



Les cafés de la statistique

"La statistique éclaire-t-elle les questions de société" ?

Soirée du 3 février 2015

Comment s'assurer de la validité des arguments statistiques contenus dans les études scientifiques ?

Synthèse des débats ^[*]

De plus en plus d'études scientifiques dont les conclusions reposent sur une approche statistique standard sont contestées, parfois avec raison. Cela pose la question de la validation de ces études, fondée pour partie sur leur reproductibilité et le débat critique entre chercheurs. Partant d'un même ensemble de données, peut-on reproduire quantitativement ou qualitativement, avec ou sans utiliser les mêmes méthodes, les résultats publiés ? L'accès aux données est-il garanti ? Selon quels critères juge-t-on qu'un résultat est plus ou moins fiable ?

Invité :

Stéphane Gregoir,

Insee, ancien directeur de la recherche de l'EDHEC (Ecole des hautes études commerciales) et du CREST (Centre de recherche en économie et statistique)

Exposé introductif :

Aujourd'hui, la question de la validation des études scientifiques est posée. Il y a un débat dans la presse, américaine et anglaise principalement, mais avec des échos partout, même en France. Comment construit-on un savoir scientifique, comment accumule-t-on progressivement des connaissances fiables ? Peut-on faire confiance pour cela aux règles et aux procédures suivies dans le monde scientifique ?

^[*] Tant l'exposé liminaire que le contenu des échanges sont structurés en quelques thèmes, sans suivre l'ordre chronologique. Par ailleurs, l'identité des intervenants n'était pas toujours connue et l'on a choisi de ne pas attribuer nominativement les propos. Au reste, ceux-ci ont été reconstitués à partir des notes du secrétariat sans reprendre leur formulation détaillée. Pour retracer le débat, les thèmes sont souvent introduits sous forme d'une question : ce qui vient ensuite n'est pas la seule réponse de l'invité, mais l'ensemble des contributions des participants.

La controverse est apparue dans le monde des sciences biomédicales, en y incluant les neurosciences et la psychologie. En 2012, un article signé de trois chercheurs nord-américains¹ est publié par l'Académie des sciences des Etats-Unis : il analyse les retraits d'articles de recherche publiés dans des revues de bio-médecine et sciences de la vie. Deux mille articles publiés depuis 1975 ont été retirés ultérieurement, et le nombre des retraits est en forte augmentation, surtout depuis le début du présent siècle. Pourquoi ces retraits ? Les erreurs constatées après publication ne comptent que pour un quart ; dans les trois quarts des cas, il s'agit de problèmes touchant au comportement des auteurs : plagiats, répliquions d'études déjà publiées par ailleurs, et surtout fraude. La fraude ici consiste essentiellement en la construction de données erronées, fabriquées ou manipulées pour fournir le résultat désiré. Le taux de retrait pour fraude a été multiplié par un facteur dix entre 1975 et 2010. Et les revues les mieux cotées sont aussi concernées: un nombre important de retraits pour fraude est constaté dans des revues dont « l'impact factor »² est élevé.

Pour comprendre comment on peut en arriver là, il faut d'abord revenir sur la nature de l'exercice qu'on fait quand on construit un savoir au moyen d'une expérience incluant un test statistique.

Le principe d'un test statistique

Une expérience est faite pour examiner s'il existe une relation entre deux phénomènes. Le résultat de l'expérience est, pour partie, aléatoire. A cause de cela, la décision que le chercheur doit prendre pour interpréter ce résultat peut être fautive. Les tests statistiques à utiliser sont les meilleures règles pour minimiser la fréquence des erreurs, a priori et indépendamment du résultat.

On peut comparer le chercheur à un juge, qui devrait se prononcer sur la relation éventuelle entre un accusé et un crime. L'hypothèse d'absence de relation signifie l'innocence de l'accusé (« hypothèse nulle »). Le résultat de l'expérience, ce sont les témoignages. Le juge est exposé à deux risques d'erreur : condamner un innocent – en l'occurrence, affirmer l'existence d'une relation alors qu'elle n'existe pas vraiment ; et acquitter un coupable – en l'occurrence, refuser l'existence d'une relation qui existe bel et bien. S'il commet le premier type d'erreur (« erreur de première espèce »), il crée un « faux positif » : un innocent se retrouvera en prison. S'il commet le deuxième type d'erreur (« erreur de deuxième espèce »), il crée un « faux négatif » : un coupable circule en liberté.

