

“Quels étudiants pour quelles universités? Analyses empiriques de mécanismes d'allocation centralisée.”

Thierry Magnac

Quels étudiants pour quelles universités? Analyses empiriques de mécanismes d'allocation centralisée.

Thierry Magnac*

March 16, 2018

Résumé: On décrit, dans cette revue, les articles récents qui se sont attachés à construire des méthodes empiriques structurelles pour estimer les préférences des agents, étudiants et universités, ou élèves et écoles, qui participent à des mécanismes d'allocation centralisée. Ces recherches empiriques s'attachent, d'abord, à confirmer ou infirmer les hypothèses et prédictions des modèles théoriques développés dans les vingt dernières années. Elles permettent aussi des analyses concrètes des mécanismes existants, par l'estimation de l'impact contrefactuel de changements dans les mécanismes utilisés et la comparaison entre divers mécanismes. Elles conduisent enfin à des arguments et conditions d'application qui justifient l'utilisation de tels ou tels mécanismes.

Mot clés: Education, appariement, mécanisme d'allocation, économétrie appliquée

Keywords: Education, matching, mechanism design, applied econometrics

JEL codes: C18, C57, D47, D82, I21

*Toulouse School of Economics (TSE), Université de Toulouse Capitole, Manufacture des Tabacs, 21 allée de Brienne, 31015 Toulouse, France; thierry.magnac@tse-fr.eu.

"[...M]arketplaces work well when they provide thickness to the market, help it deal with the congestion that thickness can bring, and make it safe for participants to act effectively on their preferences. Centralized clearinghouses organized around the deferred acceptance algorithm can have these properties, and this has sometimes allowed failed markets to be reorganized."

Alvin Roth, "Deferred acceptance algorithms: history, theory, practice, and open questions", *Int J Game Theory* [2008].

1 Introduction¹

En juillet 2017, au moment d'écrire la première version de cet article, les médias bruissaient des problèmes rencontrés par les étudiants après que les résultats de la procédure Admission Post Bac sont sortis (par exemple, l'éditorial du Monde du 19 juillet 2017²) et plus particulièrement le nombre important de candidats qui n'avaient pas reçu d'offres, une dizaine de pourcents, comme en 2016 à vrai dire. Ces mécanismes d'allocation centralisée d'étudiants à des écoles ou à des universités sont utilisés de plus en plus souvent et sont devenus l'objet de l'attention particulière des économistes dans de nombreuses régions du monde et en particulier en France: la procédure d'allocation des candidats à des postes de maîtres de conférence (Haeringer et Iehlé [2010]), l'allocation des élèves aux collèges et lycées (Hiller et Tercieux [2014]) et l'admission Post Bac elle-même dans des travaux en cours par Haeringer et Iehlé. C'est bien sûr aussi le cas dans de nombreux autres pays comme aux Etats-Unis où l'analyse du programme de Medical Residency Programme a valu à Alvin Roth sa récompense prestigieuse de 2012 dont le co-récipiendaire est Lloyd Shapley, l'un des pionniers avec David Gale de l'analyse conceptuelle de ces mécanismes (voir Forges, Haeringer et Iehlé [2013] pour une revue de ces travaux). Le mécanisme utilisé en Turquie pour allouer les étudiants après le lycée à leur spécialité

¹**Remerciements:** Cet article a bénéficié des nombreuses discussions que j'ai pu avoir sur ces sujets avec Yinghua He et des conseils de Julien Grenet, Gabrielle Fack et d'un rapporteur anonyme. Cette recherche a reçu le soutien financier du Conseil Européen de la Recherche (European Community's Seventh Framework Program FP7/2007-2013 grant agreement N°295298). Je reste le seul responsable de toutes omissions et erreurs.

²http://www.lemonde.fr/idees/article/2017/07/19/admission-postbac-mettre-fin-a-l-hypocrisie_5162427_3232.html

universitaire a été souvent l'objet d'analyse des procédures d'acceptation différée à la Gale & Shapley ("Deferred acceptance") en commençant par Balinski et Sonmez [1999] jusqu'à Akyol et Krishna [2017]. La taille gigantesque du système analogue utilisé en Chine est devenu le centre d'intérêt pour la recherche académique (par exemple, Chen et Kesten [2017]).

Toutes ces analyses utilisent des outils de conception de mécanismes ("mechanism design") développés dans la littérature économique. Cette boîte à outils définit en partie le métier des économistes au titre d'ingénieurs de conception des marchés.

Les marchés centralisés qui allouent les étudiants aux universités sont des marchés sans prix qui arbitrent offre et demande ou à tout le moins où les "prix", comme des allocations de bourse dont le montant est fixé à l'avance, ont un rôle mineur. Ensuite, la construction d'un mécanisme centralisé d'allocation requiert de connaître les préférences qu'ont les universités pour les étudiants et les préférences des étudiants pour les universités et de chercher la meilleure adéquation entre les deux parties. Mais cette allocation crée des conflits de répartition entre étudiants, entre universités et entre universités et étudiants et les appariements auxquelles ces mécanismes conduisent, doivent être comparés entre eux au titre de leurs propriétés souhaitables. La Pareto-optimalité, le caractère non manipulable et la stabilité sont ainsi les trois propriétés les plus mises en avant par la littérature théorique qui s'est développée depuis plus de trente ans et continue à l'être. Mais certains résultats théoriques ne sont pas dénués d'ambiguïtés qui ne peuvent être levées sans des apports expérimentaux ou empiriques.

Ce sont ces derniers aspects empiriques qui sont le centre d'intérêt de cette revue de littérature. J'y étudie la manière dont les analyses empiriques récentes ont utilisé des données de mécanismes existants pour estimer les primitives de l'allocation que sont les préférences des étudiants et des universités, et leur degré de rationalité. Les préférences des universités pour les étudiants, en d'autres termes leur sélection, peuvent être déterminées par leurs notes à un examen centralisé comme en Turquie et en Chine, par leurs notes sur dossier comme l'accès aux classes préparatoires en France, par un système de priorités, par exemple celles données aux étudiants habitant l'académie pour l'entrée à l'Université en France ou enfin par leur tirage au sort qui est utilisé dans certaines filières universitaires en France dans les années récentes. Les soumissions des préférences des étudiants pour les spécialités et la localisation universitaire peuvent être libres comme à Taiwan où les étudiants peuvent proposer de longues listes de

choix (Fack, Grenet et He [2015]) ou contraintes comme dans le processus APB où la première étape s'effectuait à partir de listes de vingt-quatre choix (ou 12 par formation) soumises par les étudiants.

Un des aspects les plus intéressants de ces analyses empiriques vient de ce que les agents n'ont pas forcément intérêt à révéler leurs vraies préférences, autrement dit quand le mécanisme est manipulable. Les auteurs se sont attachés à formuler des méthodes qui permettent de recouvrer, même dans ce cas, les paramètres gouvernant les vraies préférences. Ces estimations des préférences en retour ont été utilisées pour estimer l'impact qu'auraient des changements de mécanisme sur l'allocation des étudiants mais aussi sur le bien-être des agents. Cette démarche empirique complète donc les analyses théoriques de façon concrète.

Cette présentation des analyses empiriques portant sur l'organisation de l'allocation des étudiants dans les spécialités universitaires restera peu technique et s'appuiera sur des exemples du mieux qu'elle pourra. Je reviens d'abord sur le cadre d'analyse théorique et l'analyse des politiques en place dans la Section 2 puis sur les caractéristiques principales des modèles empiriques dans la Section 3 qui est le coeur de cette revue. J'étudie les problèmes spécifiques de calcul des contrefactuels dans la Section 4 et donne un catalogue des résultats importants de ces analyses empiriques dans la Section 5. La Section 6 discute et conclue cette présentation.

2 Préférences, stratégies et mécanismes d'allocation

Dans cette section, on expose de façon brève le modèle de base et les questions que se posent les économistes qui analysent les mécanismes d'allocation d'étudiants à des universités.³

2.1 Appariements et leurs propriétés

On utilise les notations usuelles pour modéliser le choix d'écoles par des étudiants (Roth et Sotomayor [1990]). Chacune des écoles, j , en nombre J , offrent q_j places au plus ("capacité") aux étudiants. Ceux-ci peuvent aussi préférer une option extérieure, notée par l'indice 0, et qui a une capacité infinie. Ce peut être une école qui ne participe pas à l'appariement centralisé

³J'utiliserai de façon interchangeable les couples étudiants/élèves ou écoles/universités même s'il y a des différences importantes pour la construction des mécanismes (notes et priorités, par exemple).

ou toute autre activité non scolaire. Les étudiants, appelés i , sont en nombre I et demandent une place au plus dans l'une des écoles ou choisissent l'option extérieure.

Dans les modèles considérés ici, les préférences des écoles sont exprimées en fonction des caractéristiques spécifiques de chaque étudiant, indépendamment des autres. De même, les étudiants ne prennent pas en compte l'allocation d'autres étudiants, par exemple leurs fratrie ou leurs amis (effets de "pairs"). Leurs préférences sont exprimées de façon ordinaire ou cardinale suivant le type de résultats que les auteurs voudront mettre en avant.

