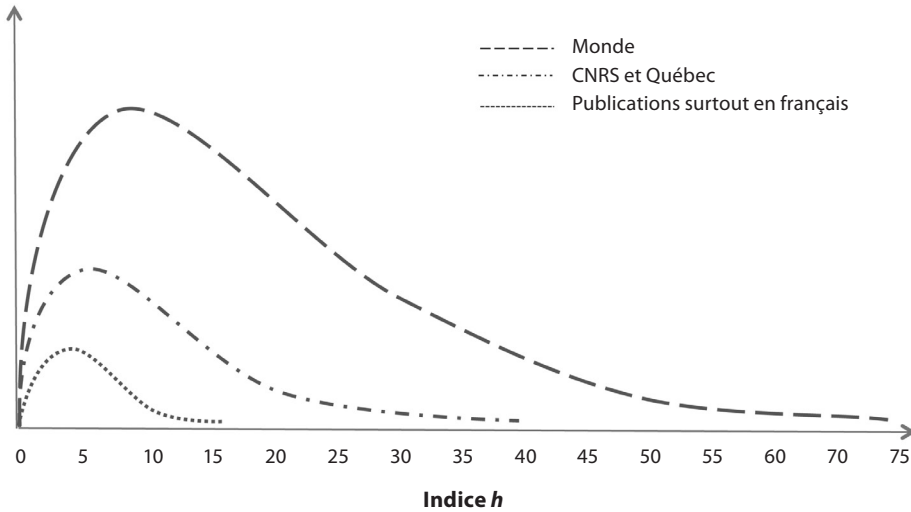
L'effet est encore plus saisissant si nous élargissons la comparaison à l'ensemble des politologues du monde. Pour nous faire une idée de ce à quoi ressemble cette distribution, nous avons tenté d'identifier les politologues qui sont susceptibles d'occuper le premier rang de nos mesures de performance. À cette fin nous avons demandé à quelques collègues de nous suggérer le nom de deux ou trois politologues qui devraient figurer parmi les dix plus grandes vedettes de la science politique d'aujourd'hui. Leurs réponses nous ont permis de générer une liste de 18 chercheurs-vedettes¹⁶. Les performances maximales de ces chercheurs sont de 456 publications répertoriées pour E. Ostrom, 62 892 citations pour R.D. Putnam et un indice h de 75 pour E. Ostrom. Nous pouvons donc considérer un indice h de 75 comme une approximation du maximum dans notre discipline. Par conséquent, pour représenter la distribution de l'indice h des politologues du monde, nous devons allonger la courbe de 37 points puisque l'extrémité passe de 38 à 75. L'effet de niche s'en trouve plus prononcé d'autant.

15. Pour simplifier la discussion, nous nous limitons à l'indice h , la mesure la plus synthétique de performance. *Mutatis mutandis*, les conclusions sont les mêmes pour les autres mesures.

16. Il s'agit, classés du plus grand (75) au plus petit (17) indice h , de : E. Ostrom, R. Keohane, S.P. Huntington, R.D. Putnam, A. Hirschman, C. Hood, W. Kymlicka, M. Olson, R.A. Dahl, W.H. Riker, A. Moravcsik, D. Easton, P.A. Hall, B.R. Barber, K.N. Waltz, J. Tully, M. Crozier, B. Bueno de Mesquita. Bien que mentionné, le nom de Charles Taylor a dû être rejeté de cette liste à cause de l'existence d'un trop grand nombre d'homonymes, ce qui nous a empêchés d'arriver à un compte relativement exact étant donné nos moyens.

FIGURE 2

Représentations théoriques de la distribution de l'indice h de trois ensembles de chercheurs en science politique



Pour conclure sur cet aspect des choses, qu'on nous permette une longue citation du sociologue français Georges Gurvitch qui résume bien notre conviction :

[I]l faut se garder de toute division en « sociologies nationales » et en « écoles sociologiques », car elle ne représenterait qu'un obstacle de plus pour la sociologie qui se propose de devenir une science dans le plein sens du terme [...] Il faut rappeler qu'il n'existe point [...] de division nationale dans les sciences proprement dites (pures ou appliquées) qui ont atteint un certain âge. Les conflits d'écoles apparaissent alors comme des maladies « infantiles » qui témoignent seulement de la difficulté des problèmes, des aspects et des perspectives qu'on n'a pas encore su dominer. (1957 : 5)

Faut-il alors cesser de publier en français ? Non, sans doute. Car il faut assurer la diffusion des travaux importants auprès du grand public. Nous voulons aussi initier les étudiants du premier cycle universitaire à la recherche en science politique. La barrière de la langue pourrait les décourager. Les étudiants des cycles supérieurs de science politique savent déjà qu'ils ne peuvent ignorer les travaux publiés en anglais ; sinon, ils l'apprendront vite à leurs dépens. Quant à nos résultats de recherche, ils devraient être diffusés auprès du plus grand nombre, c'est-à-dire en anglais... sauf, peut-être les textes qui, comme celui-ci, s'adressent seulement à un auditoire francophone.

Conclusion

La langue de publication est associée de façon systématique aux mesures de performance des chercheurs. C'est là notre principale conclusion. Les chercheurs qui publient surtout en français publient moins et sont moins cités que les autres. D'autres facteurs sont reliés à la performance : le genre, l'expérience et le milieu. Les femmes publient moins et sont moins citées que les hommes, les jeunes que leurs aînés. Les conclusions concernant les unités de rattachement sont plus nuancées. Les chercheurs de l'Université Laval et de l'Université de Montréal publient plus que les autres, mais leur performance en ce qui a trait au nombre de citations tend à être comparable à celle de leurs collègues du CNRS. Quant à la mesure synthétique qui combine le nombre de publications et de citations, l'indice h , elle permet d'identifier deux types de milieux : ceux qui semblent favoriser une plus grande performance (le CNRS, l'Université Laval et l'Université de Montréal) et les autres.