Procédure de test d'hypothèse

Quelles sont les erreurs possibles ?

| Issues possibles d'un test statistique | | |
|--|--|--|
| Décision | Situation réelle inconnue | |
| | H_0 : il n'y a pas de relation | H_0 est fautive : il y a une relation |
| Maintenir H_0 | Maintenir à raison 😊 | Maintenir à tort 😞 |
| Rejeter H_0 | Rejeter à tort Conclusion erronée d'une relation 😞 | Rejeter à raison Présomption de l'existence d'une relation 😊 |

¹ "Misconduct accounts for the majority of retracted scientific publications" Ferric C. Fang, R. Grant Steen, and Arturo Casadevall - Proceedings of the National Academy of Science of the United States of America Volume 109, n° 42 – 2012. Voir aussi : « La fraude scientifique est plus répandue qu'on le croit » Pierre Bathélémy – Le Monde 3 octobre 2012

² « L'impact factor » - en français « facteur d'impact » - est un indicateur de la visibilité d'une revue scientifique, calculé en dénombrant les citations d'articles de cette revue par des articles d'autres revues.

Procédure de test d'hypothèse

Comment choisir ?

| Issues possibles d'un test statistique | | |
|--|------------------|--------------------|
| Décision | Situation réelle | |
| | H_0 est vraie | H_0 est fausse |
| Maintenir H_0 | 😊 | Maintenir à tort 😞 |
| Rejeter H_0 | Refuser à tort 😞 | 😊 |

Objectif du statisticien:
Minimiser les probabilités de commettre les erreurs de première et de seconde espèce

Principe: Minimiser l'erreur de seconde espèce pour un niveau d'erreur de première espèce choisi a priori

Il n'existe évidemment pas de règle qui permette de minimiser les deux risques d'erreur à la fois : il faut arbitrer entre les deux. La pratique courante en statistique depuis des décennies³ consiste à limiter le risque de « faux positif » à un niveau fixé a priori (en sciences économiques et sociales, on retient habituellement le niveau 5 %) et à choisir la règle qui minimise alors le risque de « faux négatif ». Le complément à l'unité de ce dernier risque est de ce fait maximisé : c'est la probabilité de détecter à juste titre une relation qui existe, que l'on appelle la puissance du test statistique. Pour disposer d'une puissance satisfaisante, il faut beaucoup d'observations, surtout si la relation (qui existe vraiment) est d'intensité faible (pas très éloignée de l'absence de relation).

Comment la sociologie de la recherche pousse à la faute

La publication de résultats fragiles est souvent la conséquence de la petite taille des échantillons et du choix des hypothèses testées. Pour être publié, l'auteur est tenté de rechercher des hypothèses nouvelles, voire bizarres, de façon à produire « du nouveau ». Un grand nombre de ces hypothèses devraient être rejetées. Par ailleurs, le coût des expériences augmente avec le nombre des observations. Si le budget pour une expérience est limité, la puissance du test statistique le sera aussi. Des chercheurs ont examiné la puissance moyenne des tests conduits en neurosciences : elle n'était que de 21 % (autrement dit, dans presque huit cas sur dix un effet réel risquait de ne pas être détecté). En psychologie de 35 %. Face à des risques aussi élevés de non-détection d'effets réels, une réaction compréhensible de la part des chercheurs est de diversifier les essais, quitte à aller chercher des jeux d'hypothèses bizarres, au risque de multiplier les faux positifs.

En effet, une autre caractéristique de la recherche actuelle, c'est que les résultats négatifs sont peu porteurs. Vers le milieu des années 1990, 35 % des articles présentaient des résultats négatifs ; actuellement on n'en compte plus que 13 %. Publier un résultat négatif n'apporte guère de gloire ! Au contraire, un résultat positif, même à la limite de la significativité⁴, peut être souvent cité. Et on constate en effet parmi les résultats publiés accompagnés de leurs risques de première espèce une accumulation des résultats dont ce risque se situe juste en dessous du seuil de 5%.

³ C'est le « règle de Neyman- Pearson », proposée vers 1935 par les statisticiens Jerzy Neyman (1894-1981) et Egon Pearson (1895-1980).

⁴ C'est-à-dire un résultat qui a à peine moins que 5 % de chances de se produire si la relation n'existe pas, si 5 % est la valeur maximale retenue pour le risque de première espèce.