Chaque agent connaît ces propres préférences mais nous le verrons plus loin, la connaissance des préférences des autres agents diffère suivant les cas de figure. Néanmoins, apprendre ses propres préférences peut entraîner des coûts, de sélection, comme la lecture de dossiers ou l'organisation d'examens de la part des écoles et des coûts de recherche de la part des étudiants. Il peut y avoir aussi des différences entre des établissements qui sont maîtres du mode de choix de leurs préférences (par exemple, les classes préparatoires aux grandes écoles en France) et ceux qui ne le sont pas. Par exemple, il est courant que les écoles primaires n'aient que des préférences succinctes, imposées d'ailleurs souvent par une autorité de tutelle, comme un système de priorités dérivé d'une carte scolaire (Fack et Grenet [2010]) ou d'avantages attachés au rassemblement de la fratrie. Les écoles sont alors souvent indifférentes entre certains élèves et il est nécessaire de les départager, par exemple, par un système de tirage au sort, comme dans certaines filières universitaires en France dans le système APB, ou dans les écoles de Boston (Abdulkadiroglu et al. [2006]) ou de Pékin (He [2015]).

On peut maintenant décrire un appariement possible qui alloue les étudiants à des écoles. C'est une application, μ , de $\mathcal{I} = \{1, \dots, I\}$ dans $\mathcal{J} = \{0, 1, \dots, J\}$ qui associe à chaque étudiant une école $j = \mu(i)$ (ou l'option externe, $\mu(i) = 0$). Notons que sa correspondance inverse $\mu^{-1}(j)$ associe à toute école, l'ensemble de ses étudiants appariés, qui sont en nombre inférieur ou égal à sa capacité, q_j . De plus, l'algorithme qui permet l'implémentation de cet appariement à partir des soumissions de listes ordonnées de préférences par chacun des acteurs (écoles et étudiants) est appelée mécanisme, comme le plus célèbre d'entre eux, qui est celui d'acceptation différée de Gale et Shapley [1962].

Cet algorithme fonctionne de la manière suivante dans le cas où les étudiants sont les agents qui proposent leurs choix en premier. À l'étape initiale, chaque école j considère tous les

étudiants qui l'ont placée en première dans leurs listes, classent ces étudiants dans l'ordre de ses préférences et rejette ceux qu'elle préfère le moins, si sa capacité est atteinte, en ne gardant que les q_j qu'elle préfère le plus. Mais cette "acceptation" n'est que différée. En effet, au deuxième tour, les étudiants rejetés par leur premier choix, candidatent à leur deuxième choix sur leur liste d'écoles. Certains de ces étudiants rejoignent les étudiants retenus par l'école j à la première étape. L'école rejette alors les candidats surnuméraires par rapport à sa capacité et ceci dans l'ordre de ses préférences comme à la première étape. Ces étapes se répètent jusqu'à ce que tous les étudiants soient assignés ou aient choisi l'option extérieure. L'algorithme converge en un nombre fini d'étapes et l'allocation obtenue est l'appariement final, $\mu(\cdot)$.

Il existe un nombre fini, néanmoins très grand, d'allocations et on les étudie en fonction de trois propriétés formelles: l'efficacité au sens de Pareto, la stabilité et le caractère non manipulable. La première notion et la dernière sont usuelles, la dernière se basant sur le fait que la révélation des vraies préférences n'est pas dominée par une autre stratégie. La deuxième est moins connue et s'appuie sur l'absence de jalousie légitime ("justified envy"). Une allocation est stable s'il n'existe aucune paire non-appariée, école et étudiant, (j, i) , $j \neq \mu(i)$, qui préférerait (strictement) être appariée à leurs (à un de leurs, pour les écoles) appariements courants.

Néanmoins, en général, aucun mécanisme ne conduit à des allocations qui ont les trois propriétés (Balinski et Sonmez [1999]).⁴ Il faut se contenter de deux propriétés seulement: les allocations obtenues par acceptation différée (Gale et Shapley) sont stables et non manipulables (d'un côté seulement, écoles ou étudiants suivant celui qui propose en premier); les allocations obtenues par "cycle d'échange optimal" (Top Trading Cycles ou TTC) sont Pareto optimales et non manipulables (Abdulkadiroglu et Sonmez [2003]). De plus, le nombre d'allocations stables est multiple, comme le nombre d'*optima* de Pareto.

Il existe aussi des mécanismes qui n'ont aucune de ces propriétés. Par exemple, dans la famille des mécanismes à acceptation différée, certains mécanismes sont contraints, par exemple ceux qui requièrent que la liste ordonnée de choix d'écoles soumise par les étudiants soit

⁴Il y a des cas particuliers. Par exemple, si toutes les écoles ont les mêmes préférences, l'allocation obtenue par le mécanisme de dictature séquentielle a les trois propriétés. Certaines de ces propriétés dépendent aussi de la considération que les écoles ne sont pas un agent mais seulement un objet à consommer par les étudiants. Dans ce cas, les préférences des écoles n'interviennent pas dans les calculs d'optima de Pareto.

tronquée. La plupart de ces mécanismes ont des listes d'une dizaine de choix sauf cas exceptionnel comme l'entrée à l'université à Taiwan citée plus haut. D'autres mécanismes imposent des coûts d'applications pour toute école supplémentaire sur la liste comme à l'entrée à l'université en Hongrie.⁵

Ces deux types de mécanismes ne sont pas stables et sont manipulables, par exemple, en omettant des options dont les probabilités d'obtention sont extrêmement faibles. C'est aussi le cas pour le mécanisme dit de Boston, pourtant assez courant et qu'on pourrait appeler d'acceptation immédiate. Son algorithme, en effet, est le même que celui d'acceptation différée évoqué plus haut, à ceci près qu'on remplace l'acceptation différée par une acceptation immédiate. Les capacités des écoles sont ainsi saturées beaucoup plus vite et il est très peu probable qu'un élève obtienne une école populaire s'il ne l'a pas mise en premier choix.

2.2 Structures de marché

Roth dans une série de travaux, par exemple Roth et Xing [1994] et McKinney et al. [2005], a tenté de décrire les différents stades d'organisation de marchés "sans transfert" ou "sans prix", où se rencontrent une offre et une demande et conduisent à un équilibre, comme une allocation d'étudiants à des universités. Ces marchés, tant qu'ils sont petits, resteraient décentralisés. Chaque étudiant peut aller voir toutes les écoles qui l'intéresse et chaque école, évaluer ces étudiants. Cela crée pourtant rapidement de la congestion si le nombre d'intervenants sur le marché grandit puisque chaque agent doit consacrer du temps ou des ressources à rechercher de l'information ou à communiquer avec beaucoup d'agents de l'autre côté du marché. De plus, comme l'allocation finale est incertaine, les agents d'un côté du marché consacrent probablement trop de temps et de ressources à évaluer des agents de l'autre côté du marché qui n'ont pas d'intérêt marqué pour eux. C'est pourtant comme cela qu'est organisée l'allocation des étudiants aux collèges aux Etats-Unis et une petite littérature s'est développée sur ce sujet (Chade, Lewis and Smith [2014] et Che et Koh [2016]).

Si le mécanisme est décentralisé, il existe toujours la tentation d'"anticiper" le marché, c'est à dire de faire affaire avec un étudiant ou une université avant tous les autres et ceci de plus en plus tôt. C'est pourquoi, d'après Roth et Xing [1994], certaines institutions se créent pour réguler

⁵Voir le site www.matching-in-practice.eu qui recense différents systèmes d'allocation.

le marché en instaurant un calendrier de recrutement. Mais cela ne suffit souvent pas pour discipliner le marché car les solutions sont instables. Certaines parties ont un intérêt prononcé et peu de coûts, à ne pas respecter le calendrier. C'est alors qu'une centralisation de l'allocation des étudiants aux universités peut se mettre en place, comme le Medical Residency Program aux Etats-Unis qui existe depuis plus de 60 ans (Roth [1982]). Par exemple, nous reviendrons sur les résultats d'Abdulkadiroglu, Agarwal et Pathak [2017] qui montre que la centralisation de l'allocation des élèves aux lycées de New York en 2003 a donné des résultats meilleurs que la coordination décentralisée qui existait auparavant. Machado et Szerman [2015] montrent aussi que la plupart des institutions éducatives au Brésil qui ont adopté la centralisation proposée par le gouvernement brésilien depuis 2008, recrutent des étudiants que ces institutions préfèrent. Les étudiants parcourent une distance plus grande, entre lieu de résidence et université d'accueil, dans ce système d'allocation centralisée.

Néanmoins, certains de ces marchés centralisés peuvent ne pas être stables. Roth et Xing [1994] identifie un certain nombre de facteurs d'instabilité. L'offre peut être beaucoup plus grande que la demande comme sur le marché des gastroentérologues aux Etats-Unis dont la centralisation a échoué. La mauvaise connaissance des règles du mécanisme, comme pour la procédure d'allocation des doctorants au Brésil (Bardella et Sotomayor [2014]) ou de docteurs aux postes de maitres de conférences (Haeringer et Iehlé [2012]) peuvent mener à des choix incompatibles avec les vraies règles. Il peut y avoir aussi des tentatives de manipulations de la part de certaines institutions.