Malgré leur caractère convaincant, les résultats que nous rapportons ici ne sont pas suffisants pour donner un aperçu complet de la situation. Il faudrait, entre autres, considérer les différents champs de recherche et les diverses approches théoriques et méthodologiques, car il se peut que l'effet de la langue de publication varie d'un champ ou d'une approche à l'autre. Inclure ces variables dans nos régressions pourrait modifier nos conclusions. Il serait donc justifié de poursuivre la recherche dans cette direction.

Bibliographie

- Anderson, James D., 1974, « Foreign Language Barriers in Information Transfer », *Journal of Education for Librarianship*, vol. 14, n° 3, p. 171-184.
- Baneyx, Audrey, 2008, « 'Publish or Perish' as Citation Metrics Used To Analyze Scientific Output in the Humanities: International Case Studies in Economics, Geography, Social Sciences, Philosophy, and History », *Archivum Immunologiae et Therapiae Experimentalis*, vol. 56, p. 363-371.
- Bland, Carole J., Bruce A. Center, Deborah A. Finstad, Kelly R. Risbey et Justin G. Staples, 2005, « A Theoretical, Practical, Predictive Model of Faculty and Department Research Productivity », *Academic Medicine*, vol. 80, n° 3, p. 225-237.
- Bland, Carole J., Elizabeth Seaquist, James T. Pacala, Bruce A. Center et Deborah Finstad, 2002, « One School's Strategy to Assess and Improve the Vitality of its Faculty », *Academic Medicine*, vol. 77, p. 368-376.
- Gingras, Yves et Sébastien Mosbah-Natanson, 2010, « Les sciences sociales françaises entre ancrage local et visibilité internationale », *European Journal of Sociology / Archives européennes de sociologie*, vol. 51, n° 2, p. 305-321.
- Green, William H., 2005 [5^e éd.], *Économétrie*, Édition française dirigée par Didier Schlachter, Upper Saddle River (NJ), Prentice Hall, Pearson Education Inc.
- Gurvitch, George, 1957, *La vocation actuelle de la sociologie*, Paris, Presses universitaires de France.
- Hanson, Christopher W., 1962, *The Foreign Language Barrier in Science and Technology*, London, Aslib.

- Hesli, Vicki L. et Jay Mook. Lee, 2011, « Faculty Research Productivity: Why Do Some of our Colleagues Publish More than Others? », *PS: Political Science & Politics*, vol. 44, p. 393-408.
- Hirsch, J.E., 2005, « An Index to Quantify an Individual's Scientific Research Output », *Proceedings of the National Academy of Sciences of the United States of America*, vol. 102, p. 16569-16572.
- Imbeau, Louis M. et Jérôme Couture, 2010, « Pouvoir et politiques publiques », dans Stéphane Paquin, Luc Bernier et Guy Lachapelle (sous la dir. de), *L'analyse des politiques publiques*, Montréal, Presses de l'Université de Montréal, p. 37-72.
- Ioannidis, John P.A., 2005, « Why Most Published Research Findings Are False », *PLOS Medicine*, vol. 2, n° 8, p. 696-701.
- Ioannidis, John P.A., 2008, « Why Most Discovered True Associations Are Inflated », *Epidemiology*, vol. 19, n° 5, p. 640-648.
- Jacso, Peter, 2008, « Testing the Calculation of a Realistic *h*-index in Google Scholar, Scopus, and Web of Science for F.W. Lancaster », *Library Trends*, vol. 56, n° 4, p. 785-815.
- Kosmopoulos, Christine et Denise Pumain, 2007, « Citation, Citation, Citation: Bibliometrics, the Web and the Social Sciences and Humanities », *Cybergeo: European Journal of Geography*, n° 411, consulté sur Internet (<http://www.cybergeo.eu/index15463.html>) le 17 septembre 2011.
- Lotka, Alfred J., 1926, « The Frequency Distribution of Scientific Productivity », *Journal of the Washington Academy of Sciences*, vol. 16, n° 12, p. 317-324.
- Montpetit, Éric, André Blais et Martial Foucault, 2008, « What Does it Take for a Canadian Political Scientist to be Cited? », *Social Science Quarterly*, vol. 89, n° 3, p. 802-816.
- Murphy, Larry J., 1973, « Lotka's Law in the Humanities? », *Journal of the American Society for Information Science*, vol. 24, n° 6, p. 461-462.
- Murphy, Raymond, 1988, *Social Closure: The Theory of Monopolization and Exclusion*, Oxford, Clarendon Press.
- Ouimet, Mathieu, Pierre-Olivier Bédard et François Gélinau, 2011, « Are the *h*-index and Some of its Alternatives Discriminatory of Epistemological Beliefs and Methodological Preferences of Faculty Members? The Case of Social Scientists in Quebec », *Scientometrics*, vol. 88, n° 1, p. 91-106.
- Tilly, Charles, 1998, *Durable Inequality*, Berkeley, University of California Press.
- Weber, Max, 1922, *Economy and Society*, Los Angeles, University of California Press.
- Wood, D.N., 1967, « The Foreign-language Problem Facing Scientists and Technologists in the United Kingdom – Report of a Recent Survey », *Journal of Documentation*, vol. 23, n° 2, p. 117-130.
- Yitzhaki, Moshe, 1998, « The 'Language Preference' in Sociology: Measures of 'Language Self-Citation,' 'Relative Own-Language Preference Indicator,' and 'Mutual Use of Languages' », *Scientometrics*, vol. 41, n°s 1-2, p. 243-254.