Dans un contexte où la carrière d'un chercheur et son accès aux sources de financement dépendent du nombre de ses publications et de leur audience, ces caractéristiques des travaux publiés poussent les chercheurs au bord de la « mauvaise conduite ». Sélectionner des hypothèses exotiques, réaliser des tests peu puissants, accepter un biais en faveur des résultats positifs : cela ouvre la voie aux « mauvaises conduites de recherche » qui conduisent au constat effectué ces dernières années.

Le garde-fou des « referees »

Pourtant, les articles ne sont pas publiés sans examen préalable dans les revues scientifiques. Celles-ci font appel à des « referees » pour évaluer les « papiers » qui leur sont soumis : un, deux ou trois selon les cas, censés être de bons connaisseurs du domaine étudié. Comment peuvent-ils laisser passer des erreurs ou des fraudes ?

Tout d'abord, l'information fournie aux referees n'est le plus souvent pas complète. Ils n'ont en général pas accès aux protocoles d'expérience ; et jusqu'à une période récente, peu de revues demandaient à avoir accès aux données (et par conséquent les referees n'en disposaient pas). Dans beaucoup de disciplines, la mise en place des expériences est un investissement qu'il faut protéger afin de ne pas communiquer des informations précieuses à des concurrents éventuels. Le même argument vaut pour les programmes informatiques, qui peuvent même dans certains cas être brevetés : le referee n'en dispose pas en général.

Par ailleurs, le referee ne travaille pas pour lui-même, mais pour la collectivité scientifique. C'est un travail astreignant, peu visible et donc peu valorisé⁵. Du coup, les rapports demandés se font attendre : 6 mois, un an, un an et demi... Certains directeurs de laboratoire « sous-traitent » les rapports qui leur ont été demandés à de jeunes docteurs, voire à des doctorants qui n'ont qu'une vue très partielle du domaine considéré.

Tout cela fait que le système des referees est fragile.

Quand l'idéologie s'en mêle

L'erreur est naturelle et dans une certaine mesure inévitable dans la recherche scientifique. Mais parfois l'idéologie s'en mêle : on veut prouver « à tout prix » un résultat qui conforte une position générale que l'on a sur l'économie ou sur la société. L'exemple de l'article de Kenneth Rogoff et Carmen Reinhart, paru dans l'*American Economic Review* peut être donné : cet article prétendait démontrer qu'au-delà d'un certain seuil – 90 % du PIB – l'endettement public devenait un obstacle à la croissance. Il a été utilisé tant par le Fonds monétaire international que par la Commission européenne pour justifier des politiques d'austérité budgétaire. C'est un doctorant en sciences économiques de 28 ans dans une université du Massachusetts, à qui ses professeurs avaient demandé de refaire l'étude, qui a découvert des erreurs de calcul et des exclusions arbitraires de certaines données : si l'on corrigeait ces erreurs, les conclusions ne pouvaient pas être reproduites. L'article initial avait été écrit pour soutenir une thèse a priori.

Autre exemple : il y a une dizaine d'années, un livre très populaire, « *Freakonomics* » de Steven Levitt et Stephen Dubner, vendu à plusieurs millions d'exemplaires dans le monde, a donné des exemples d'application des techniques microéconomiques à des problèmes de société. Parmi celles-ci, reprenant un article de Donohue et Levitt antérieurement paru dans le très sérieux *Quarterly*

⁵ Même si certaines revues ont pris l'habitude de rémunérer leurs referees

Journal of Economics, figurait une étude de l'impact de la légalisation de l'avortement sur la criminalité une vingtaine d'années plus tard : moins de naissances non désirées conduisait à moins de crimes, selon les auteurs. L'analyse a été reprise par des chercheurs de la Banque fédérale de réserve de Boston, qui ont trouvé de multiples erreurs de données et de méthode : au bout de quelques années de controverse, le sens de l'effet éventuel de la légalisation de l'avortement sur le taux de criminalité reste incertain. Le succès initial de l'étude et du livre n'était pas sans rapport avec une prise de position idéologique forte sur un sujet controversé.

Quelques pistes

La multiplication des mises en cause d'études scientifiques a provoqué une prise de conscience et un appel à plus d'exigences dans l'établissement et la publication des résultats.

La formation des scientifiques aux méthodes statistiques devrait être renforcée : trop souvent, certains appliquent mécaniquement des recettes en utilisant directement les sorties des logiciels sans en maîtriser le sens.