Il existe enfin des frictions de marchés comme celles induites par des coûts de recherche et d'application de la part des étudiants et des coûts de sélection de la part des universités. Or les coûts de recherche et d'application ont dramatiquement baissé dans les années récentes à cause d'Internet et des plateformes de candidature en ligne. Cela accroît fortement les coûts de sélection pour les universités puisqu'elles se retrouvent face à de très nombreuses candidatures, dont la plupart sont de circonstance. Il semble alors optimal pour les écoles d'imposer des coûts de candidature, par exemple des visites dans les universités comme aux Etats-Unis ou de longs questionnaires à remplir en ligne pour des offres d'emploi. Arnosti, Johari et Kanoria [2016] montrent de façon surprenante que des coûts d'application plus grands peuvent permettre d'augmenter le bien être des candidats.

3 Modèles empiriques et estimation

Les données disponibles peuvent être de type expérimental, en laboratoire ou sur le terrain, ou recueillies directement à partir de mécanismes existants. Elles décrivent au moins indirectement, les préférences, c'est-à-dire une liste ordonnée de choix d'universités soumise par les étudiants. On dispose aussi de données sur la façon dont les universités classent les étudiants.

Je laisserai de côté ici une large part de la littérature en économie de l'éducation, comme Arcidiacono [2005], Epple, Romano et Sieg [2006], Hastings, Kane et Staiger [2009], Burgess, Greaves, Vignoles et Wilson [2015] ou Gazmuri [2016] qui n'a pas pour but de comprendre et d'évaluer les mécanismes d'allocation mais plutôt les déterminants de la demande et de l'offre, comme les capacités des filières universitaires, les aides financières ou les frais d'inscription. Je laisserai aussi de côté la littérature utilisant les discontinuités générées par les mécanismes (Abdulkadiroglu, Angrist, Narita & Pathak [2017b]). Cette revue de la littérature porte sur la construction d'analyses contrefactuelles de l'impact de la modification du mécanisme d'allocation sur les résultats de l'allocation et le bien-être des étudiants et non de l'impact d'autres changements dans l'environnement.

Les acteurs de l'allocation sont les écoles $j = 1, \dots, J$ et les étudiants $i = 1, \dots, I$. Néanmoins, les préférences affichées des écoles sont en général supposées non-stratégiques et révéler leurs "vraies" préférences.⁶ Elles sont maintenues fixes, en particulier dans les expérimentations contrefactuelles que les auteurs construisent après estimation. Ceci est une autre différence entre cette littérature sur l'impact des mécanismes d'allocation et les articles sur le choix de filières universitaires comme dans Fu [2014].

Les préférences des universités se traduisent par un ensemble de listes ordonnées d'étudiants pour chaque université ou école. Je reviendrai plus tard sur la connaissance complète ou incomplète de ce classement par l'économètre et par les agents. Les écoles classent les élèves par des priorités, conférées par les fratries ou les zones de carte scolaire et complétées éventuellement par un tirage aléatoire qui départagent les ex-aequo (Calsamiglia et al. [2017], Agarwal et

⁶C'est une hypothèse justifiée pour des écoles dont les priorités sont données par une autorité de tutelle. Cela l'est beaucoup moins pour des établissements qui auraient intérêt à manipuler les critères d'entrée. L'"objectivité" de notes d'examen peu manipulables, peut être une forme d'engagement à ne pas jouer stratégique.

Somainsi [2017]). Les universités classent aussi les candidatures d'étudiants par leurs notes (Ajayi [2015], Carvalho et al. [2017]) ou une version bruitée de leurs notes (He et Magnac [2017]).

Ce n'est donc pas du côté des écoles et des universités que les classements changent quand le mécanisme change mais du côté des élèves ou étudiants. Ceux-ci ont leurs préférences que l'on observe au mieux de façon ordinaire et partielle et des croyances sur leurs probabilités de réussite à telle ou telle école et ceci en fonction du mécanisme utilisé. C'est l'identification et l'estimation de ces deux types d'objets, préférences et croyances, qui font l'objet de cette section. Mais il est en général difficile d'identifier séparément les préférences et les croyances, et l'estimation des préférences se fait en fixant les croyances ou la manière dont elles se calculent ("anticipations rationnelles"). Néanmoins, il existe au moins deux exemples d'identification directe des préférences et des croyances. Kapor et al. [2017] dispose en effet de données d'enquêtes sur les croyances subjectives des agents. Lufade [2018] quant à elle identifie directement les préférences pour une sous-population.

He [2015] propose une distinction très utile entre modélisations selon deux hypothèses qui sont satisfaites ou non: d'abord, l'hypothèse d'anticipations rationnelles ou de "sophistication" des agents, c'est-à-dire, quand leurs croyances subjectives et objectives coïncident et sont homogènes ("common priors"); ensuite, l'hypothèse d'observabilité des probabilités de succès ou d'estimabilité de ces probabilités par l'économètre. L'abandon de l'hypothèse d'anticipations rationnelles introduit de nouveaux paramètres à estimer comme ceux gouvernant la probabilité que les étudiants soient sophistiqués. Quand on abandonne ces deux hypothèses, sophistication et probabilités de succès estimables, He [2015] et Hwang [2016] montrent que les préférences restent partiellement identifiées.

D'autre part, on observe au mieux les préférences ordinaires des étudiants puisque les écoles constituent un ensemble de choix discrets pour ceux-ci. Or, pour presque tout mécanisme contrefactuel, on a besoin d'identifier les préférences cardinales pour calculer les nouvelles stratégies mises en oeuvre par les agents sauf si on fait l'hypothèse que tous les étudiants sont toujours non-stratégiques (ou "naïfs").

On commence cette section par l'étude d'un cas de figure particulier d'acceptation différée (les étudiants proposant), un mécanisme pour lequel une stratégie faiblement dominante pour les étudiants est de déclarer leurs vraies préférences ordinaires (Fack, Grenet et He [2015]). Ceci

permet une estimation directe des préférences cardinales en se dispensant de toute hypothèse informationnelle. Si cela n'est pas le cas, il faut préciser le cadre informationnel et le concept de solution du jeu d'allocation joué par les étudiants car certains étudiants peuvent ne pas révéler leurs vraies préférences et jouer de façon stratégique. On revoit ensuite l'estimation des croyances qu'elles soient rationnelles ou non.

3.1 Estimation directe de la demande

La révélation des vraies préférences s'obtient, sous certaines conditions, quand le mécanisme en place est un mécanisme d'acceptation différée sans restrictions de longueur de listes et sans coûts d'application. Même dans ce cas pourtant, les étudiants pourraient ne pas révéler leurs vraies préférences car ce n'est qu'une stratégie faiblement dominante que de révéler la vérité. En effet, certaines alternatives ont une probabilité nulle d'occurrence ou d'autres alternatives sont dominées par des options extérieures et leur présence sur la liste de choix n'a pas d'importance (Fack et al. [2015]).

Observer, de manière complète, les listes ordonnées de choix des écoles, incluant une possible option extérieure, est d'ailleurs rare sauf dans un cadre expérimental (He et Magnac [2017]). Si c'est le cas, l'estimation des préférences cardinales se dérive de modèles à choix discrets utilisés en économie industrielle (Berry et Haile [2014] par exemple). Notons néanmoins que l'information sur les préférences est plus complète qu'en économie industrielle puisqu'elle consiste, pour chaque étudiant, en des listes d'alternatives ordonnées au lieu d'un seul choix.

Ces listes ordonnées peuvent ou non inclure une option extérieure. Certains auteurs supposent que les élèves ont un ensemble fermé de choix (par exemple une liste d'écoles comme dans Abdukadiroglu et al. [2017a], ou Agarwal et Somaini [2017]) et négligent la possibilité d'une option extérieure. Cette hypothèse n'est pourtant pas toujours crédible puisqu'il existe très souvent des options extérieures, comme des écoles privées. De plus, l'existence d'une option extérieure fait que certaines alternatives sont inacceptables ce qui rationalise le fait que les listes soumises par les élèves peuvent ne pas avoir la longueur maximale (He et Magnac [2017]) ou que les élèves changent d'avis ex-post (Calsamiglia et al. [2017]). On peut aussi faire ou non l'hypothèse que les alternatives inacceptables ne sont pas classées (He [2015]) et la seule information dont on dispose est qu'elles sont inacceptables. Nous verrons que l'existence d'une

option extérieure complique l'identification des modèles dans la Section 3.5.

En reprenant un cadre de choix discrets à utilité aléatoire, les auteurs postulent, pour un étudiant $i = 1, \dots, I$, des utilités propres à chaque alternative $j = 0, \dots, J$, qui s'écrivent en fonction des caractéristiques des écoles z_j et des étudiants x_i comme par exemple:

$$v_{ij} = z_j(\gamma_i + x_i\beta) + \zeta_j + \varepsilon_{ij}. \quad (1)$$

Cette spécification des vraies préférences, même quand elles ne sont pas observées, est celle retenue par tous les auteurs que nous étudions avec quelques variantes. Le degré d'hétérogénéité observable et inobservable entre étudiants (γ_i et x_i) varie et le choc idiosyncratique, usuellement Logit mais quelquefois normal (Calsamiglia et al. [2017]), peut ne pas être présent (Agarwal [2015]). Toute liste ordonnées de préférences de l'étudiant i , $j_1^{(i)}, j_2^{(i)}, \dots, j_{J+1}^{(i)}$ aura donc une probabilité dérivée d'un modèle appelé "Logit explosé" ("exploded Logit") que l'on peut intégrer par rapport à une distribution des γ_i , ou en utilisant une approximation par méthodes de simulation. La prise en compte de l'option extérieure, $j = 0$, quand elle existe, se fait aisément même si les alternatives inacceptables ne sont pas classées. Les estimateurs sont soit de type Bayésiens (Abdulkadiroglu et al. [2017a], Agarwal et Somaini [2017]) soit de maximum de vraisemblance simulé ou non (He et Magnac [2017]).

Il existe aussi des mécanismes qui produisent des listes ordonnées qui peuvent ne pas être complètes si le mécanisme utilisé est un mécanisme d'acceptation différée tronquée ou comportant des coûts liés à la longueur de la liste. Dans le cas de mécanisme à acceptation différée, on sait alors que l'ordre de la liste reste conforme à l'ordre des préférences (Haeringer et Klijn [2009]) mais cela ne suffit pas. Les listes incomplètes qui sont soumises sont non seulement incomplètes mais leur degré d'incomplétude est aussi stratégique (Fack et al. [2015]). Certains articles pourtant négligent ce problème, qui est un problème de sélection, et font l'hypothèse incorrecte que la probabilité d'observer ces listes est égale à la probabilité d'observer les vraies préférences (Abdulkadiroglu et al. [2017a]).

Quand le mécanisme conduit les étudiants ou certains étudiants à des choix stratégiques, il faut modéliser les anticipations des probabilités de réussite que les étudiants forment à partir de leur connaissance du jeu. Ces probabilités de succès sont conditionnées aux ensembles d'information dont les agents disposent et c'est par cela que nous commençons.

3.2 Structure informationnelle du jeu

Pour les étudiants, l'incertitude a deux sources. D'abord, celle relevant de leur position dans les listes de préférences des écoles ou filières universitaires. Ensuite les préférences et croyances des autres étudiants créent de l'incertitude et son effet dépend de chaque mécanisme.

Pour fixer les préférences des écoles, on supposera dans la suite que chaque école ou université, j , établit un score spécifique à chaque étudiant, s_{ij} , qui est donné par:

$$s_{ij} = w_{ij} + \eta_{ij}, \quad (2)$$

où un score plus élevé $s_{ij} > s_{i'j}$ signifie que l'étudiant i est préféré par l'école j à l'étudiant i' . Les deux termes sont toujours connus par les écoles alors que, dans la plupart des cas, η_{ij} n'est connu ni par les étudiants ni par l'économètre. Le terme w_{ij} , est connu par l'économètre et peut l'être par les agents. En règle générale, les priorités pour les écoles primaires ou secondaires, c'est-à-dire w_{ij} , sont supposées connues par les étudiants (Calsamiglia et al. [2017], Fack et al. [2015]) quand les élèves font leurs choix mais tout tirage aléatoire, c'est-à-dire η_{ij} , servant à départager les ex-aequo ne le sont pas.⁷ Les applications empiriques sur le choix de filières universitaires utilisent des modèles similaires quand les examens ont lieu après les choix des filières (Ajayi [2015], Carvalho et al. [2017]) et w_{ij} est alors une fonction d'un certain nombre de caractéristiques de l'étudiant et de la filière à laquelle il candidate. L'incertitude n'est pas toujours présente puisque les notes, s_{ij} , peuvent être connues au moment du choix des étudiants comme en Turquie (Akyol et Krishna [2017]) ou en Tunisie (Lufade [2018]).

Dans le cas où les notes sont partiellement connues, un modèle de détermination de ces notes comme l'équation (2) est spécifié et ce modèle, ainsi que la valeur de ses paramètres, est supposé être de connaissance commune pour les étudiants. Rappelons que les écoles sont supposées être un acteur passif du jeu et leurs listes de préférences sont supposées ne pas varier même si les notes qui en sont un déterminant peuvent être aléatoires pour les étudiants, avant qu'ils fassent leurs choix.

Les étudiants sont supposés avoir des croyances, rationnelles ou non sur la détermination des préférences des écoles ou des universités.

⁷Pour modéliser exactement ce cas, il faudrait ajouter des contraintes sur la variation des η_{ij} pour respecter les priorités strictes en fonction des w_{ij} .

Anticipations rationnelles L'information que les étudiants ont les uns sur les autres est la plus incomplète. Certains articles empiriques posent l'hypothèse que les étudiants sont rationnels, qu'ils connaissent la structure du jeu et que cette information est connaissance commune – on dit aussi qu'ils sont sophistiqués. D'autres auteurs supposent que certains étudiants sont sophistiqués et d'autres non. D'autres enfin observent directement les croyances subjectives (Kapor et al. [2017]).

Les étudiants sont supposés connaître la distribution des covariables dans la population des étudiants et des universités.⁸ De plus, les articles objets de cette revue font des hypothèses différentes quant à la connaissance par les étudiants, des préférences ou actions des autres étudiants. Des concepts de solution de type équilibre de Nash ou Nash-Bayésien et diverses variantes de l'ensemble d'information peuvent être trouvés dans la littérature:

- Equilibre de Nash: Les étudiants connaissent les choix des autres étudiants et réagissent à ces choix. Ce cas semble particulièrement adapté au cas où le jeu est bien connu et a été joué à de nombreuses reprises dans le passé (Carvalho et al. [2017]).
- Equilibre Nash-Bayésien: Les étudiants ne connaissent que la distribution des préférences des autres étudiants et réagissent aux anticipations des actions de ceux-ci (Fack et al. [2015]).

Il est important de noter que le jeu joué par les étudiants est un jeu de congestion (*pseudopotential games* voir Dubey et al. [2006]) et l'équilibre, pour la plupart des mécanismes utilisés, peut se résumer par des seuils de notes, ou de rangs, propres à chaque école (Azevedo et Lechno [2017]). Par exemple, si un mécanisme d'acceptation différée est utilisé, le seuil de l'école j détermine les étudiants qui peuvent avoir, à l'équilibre, accès à cette école, la condition étant que la note ou rang de cet étudiant est plus élevée que ce seuil. Ainsi la seule chose qui importe pour chaque étudiant est l'ensemble des écoles dans lesquelles il est accepté et tout se passe comme s'il peut choisir son école préférée parmi les écoles où il est accepté.

En ce sens, les stratégies des étudiants s'écrivent en fonction de statistiques agrégées relatives aux autres étudiants et cela fait que la différence entre des équilibres de Nash ou Nash-Bayésien

⁸Comme le jeu a beaucoup d'acteurs, il importe peu en général que cette distribution soit la distribution empirique dans l'échantillon ou la distribution dans la population.

n'est pas toujours très claire. Comme le jeu met en oeuvre chaque étudiant vis-à-vis des actions agrégées des autres acteurs, l'équilibre de Nash peut être réinterprété en disant que la distribution empirique des choix conditionnellement aux covariables est très proche de la distribution des choix dans la population. Ceci concerne les équilibres dont le calcul sera étudié dans la Section 4.

Autres croyances Les anticipations peuvent être adaptives comme dans Agarwal et Somaini [2017] ou Luffade [2018]. Ce sont les probabilités de réussite des années précédentes qui sont prises comme croyances par les étudiants. Il se peut aussi que les étudiants n'utilisent que des règles grossières comme des anticipations adaptives sans considération des caractéristiques individuelles (Agarwal et Somaini) ou utilisent des règles d'élimination de stratégies dominées (He [2015]) ou des règles heuristiques comme chez Hwang [2016]. Ces dernières règles ne permettent que d'identifier partiellement les préférences. C'est la même chose chez Fack et al [2015] dans leur modèle le moins contraint qui utilise la seule restriction que les listes déclarées sont ordonnées en conformité avec les vraies préférences.

3.3 Probabilités d'acceptation

Si les agents sont sophistiqués, les croyances doivent être calculées (Calsamiglia et al. [2017]) ou estimées directement à partir des données observées, comme dans Carvalho et al. [2017] ou Agarwal et Somaini [2017].

On considère ici le mécanisme d'allocation décrit par Carvalho et al. [2017]. Sa généralisation au cas de mécanismes où l'équilibre Nash-bayésien est déterminé par des seuils d'acceptation est dérivée par Agarwal et Somaini [2017]. Cette classe contient les mécanismes usuels sauf les mécanismes de type cycle d'échange optimal (TTC).