La communication des données et des programmes tend à devenir de règle, dans le cadre d'une exigence de transparence sur l'ensemble des procédures et des traitements.

Les tailles d'échantillon demandées augmentent, de façon que les tests soient plus puissants, sachant qu'à budget constant cela implique moins d'expériences et donc moins de résultats.

Simultanément, il convient d'être plus exigeant dans la procédure de test, en fixant des risques de première espèce plus petits que le « $\alpha = 5\%$ » rituel. Sans aller jusqu'aux niveaux requis par les physiciens du nucléaire⁶, certains proposent de fixer ce risque plutôt à 0,5 %.

L'approche bayésienne des tests rencontre un succès croissant, en particulier parce que les nouvelles possibilités de calcul permettent de la mettre en œuvre ; il s'agit d'une approche qui traite symétriquement les deux hypothèses entre lesquelles il faut choisir.

Enfin, la sévérité à l'égard des fraudeurs se renforce. Aux Etats-Unis, s'agissant du domaine biomédical, un « Office of Research Integrity⁷ » a été créé pour détecter la fraude et poursuivre les fraudeurs : treize ont été condamnés en 2014.

Débat :

Sur la gravité de la situation

La situation est-elle si grave ? L'importance des fraudes commentées par l'invité peut le laisser penser : plusieurs milliers d'articles dans la littérature biomédicale depuis 1973. Mais ce chiffre doit être rapproché de la production totale d'articles, estimée à 25 millions depuis 1940 ! La fraude reste un phénomène marginal. Cependant, les effets d'une seule fraude peuvent être dévastateurs pour toute une production scientifique, lorsque le doute s'installe.

⁶ Dans l'expérience visant à détecter le boson de Higgs, le risque de première espèce a été limité à 1 pour cent mille. Mais à ce niveau de risque, on ne pourrait pratiquement plus rien valider en sciences humaines.

⁷ <http://ori.hhs.gov/>

Un participant fait part de son étonnement au sujet des informations qu'il a lues récemment à propos du « bisphénol A » : en France, l'Agence nationale de sécurité sanitaire de l'alimentation, de l'environnement et du travail (Anses) a estimé au vu d'un certain nombre d'études que ce produit présente un danger, alors que l'Agence européenne de sécurité des aliments estime le contraire, tout en divisant par dix la dose maximale jugée acceptable dans le corps humain. Que penser ?

Le cas des sciences physiques et des sciences de la vie doit sans doute être distingué du cas de l'économie et de la sociologie. Dans les sciences « dures », les articles proposés aux revues ont beaucoup de chances d'être acceptés : pour la *Physical Review*, le taux d'acceptation avoisine 80 %. Dans les sciences humaines, les meilleures revues pratiquent couramment des taux compris entre 10 % et 30 %. On n'y cite pas de cas de fraude, mais le problème majeur est l'interférence entre l'idéologie et la production des résultats. A ce propos, certains se demandent s'il peut exister dans ces disciplines des jeux d'hypothèses ne comportant aucune idéologie ? Peut-être pas, répond l'invité : mais la validation empirique, le test en tant que tel ne doit pas être idéologique. Et ce n'est pas toujours ainsi que les choses fonctionnent : tel économiste renommé déclare ouvertement n'avoir pas publié un résultat parce que celui-ci ne cadrerait pas avec le point de vue qu'il voulait mettre en avant. Dans beaucoup d'études, la sophistication de la méthodologie masque un manque de robustesse du résultat : il suffit d'enlever un point aux observations pour que le résultat ne tienne plus du tout. Par ailleurs, beaucoup de chercheurs en sciences humaines ou sociales manient sans connaissance suffisante la statistique, qui n'est pas une science intuitive : ainsi, on trouve dans des articles des modèles auto-régressifs⁸ estimés comme s'il s'agissait de modèles ordinaires, sans que soient corrigés les biais associés à ce type de modèle.