Le mécanisme de Carvalho et al. [2017], en le simplifiant, suppose que n étudiants doivent choisir de concourir à une seule de deux écoles médicales, que l'on note S et F qui ont un nombre de places q_S et q_F . On supposera que n est beaucoup plus grand que les places disponibles, $q_S + q_F$. Après avoir choisi de concourir à S et F , ce choix étant dénoté D_i , les étudiants passent l'examen qui conduit à leur classement par les écoles. On observe leurs notes, s_i , qui ne dépendent pas de l'école considérée.

Dans ce cadre simple, l'équilibre (de Nash) s'écrit, sous certaines conditions,⁹

$$\begin{aligned} \sum_{i=1}^n \mathbf{1}\{s_{ij} > t_S, D_i = S\} &= n_S, \\ \sum_{i=1}^n \mathbf{1}\{s_{ij} > t_F, D_i = F\} &= n_F, \end{aligned} \tag{3}$$

dans lesquels $\mathbf{1}\{\cdot\}$ est l'indicateur de la condition, la variable t_S (respectivement t_F) est le seuil d'acceptation des étudiants ayant choisi l'école S (resp. F). En d'autres termes, ces seuils sont les quantiles d'ordre $1 - n_j/n$ de la distribution de s_i dans la population d'étudiants qui ont choisi $j \in \{S, F\}$.¹⁰

La probabilité d'acceptation dépend alors de deux types d'événements aléatoires. D'abord, on peut conditionner par rapport à tout seuil possible (t_S, t_F) et la probabilité de réussite dans l'une ou l'autre école est donnée par:

$$P_i(t_j) = \Pr(s_i > t_j). \tag{4}$$

L'aléatoire est propre à l'individu, s_i , et vient soit de tirages aléatoires pour départager les ex-aequo dans le cas des priorités des écoles ou de notes aléatoires pour l'entrée à l'université. Si s_i est connu par l'étudiant comme dans le cas des universités en Turquie, il n'y a pas d'incertitude individuelle.

Ensuite, l'aléatoire vient des seuils (t_S, t_F) dont la distribution est donnée par les équations (3). Dans celles-ci, il peut y avoir deux sources d'incertitudes: celles liées aux chocs sur les scores $\{s_{i'}\}_{i' \neq i}$ des autres étudiants, celles liées aux préférences cardinales des autres étudiants à travers leurs choix, $D_{i'}$. Quand on intègre cette incertitude, on obtient les probabilités de réussite, anticipées par des agents rationnels:

$$P_i^{(r)} = E_{(t_S, t_F)} P_i(t_j) \tag{5}$$

C'est à ce dernier stade que les stratégies empiriques des papiers diffèrent:

⁹Ces conditions sont liées à la saturation des contraintes de capacité, et elles sont satisfaites ici car le nombre d'étudiants est beaucoup plus grand que les capacités. Au prix de quelques manipulations analytiques, il n'est pas difficile d'écrire le cas où certaines écoles ne saturent pas leurs capacités (Agarwal et Somaini [2017]).

¹⁰Les quantiles ne sont pas définis de manière unique dans l'échantillon et on utilise une des conventions habituelles pour les rendre uniques et être en conformité avec les écritures ci-dessous.

- Carvalho et al [2017]: on considère un équilibre de Nash et les valeurs des D_i sont fixes quand on rééchantillonne dans les équations (3) pour obtenir la distribution de (t_S, t_F) . L'aléatoire qui sous-tend cette distribution vient donc seulement de la distribution de $\{s_\nu\}$.
- Fack, et al. [2015], Calsamiglia et al. [2017]: on suppose que le jeu est joué par un continuum d'agents et que $n \rightarrow \infty$ en fixant le nombre d'écoles et en adaptant les capacités des écoles de telle façon à maintenir les n_j/n constants (à la variation induite par le problème d'entiers naturels près). Fack et al. [2015] montre que pour des mécanismes d'acceptation différée les seuils (t_S, t_F) convergent à la vitesse \sqrt{n} vers leur limite en probabilité, (t_S^∞, t_F^∞) , qui est déterministe (voir aussi Azevedo et Leshno [2017]). Les probabilités d'acceptation de l'équation (5) s'expriment comme dans l'équation (4) où on a remplacé t_j par t_j^∞ . Il est à noter que, dans toute estimation ultérieure, la valeur inconnue $P_i(t_j^\infty)$ est alors remplacée par un estimateur convergent, c'est à dire par $P_i(t_j)$ si t_j est la valeur observée dans l'échantillon. Les auteurs conditionnent alors l'analyse par rapport aux seuils observés. Pour l'inférence statistique, il faut prendre en compte qu'ils sont aléatoires.
- Agarwal et Somaini [2017]: on rééchantillonne dans la distribution des listes soumises par les étudiants et dans la distribution des priorités aléatoires ou des notes obtenues par les étudiants. Puis on recalcule les équilibres obtenus par les équilibres d'offre et demande décrits par l'équation (3) pour chacun de ces échantillons. Ceci conduit à la distribution des seuils que l'on peut alors utiliser pour estimer (5).

Ces différentes procédures se généralisent à des cadres divers correspondant à la variété des mécanismes utilisés dans la réalité. Ces probabilités d'acceptation sont utilisées comme estimateurs de première étape pour obtenir des estimateurs des paramètres de préférences quand les choix sont stratégiques.

3.4 Estimation indirecte des préférences

Si le chercheur suspecte que les préférences révélées ne sont pas les vraies préférences des agents car le mécanisme induit probablement des comportements stratégiques chez certains étudiants,

il faut au préalable, pour obtenir une estimation des vraies préférences qui serviront à calculer les contrefactuels, estimer les probabilités de succès dans chacune des alternatives ou raisonner à partir de restrictions identifiantes spécifiques.

Je commence ici par décrire ce que les auteurs développent sous des hypothèses de croyances grossières puis je passe à l'estimation sous hypothèse de stabilité c'est-à-dire robuste à tout type de croyances tant que celles-ci conduisent à un équilibre stable. Nous revenons à la fin à un cadre où nous connaissons ou pouvons estimer les croyances des étudiants, et où nous pouvons estimer les préférences à partir de ces croyances.

Croyances grossières He [2015] et Hwang [2016] ont montré qu'il était possible d'estimer les préférences des étudiants à partir de données dérivées du mécanisme de Boston et d'hypothèses faibles. He [2015] par exemple utilise la règle que des stratégies dominées ne sont pas jouées. Il prend aussi en compte le fait que certaines observations de listes ordonnées de choix peuvent contenir des options inacceptables ex-post mais dont le coût d'inclusion dans la liste est zéro. Il impose des règles spécifiques aux stratégies mixtes sous-tendant ces listes. Il en déduit des bornes sur les probabilités de choix de telle ou telle liste ordonnée de choix ce qui mène à l'identification partielle des paramètres de préférences.

Néanmoins, ce qui anime He [2015] sont des tests de sophistication des agents. En particulier il regarde deux types d'hypothèses emboîtées. Le premier ensemble oppose une hypothèse nulle d'équilibre Nash-bayésien – c'est à dire d'anticipations rationnelles communes – à une hypothèse alternative de croyances hétérogènes avec probabilités de réussite dans chaque école non dégénérées (c'est-à-dire non nulles). Cette dernière hypothèse sert aussi d'hypothèse nulle face à l'hypothèse alternative de croyances hétérogènes partiellement dégénérées. Le deuxième ensemble emboîté d'hypothèses part, non de l'hypothèse nulle Nash-bayésienne, mais d'une hypothèse nulle naïve où les étudiants révèlent leurs vraies préférences, le reste de la séquence d'hypothèses restant le même. Ce qui l'intéresse donc, tout particulièrement, est la rationalité des agents.

Hwang [2016] de son côté utilise un cadre où les données viennent d'un mécanisme de Boston et où l'identification des préférences restent partielles. Il considère deux règles heuristiques. Les étudiants anticipent correctement la popularité (c'est-à-dire la congestion) des écoles. Deux-

ièrement, les étudiants ne classent une école avant une autre de probabilité d'acceptation plus grande que si la première est strictement préférée.

Ajayi [2015] utilise aussi ce type de restrictions dérivées de Chade et Smith [2006] qui font que les listes sont ordonnées de telle façon que les préférences décroissent et que les niveaux de sélectivité aussi. Ces derniers sont mesurés par des seuils de notes d'entrée pour chacune des écoles. Ajayi [2015] n'utilise pourtant qu'une implication de cette règle pour l'estimation des préférences. Elle contraste le choix de la première école d'une liste vis à vis d'écoles placées dans la liste et de sélectivité approximativement équivalente.

Propriété de stabilité Fack et al. [2015] et Agarwal [2015] utilisent la propriété de stabilité de l'équilibre pour estimer indirectement la demande et se dispensent donc du calcul des probabilités d'accès. C'est une hypothèse qui est courante dans la littérature sur l'appariement monogame comme le mariage (par exemple Galichon et Salanié [2015]) – car on n'observe que les mariages effectifs et non les préférences des agents – ou sur l'appariement polyvalent comme des relations d'échange entre entreprises, les unes vendant ou achetant aux autres des biens intermédiaires (Fox [2010]).

Agarwal [2015] utilise cette notion car la seule information dont il dispose est l'appariement final d'un élève médecin avec une discipline médicale dans une université (Medical Residence Program, un mécanisme étudié et façonné par Alvin Roth [2008]). Agarwal [2015] dispose aussi des caractéristiques des deux parties et utilise comme restriction pour l'estimation, la notion de stabilité par paire ("pairwise stability") et l'absence de paire bloquante. Aucune université, appariée à une étudiante, ne pourrait préférer une autre étudiante, qui la préférerait aussi, à l'université à laquelle cette étudiante est appariée. Agarwal [2015] spécifie de manière particulière les préférences des agents. Les universités ont des préférences verticales pour les étudiants incluant les caractéristiques des étudiants résidents et un terme d'hétérogénéité inobservable constant. Les préférences des étudiants, quant à elles, sont restreintes au modèle de caractéristiques pures, c'est-à-dire, qu'elles ne dépendent pas d'un terme aléatoire spécifique à l'appariement entre université et étudiant comme dans Berry et Haile [2014].

Fack et al. [2015] utilise un argument différent. Les données que les auteurs utilisent viennent d'un mécanisme d'acceptation différée tronqué pour lequel la révélation des préférences

n'est pas nécessairement une stratégie dominante même faiblement. Ils montrent néanmoins que l'équilibre Nash-bayésien, en grand échantillon, est stable. Les auteurs contrastent des estimations qu'ils obtiennent sous hypothèse de révélation des vraies préférences et sous hypothèse de stabilité de l'équilibre. De plus, ils disposent d'informations supplémentaires par rapport à Agarwal puisqu'ils observent les listes, certes partielles, de préférences des étudiants et non pas seulement les caractéristiques de leurs appariements. Fack et al [2015] utilisent alors ces restrictions supplémentaires dans un cadre d'inégalités de moments. Ils testent d'ailleurs les restrictions dérivées de la stabilité et les contraintes à l'inégalité par rapport à une situation où seules les contraintes sur les probabilités de choix si les listes sont partielles, sont utilisées.

Anticipations rationnelles Calsamiglia et al. [2017] explique comment calculer, à moindre frais, les stratégies des élèves quand un mécanisme de Boston est utilisé et cette technique peut s'étendre à d'autres mécanismes. Les auteurs se servent du fait que les préférences des écoles sont données par un système de priorités complétées par un système de tirage aléatoire. En conditionnant par rapport à un système de probabilités d'acceptation dans chaque école, ils écrivent la valeur et le choix optimal de la liste ordonnée d'écoles en utilisant des outils de programmation dynamique à la Bellman. L'espace d'état est de dimension plus petite que le calcul des valeurs et choix optimaux par une méthode d'évaluation complète de la valeur de chaque liste.

Si la liste ordonnée est de taille R au plus, on commence par chercher la meilleure école parmi celles qui restent et ceci en utilisant les probabilités de succès déterministes qui sont fonction des seuils observés (voir Section 3.3). Cette meilleure école est alors obtenue avec une certaine probabilité et la valeur du programme à l'étape R s'en déduit comme une somme de l'espérance de valeur de cette école et de l'option extérieure, pondérée par leurs probabilités respectives. Puis, on raisonne par récurrence arrière en partant de la valeur obtenue à l'étape r et en en déduisant la valeur à l'étape $r - 1$ en utilisant le même type de raisonnement.

De façon générale, Agarwal et Somaini [2017] se servent de l'idée qu'une liste ordonnée, $L_i = (j_{i1}, \dots, j_{iR})$ d'écoles que les étudiants soumettent (de longueur réduite dans leur cas) forment un portefeuille, puisqu'une probabilité d'obtention, $\pi(j_{ir}; L_i)$ est attachée, à chacune des écoles,

j_{ir} de la liste L_i . Si on note $v_i(j_{ir})$ la valeur de cette école, la valeur de la liste s'en déduit:

$$v_i(L_i) = \sum_{r=1}^R \pi(j_{ir}; L_i) v_i(j_{ir}). \quad (6)$$

La liste observée et supposée optimale, L_i^* , est donc obtenue comme l'argument de $\max_{L_i} v_i(L_i)$. Comme les auteurs estiment de façon convergente $\pi(j_{ir}; L_i)$ dans une première étape décrite en Section 3.3, cette valeur ne dépend plus que des paramètres de préférences qui sont estimés de façon bayésienne dans un cadre de choix discret.

Agarwal et Somaini font l'hypothèse qu'il n'existe pas d'options extérieures. Carvalho et al. [2017] par contre tiennent compte d'une option extérieure qui complique l'estimation. D'abord, leurs données sont échantillonnées en fonction du choix des étudiants puisque seuls les étudiants intéressés par l'une des deux écoles (S ou F) sont observés. Si on normalise la valeur de l'option extérieure à 0 et si on note v_{iS} et v_{iF} la valeur des deux écoles pour l'étudiant i , seuls les étudiants tels que $\max(v_{iS}, v_{iF}) > 0$ sont observés. De plus, le choix pour un étudiant i tel que $v_{iS} > 0, v_{iF} < 0$ ne peut être que S . De même pour les étudiants tels que $v_{iS} < 0, v_{iF} > 0$ qui choisissent toujours F . Ce n'est que pour les étudiants $v_{iS} > 0, v_{iF} > 0$ que le cadre d'Agarwal et Somaini s'applique et les étudiants choisissent l'argument de $\max(\pi(S)v_S, \pi(F)v_F)$. Ce cadre conceptuel à options externes se généralise à un cadre d'écoles multiples, au prix de complications notationnelles.

3.5 Identification

On a vu que les probabilités de succès peuvent être identifiées de façon non paramétrique en utilisant des données sur les règles de priorité ou la distribution des notes. D'autre part, l'identification des préférences dans le cas où les listes ordonnées soumises révèlent les vraies préférences est déjà connue dans la littérature dans le cas de listes réduites à un seul élément (Matzkin [1993]) et ces résultats s'étendent, bien sûr, au cas où les listes sont plus longues. L'identification des préférences dans un cadre stratégique est plus délicate. Néanmoins, l'identification des préférences sous hypothèse de stabilité s'inspire aussi des modèles à choix discrets (Berry et Haile [2014]). En particulier, Diamond et Agarwal [2016] décrit les conditions d'identification des paramètres de modèles discrets à caractéristiques pures. Agarwal [2015] qui utilise cette dernière spécification et une restriction de stabilité par paire a besoin

d'une restriction d'exclusion de chaque côté du marché. Il se sert de l'endroit de naissance pour les étudiants et de la quantité de capital humain fournie, pour les universités. Dans cette littérature utilisant la restriction de stabilité, Menzel [2015] n'utilise pas ces restrictions d'exclusion mais n'identifie que les sommes des utilités des partenaires.

Dans un cadre de choix d'écoles, Agarwal et Somaini [2017] fournit des preuves d'identification quand certaines conditions sont satisfaites. La première condition est l'absence d'options extérieures et la normalisation d'une des options à zéro ou à une valeur dérivée du cadre spécifique d'analyse. La seconde condition peut être exprimée de deux façons. Il peut d'abord exister des variables sous-jacentes aux probabilités de succès, comme la variable, w_{ij} , dans l'équation (2) et ces variables sont supposées exclues des préférences. Ceci permet dans l'expression (6) de faire varier les probabilités $\pi(j_{ir}; L_i)$ indépendamment des préférences $v_i(j_{ir})$ ce qui permet d'identifier ces dernières, à condition que la variation des probabilités soit continue et complète c'est-à-dire varie continûment de 0 à 1. Ceci n'est pas toujours le cas surtout dans un cadre d'écoles/élèves où les priorités sont discrètes en général.

Deuxièmement Agarwal et Somaini [2017] propose une autre hypothèse d'exclusion qui est celle d'un régresseur "spécial" (Lewbel [2000]). Celui-ci affecte les préférences dans chaque école comme un facteur additif à un indice général d'utilité qui reste à identifier. Sous les conditions restrictives d'un régresseur spécial, d'exogénéité, de monotonie et de variation complète (Magnac et Maurin [2007]), Agarwal et Somaini [2017] montrent l'identification des indices de préférence pour chaque école. Le régresseur spécial, suggéré par les auteurs dans leur cadre, est la distance de la résidence de l'élève à chaque école.

Carvalho et al. [2017] dérive des résultats dans un cadre d'allocation d'étudiants à des universités plus que des élèves à des écoles. C'est pourquoi ils autorisent une option extérieure. Cette option extérieure ne permet pas l'identification cardinale des utilités quand une seule des options S ou F est valorisée par les agents et que l'échantillonnage est endogène ("choice-based"). Il faudrait pour cela faire une hypothèse d'existence de régresseur spécial mais pour les étudiants, la distance joue un rôle moins important que pour des élèves. Par contre, quand les valeurs des deux écoles v_S et v_F sont positives, leur preuve d'identification repose sur la variation des probabilités de succès dues à des variables exogènes sous-jacentes aux probabilités comme les histoires éducatives lointaines et utilise les preuves classiques de Manski [1988] et

Matzkin [1993] comme le font Agarwal et Somaini [2017].