Un participant a rappelé que les études publiées en économie doivent contenir une revue de la « littérature » antérieurement publiée sur le sujet, et ainsi se situer dans un historique d'accumulation de connaissances, à l'intérieur d'un cadre théorique commun. N'y a-t-il pas là une garantie de rigueur de la démarche ? Mais il a été noté par ailleurs que le travail de recherche théorique, portant sur les modèles eux-mêmes, est difficile et long, et que beaucoup de jeunes chercheurs préfèrent commencer par analyser des données, quitte à retenir ensuite la théorie qui colle le mieux avec celles-ci. Les « apprentissages automatiques » pratiqués sur les Big Data (mégadonnées) reposent sur ces détections et exploitations systématisées : plutôt que de construire des savoirs, on y recueille des descriptions de régularités statistiques.

Sur les causes de la dérive

On est peu revenu pendant le débat sur les causes des problèmes qui sont apparus, causes relevant, selon l'invité, de la sociologie de la recherche. Tous les participants ont toutefois été d'accord pour attribuer un rôle décisif – et négatif – aux incitations adressées aux chercheurs en matière de publication : évaluation de leur travail d'après le nombre de leurs publications, au cours d'une période récente et courte et en prenant en compte de façon mécanique le nombre de citations obtenues par ces publications. Ces incitations découragent du travail « lent, sérieux et détaillé » et poussent à rechercher le résultat étonnant qui provoque un flot de reprises, voire à « bidonner » pour assurer sa carrière. On aboutit à des absurdités : l'invité cite un curriculum vitae de chercheur qui repose entièrement sur la somme des « impact factors » des revues ayant publié ses articles – somme actualisée au fil du temps et de l'évolution de ces « IF » (alors que les articles du chercheur

⁸ Modèles dans lesquels intervient la relation entre une variable observée à un instant t et la même variable observée à un instant antérieur.

sont restés les mêmes !). Un participant a cité le cas d'un chercheur français en mathématiques⁹ qui avait très peu publié pendant les dix années précédant la publication de l'article qui lui valut la médaille Fields : cette rareté de publication faillit compromettre son emploi.

Les chercheurs travaillent désormais dans un monde concurrentiel, dans lequel certains peuvent obtenir des salaires élevés. Mais c'est au prix d'un alourdissement des comptes qu'ils ont à rendre par rapport à des grilles et à des règles bureaucratiques. Peut-être faut-il regretter le temps où les chercheurs étaient plus libres, même s'ils étaient mal payés...

La réplication : un remède ?

De plus en plus souvent, la question de la réplication des études scientifiques est mise au cœur du débat sur leur validation. C'est d'ailleurs l'impossibilité de répliquer certaines études fondamentales concernant le cancer qui a vivement inquiété le public en 2012 : selon un article publié dans Nature, des chercheurs d'une firme pharmaceutique américaine ayant tenté de reproduire 53 études considérées comme fondamentales n'ont confirmé les résultats originaux que pour six d'entre elles.

Qui dit réplication dit tout d'abord accès aux données de l'étude. Une participante confirme le fait : pour la première fois, elle s'est vu demander par la revue à laquelle elle avait soumis un article de biomédecine les données de base de son travail. En ce cas précis, l'hôpital public dans lequel elle travaille a donné son accord pour cette transmission. Mais ce n'est pas toujours possible. Un participant qui professionnellement transmet à des chercheurs des fichiers de données provenant de producteurs publics le rappelle : on communique ces données aux chercheurs à la condition expresse qu'ils ne les rediffusent à personne. Comment pourraient-ils les communiquer aux éditeurs d'une revue, voire de plusieurs revues, et aux referees ? A cela l'invité répond que l'anonymat des referees n'est pas une règle pratiquée dans toutes les disciplines : pour certaines revues de mathématiques, les referees signent leurs rapports. On pourrait envisager d'étendre l'accès aux données à des referees nommément désignés et ayant pris les engagements nécessaires.

Les difficultés pour répliquer ne s'arrêtent pas à l'accès aux données. La question des programmes informatiques peut être épineuse. Par exemple, de nombreuses études récentes ont recours à des simulations répétées d'expériences aléatoires. En ce cas, pour procéder à la réplication, il faut disposer du même générateur de nombres aléatoires, utiliser les mêmes graines du simulateur, etc. ; et ensuite tirer les échantillons des distributions a posteriori selon les mêmes procédures... ; c'est rarement possible.

Pour un participant, c'est la possibilité même d'une réplication qui doit être mise en doute, sur le plan philosophique : il n'y a pas deux expériences pareilles. Sans entrer dans ce débat, l'invité signale que certains philosophes ont introduit le concept moins exigeant de « véri-similitude » pour désigner la plus ou moins bonne correspondance entre les propositions et les faits : même si la capacité à reproduire n'est pas dans la nature, la véri-similitude peut être testée, et cela constitue un chaînon essentiel de la démarche scientifique.