4 Contrefactuels

On a déjà vu que l'on a besoin de connaître les préférences cardinales des étudiants pour estimer les contrefactuels puisqu'il semble impossible de trouver des cas où les listes ordonnées de choix sont les mêmes dans le contrefactuel et dans les données réelles sauf à supposer que les agents jouent de manière naïve. On commence par analyser le cas le plus compliqué d'anticipations rationnelles. Puis on décrit rapidement comment d'autres systèmes de croyances peuvent être utilisés. Pour fixer les idées, on supposera que le contrefactuel consiste à remplacer le système en vigueur par un nouveau mécanisme. Par exemple un mécanisme d'acceptation différée sans troncature ni coûts d'application, est remplacé par un mécanisme d'acceptation différée avec troncature ou coûts d'application (He et Magnac [2017]).

Anticipations rationnelles Celles-ci reposent sur le calcul d'un équilibre Nash-bayésien. On a vu que, pour l'estimation à partir de données dérivées de la grande majorité des mécanismes, l'équilibre se caractérise, pour chaque école, par des seuils qui sont des limites inférieures aux scores des étudiants qui seront acceptés à l'équilibre. Dans un cadre d'équilibre stable comme celui de Fack et al. [2015], le choix des étudiants consiste à choisir l'école dont le seuil est inférieur à leur score et qui est leur préférée parmi les écoles satisfaisant à cette condition. Dans d'autres cadres, Calsamiglia et al. [2017] ou Agarwal et Somaini [2017] spécifient l'algorithme qui permet de choisir les listes ordonnées (sous les contraintes spécifiques du mécanisme). Mais tout ceci se fait à seuils fixés, soit à leur valeur observée dans les données, soit celles obtenues par rééchantillonnage (Carvalho et al. [2017]).

Pour les contrefactuels, les seuils ne sont pas connus et il faut donc pour déterminer le nouvel équilibre, calculer ces nouveaux seuils. Ce calcul se fait en général par une méthode de point fixe et par simulation. On considère ainsi S états de la nature qui se caractérisent par la réalisation de la valeur aléatoire des préférences cardinales pour chaque étudiant ainsi que la réalisation de leurs notes ou priorités aléatoires et ceci conditionnellement à leurs caractéristiques. Le jeu est joué par les étudiants dans chaque état de la nature $s \in \{1, \dots, S\}$ et le point fixe cherché est

la distribution des seuils, t_j , de chaque école. Les étudiants calculent leurs stratégies optimales en réponse à la distribution des seuils et la distribution des seuils est celle des seuils d'équilibre donnés par exemple par un système d'équations (3) où les stratégies sont supposées optimales. En règle générale, les auteurs font l'hypothèse que les étudiants n'ont qu'une seule meilleure réponse (avec probabilité 1) sauf He [2015] qui prend en compte la multiplicité éventuelle des meilleures réponses. Cela rend plus incertaine l'évaluation des impacts.

L'algorithme est en général initialisé aux seuils observés dans les données pour tout s . Puis on dérive les stratégies optimales des étudiants pour ces seuils et ceci sous le mécanisme contrefactuel par exemple celui qui spécifie que la liste de choix possibles est tronquée. Dans chaque état de la nature s , on dérive des stratégies optimales précédentes, les seuils d'équilibre en utilisant un système d'équations (3). Ceci fournit, de façon temporaire, la distribution empirique simulée des seuils d'équilibre. En repartant de cette distribution empirique, on reformule les stratégies optimales des étudiants et on calcule les nouveaux seuils d'équilibre. On itère ce processus jusqu'à obtenir une distribution empirique simulée qui ne varie plus, en utilisant un critère d'arrêt raisonnable.

Si l'algorithme converge, le point fixe obtenu est un équilibre Nash-bayésien ou au moins une approximation finie de cet équilibre. Si on sait que l'équilibre est unique, comme dans le cas de mécanismes d'acceptation différée avec un nombre infini de joueurs (Fack et al [2015]), on a bien caractérisé l'équilibre contrefactuel. Il est à remarquer qu'on ne dispose pas de règles précises pour choisir le nombre de simulations et il n'est pas entièrement clair que les événements de faible probabilité soient bien saisis par les algorithmes que les divers auteurs ont utilisé.

Enfin, après que cet équilibre Nash-bayésien est estimé, on peut envisager plusieurs hypothèses pour en dériver les comportements contrefactuels des étudiants. En particulier, He et Magnac [2017] distingue deux cas. Le premier est un cadre d'équilibre Nash-bayésien *dans la population* puisque on calcule les comportements des étudiants appartenant à la population de référence sans le rapporter aux comportements observés dans l'échantillon. Seuls sont pris en compte les paramètres des préférences et des notes, estimés dans l'échantillon, et les caractéristiques exogènes observées des agents. Le deuxième cas est un cadre d'équilibre Nash-bayésien *dans l'échantillon* car on recalcule les comportements optimaux des étudiants en les conditionnant par les données observées dans l'échantillon et en particulier les listes ordonnées d'écoles

soumises par les étudiants. On reste ainsi proche de l'échantillon original. Le choix d'un mode de calcul des contrefactuels dépend du type de questions économiques que l'on se pose, au niveau de la population tout entière ou au niveau de l'échantillon particulier observé.

Autres croyances L'étendue des choix est infinie mais au contraire des anticipations rationnelles, ces croyances sont constantes et ne sont pas le résultat d'un équilibre comme ce qui est décrit dans la partie précédente. Certains auteurs, comme He [2015], utilisent d'abord des hypothèses implicites sur les croyances comme celles que font des étudiants "naïfs" qui révéleraient leurs vraies préférences même s'ils n'ont pas intérêt à le faire. Les étudiants peuvent aussi être naïfs ou sophistiqués avec une certaine probabilité à estimer. Agarwal et Somaini [2017] envisagent aussi le cas où les anticipations sont adaptatives et utilisent des données de coupe transversale pour leur implémentation puisque l'allocation est observée au cours de quatre années successives. Ils examinent aussi le cas d'anticipations agrégées ou grossières. Les agents ne conditionneraient pas à leurs caractéristiques spécifiques, la formation de leurs anticipations et n'anticiperaient que des probabilités agrégées de succès. C'est aussi le cas de Lufade [2018] où les anticipations ne changent pas quand le mécanisme contrefactuel est adopté ou change à la marge seulement. Les simulations donnent alors l'effet de court-terme du changement de mécanisme à anticipations constantes.

5 Résultats empiriques majeurs

Je ne commenterai pas ici les résultats obtenus par expérimentation directe et qui n'utilisent pas d'outils spécifiquement économétriques en comparant directement des résultats d'expériences. Cette littérature qui contraste aussi les résultats des mécanismes différents, et il y en a beaucoup, se développe trop vite pour que je puisse la couvrir de manière raisonnable dans cette courte revue.

Les auteurs dont j'ai étudié les méthodes, se sont d'abord intéressés à la question centrale de l'utilité sociale des mécanismes d'allocation centralisée. Ainsi, Abdulkadiroglu, Agarwal et Pathak [2017a] utilise les données sur le passage, en 2003, dans les lycées de New York, d'un système décentralisé avec faible coordination à un système centralisé et coordonné. Le

nouveau système, qui est une variante du mécanisme d'acceptation différée, est évalué par rapport à deux étalons: l'assignation à l'école du voisinage et l'allocation utilitariste optimale qui domine nécessairement la première allocation. Les auteurs montrent que le mécanisme utilisé arrive à capturer 80% des gains obtenus en passant d'un étalon à l'autre alors que le système non coordonné n'en capturerait que 45%. De plus, ce sont les étudiants qui avaient le plus de mal à trouver une école et étaient assignés administrativement par l'ancien mécanisme, qui gagnent le plus au nouveau système.

Agarwal [2015] utilise des données sur le mécanisme d'allocation centralisé le plus célèbre, celui des programmes de résidents médicaux étudiants aux Etats-Unis (Roth [1982]) et cherche à analyser si les salaires bas observés viennent d'un pouvoir monopsonique de la part des universités. De façon remarquable, il montre que cela n'est pas le cas et qu'il est probable que cela soit la qualité de l'investissement en capital humain que les résidents reçoivent, qui motive leur choix de résidence, beaucoup plus que les incitations financières.