Sur un plan voisin, l'essor des « méta-analyses » d'études a été cité par un participant comme une voie pour arriver à de meilleures certitudes, notamment dans le domaine biomédical.

⁹ Laurent Lafforgue

Vers une déontologie de la recherche ?

Les cas de fraude sont rares, mais ils sont très souvent cités et ont un caractère dévastateur non seulement pour le chercheur en question – ce qui peut être considéré comme normal – mais pour ses collègues de laboratoire, son université, et l'image de la recherche en général.

Existe-t-il une déontologie de la recherche, qui préviendrait ce genre de comportement ? demande un participant. Et cette déontologie est-elle enseignée aux jeunes chercheurs ? Les institutions chargées de la formation des chercheurs, et notamment les écoles doctorales dans les universités, sont conscientes de ces questions et organisent des cours et des séminaires de déontologie. Mais ce sont des enseignements dispensés « à froid » : beaucoup de gens sont prêts à les oublier quand ils sont en situation... A une jeune participante qui se demande comment concilier la justesse des messages avec l'impératif de simplification des résultats de la recherche qui prévaut lorsqu'on veut populariser ceux-ci, l'invité répond que cela fait partie du métier de statisticien : c'en est la dimension « rhétorique », de simplifier pour faire comprendre le message, tout en rendant compte fidèlement de la réalité du résultat et de son degré de validité. C'est l'expression qu'il s'agit de simplifier, pas le résultat ! Si l'on estime devoir être percutant pour être audible, on bascule dans l'idéologie.

Les revues scientifiques elles-mêmes s'efforcent de dissuader les éventuels fraudeurs : un des moyens consiste à afficher une politique pour les retraits d'articles, précisant ainsi les critères qui peuvent leur permettre de conclure en ce sens. Ces critères sont-ils partagés entre ces revues, ou propres à chacune ?

Il reste que le travail de vérification est long et pénible, comme le souligne un participant, surtout que dans la grande majorité des cas on ne trouve rien de répréhensible (heureusement). Comment résoudre ce problème ? Peut-être en le confiant à des étudiants, à titre de démarche éducative certainement très profitable. De façon générale, tout ce qui favorise l'esprit critique chez les jeunes chercheurs est bon à prendre, pour leur permettre de distinguer ultérieurement ce qui est crédible de ce qui ne l'est pas.

Dans la statistique publique européenne, une « démarche qualité » a été mise en place : elle repose sur le « Code des bonnes pratiques » et sur les revues par les pairs. Attention toutefois au caractère essentiellement formel du respect de certains critères : ainsi, pour respecter le critère de cohérence des résultats, un institut de statistique n'a pas hésité à modifier toutes les micro-données d'une enquête de façon à retrouver un agrégat déterminé par ailleurs. Les règles ne sont pas grand-chose sans une prise en compte des intentions qui ont conduit à les adopter.

Un participant évoque les cas de plus en plus nombreux où une expertise statistique est recherchée dans un procès, par exemple pour établir un lien entre l'absorption d'un médicament et un trouble ultérieur. Pour ce participant, on peut de moins en moins séparer la recherche de ses interfaces avec la vie réelle, qu'il s'agisse de relations avec les organes d'opinion (question de la vulgarisation : « un intervalle de confiance, ça ne peut pas passer dans un titre de journal »), avec le monde juridique ou avec le monde de l'entreprise.

Il se peut enfin que le système du « tout recherche » pratiqué par certaines grandes universités américaines ait atteint ses limites. Ces universités, disposant d'un très gros capital et donc de

revenus indépendants de leurs services d'enseignement, attirent des « grands noms » de la recherche pour des raisons d'image : ce qui n'est pas sans lien avec la valorisation des chercheurs en fonction des citations qu'ils obtiennent. Mais ce modèle économique sera peut-être affecté par les progrès techniques et les gains de productivité que connaît le secteur de l'enseignement supérieur : développements des MOOCs¹⁰, mise à disposition d'outils informatiques avec un tutorat moins qualifié... La qualité de l'enseignement effectivement dispensé devient plus importante que le prestige des chercheurs très en vue. Peut-être cela annonce-t-il une organisation différente du système de recherche lui-même.



¹⁰ Massive open online courses