Ajayi [2015] utilise des données provenant d'un mécanisme centralisé d'allocation d'élèves d'écoles primaires à des collèges au Ghana entre 2005 et 2008. Ce mécanisme est du type acceptation différée avec troncature de listes (qui varie au cours du temps). Elle montre d'abord que le degré de sophistication des étudiants varie positivement avec la qualité de leur école primaire mesurée par le score moyen des élèves à l'examen national qui sert de "préférences" pour les collèges. Elle montre aussi que les étudiants, provenant d'écoles primaires de meilleure qualité, candidatent à des collèges de meilleure qualité, ceci conditionnellement à la note de l'élève. Néanmoins, la distance à l'école et les coûts afférents ont quantitativement plus d'importance que la qualité du collège pour expliquer la variance des préférences. Elle utilise aussi deux changements observés du mécanisme. Une liste plus longue ou des restrictions sur la structure de la liste d'écoles en fonction de catégories, amoindrissent les différences entre étudiants provenant d'écoles primaires de différentes qualités.

Carvalho et al. [2017] étudie l'interaction entre les mécanismes de sélection et d'allocation. Certaines universités fédérales brésiliennes, jusqu'à une époque récente, utilisaient un système de sélection des étudiants par des examens en deux étapes. Le premier examen par questionnaire à choix multiples cherchait à éliminer les candidats et ne restait pour le deuxième examen plus approfondi que les étudiants ayant réussi au premier tour. Ce type de système, qui diminue les

coûts de sélection de manière importante pour les écoles, est d'ailleurs utilisé dans d'autres pays du monde (Japon, Corée du Sud par exemple) sous des habillages variés (Hafalir et al. [2015], Avery, Lee et Roth [2014]). Carvalho et al. [2017] montre que le réglage fin de l'écramage ne semble pas importer. Par contre, le passage à un mécanisme où les listes ordonnées de choix seraient plus longues serait bénéfique d'un point de vue utilitarier pour les étudiants et pour affiner la sélection pour les écoles car les motivations stratégiques diminuent. Enfin, l'inversion du calendrier des choix de filières et des notes semble conduire à des comportements plus opportunistes de la part des étudiants. Cela fait écho aux résultats trouvés par Wu et Zhong [2014] en Chine où ce même calendrier inversé dégrade la sélection d'étudiants car ceux-ci ont en moyenne une moindre réussite ultérieure dans leurs études.

Ensuite, un débat important porte sur la justification de l'utilisation du mécanisme d'acceptation différée ou de Boston (ou d'acceptance immédiate). Les études théoriques montrent que l'un et l'autre ont chacun leurs avantages (Ergin et Sonmez [2006]). L'acceptation différée n'induit pas de comportements stratégiques et est plus facile à gérer qu'un mécanisme de Boston. Par contre, il ne permet pas de révéler l'intensité des préférences des étudiants pour les collègues.

Ce débat dépend du degré de sophistication des élèves ou de leurs familles (Pathak et Sonmez [2013]). He [2015] s'intéresse à la crédibilité de l'hypothèse de sophistication en utilisant des données recueillies dans des collèges à Pékin en 1999 où l'allocation se fait par un mécanisme de Boston. D'après lui, le niveau de sophistication des parents est hétérogène mais leur degré de sophistication n'est pas corrélé avec leur niveau de revenu ou d'éducation. La raison principale des différences entre croyances subjectives et objectives vient du fait que les parents sont trop prudents. Ils considèrent les écoles les plus populaires moins souvent qu'ils ne le devraient en espérance d'utilité. Enfin, quels élèves ou familles gagnent ou perdent dans le cas où le mécanisme de Boston utilisé est remplacé par un mécanisme d'acceptation différée dépend de leur degré de sophistication et du nombre d'agents sophistiqués sur le marché. Conclure que le mécanisme de Boston affecte négativement les agents naïfs et profite aux agents sophistiqués n'est donc vrai que dans certains contextes. Les résultats de Hwang [2016], obtenues avec d'autres données, sont d'ailleurs différents de He [2015] dans certaines dimensions. Le premier de ces auteurs trouve que le bien-être ex-ante est plus important sous un mécanisme de Boston que sous acceptation différée mais il trouve aussi que les étudiants naïfs sont plus souvent

affectés à des écoles moins cotées.

Calsamiglia et al. [2017] utilise des données sur les entrées en école maternelle publique ou semi-publique à Barcelone dans les années 2006 et 2007 et s'intéresse aussi au degré de sophistication des agents. En particulier, une réforme de la carte scolaire eut lieu entre les deux années d'observation. Les préférences des agents sont estimées à partir des données de 2006 et les résultats sont utilisés pour prédire les effets de la réforme de 2007 en les comparant aux données réelles. L'adéquation est très bonne dans des dimensions comme la prédiction de l'allocation des élèves à une école dans leur zone de carte scolaire.

D'autres expériences contrefactuelles intéressent les auteurs puisque le mécanisme employé à Barcelone est de type Boston ou d'acceptation immédiate. Calsamiglia et al. [2017] évalue d'abord ce que serait l'effet de remplacer ce mécanisme par un mécanisme d'acceptation différée qui sous une hypothèse de préférences strictes est meilleur d'un point de vue de Pareto (Sonmez et Unver [2011]). Néanmoins, comme le système d'accès aux écoles maternelles est un système de priorités – des points sont donnés pour la fratrie ou les écoles de quartier – il faut des tirages aléatoires pour départager certaines priorités dans certaines écoles. Cela peut renverser l'ordre au sens de Pareto entre mécanisme de Boston et d'acceptation différée (Ergin et Sonmez [2006]) et c'est ce qui se passe en pratique à Barcelone. De plus, le mécanisme d'acceptation différée augmente l'inégalité entre zones de la ville. C'est tout le contraire avec un autre mécanisme, celui de "cycle d'échange optimal" (Top Trading Cycles), qui est rarement utilisé dans le cadre d'allocation d'étudiants aux écoles mais qui est Pareto-optimal (Abdulkadiroglu et Sonmez [2003]). Il est peu surprenant que les auteurs trouvent qu'il est meilleur à Barcelone, dans un sens utilitariste, mais il semble aussi plus égalitaire entre zones. En particulier, beaucoup de ménages obtiennent leurs choix favoris hors de la zone.

Fack et al. [2015] utilise des données sur l'allocation d'élèves aux lycées dans un district de Paris en 2013 qui est effectuée au moyen de la procédure Affelnet (Hiller et Tercieux [2014]) qui est une variante d'un mécanisme d'acceptation différée. La révélation des vraies préférences est rejetée par les données au profit du concept de stabilité alors celui-ci ne l'est pas vis à vis d'une alternative instable. Les auteurs, pour estimer les préférences dans ce dernier cas, utilisent seulement des inégalités sur les probabilités de choix de telle ou telle alternative.

Le premier contrefactuel que ces auteurs étudient consiste à remplacer le système de notes et

les priorités pour les plus défavorisés par un système de notes seulement. Ceci accroît fortement la ségrégation par statut socio-économique. Par contre le remplacement du système de priorité actuel par un système de tirage aléatoire n'affecte que modérément la ségrégation et diminue bien sûr de beaucoup la sélection par le mérite. En fait, c'est un mélange entre sélection au mérite pour les écoles les plus cotées et un tirage aléatoire pour les autres qui pourrait équilibrer les deux effets mais les effets différentiels sur le bien-être des étudiants les moins et les plus académiquement méritants sont importants.

D'autre part, Agarwal et Somaini [2017] utilise des données de Cambridge (Etats-Unis) d'allocation d'élèves aux écoles primaires par un mécanisme de type Boston où trois choix seulement sont possibles. Ce mécanisme est complété par une règle d'allocation ex-post des étudiants qui ne sont pas assignés à une école. Les choix sont donc a priori stratégiques et il est peu recommandable de mettre une école très populaire en deuxième et troisième position sur la liste ordonnée de choix. Les auteurs montrent d'abord que le comportement des étudiants change de manière discontinue à la frontière des zones de résidence ce qui prouve l'existence d'effets stratégiques, sous-tendant leurs décisions. En utilisant les méthodes d'estimation vues plus haut, ils montrent que même si 83% des élèves semblent alloués à leur premier choix déclaré, ce nombre tombe à 72% quand on considère leurs vraies préférences. Néanmoins, ce mécanisme utilisé est préféré de manière utilitarienne à un mécanisme d'acceptation différée et les auteurs quantifient cette différence en termes d'un indicateur de bien-être qui est la distance à l'école. Cet avantage reste petit (0.07 miles). L'avantage du mécanisme de Boston reste surtout dû à la présence d'étudiants à qui la municipalité n'offrent pas de déjeuner gratuit, et ceci, pour des raisons de congestion. Néanmoins, ce mécanisme est instable et génère donc de la jalousie légitime pour entre 4 et 13% des étudiants, suivant la spécification.

Enfin, He et Magnac [2017] s'intéresse à l'arbitrage entre coûts de sélection dus à la congestion et coûts d'application quand on utilise des mécanismes d'acceptation différée. Les données proviennent d'une expérience de terrain parmi les étudiants de Master à l'école d'économie de Toulouse pour le choix de spécialité en deuxième année. Pour estimer les préférences des étudiants, ils utilisent les listes ordonnées soumises par les étudiants quand le mécanisme d'acceptation différée avec liste complète est utilisé. Dans des expériences contrefactuelles, ils comparent les résultats du mécanisme existant, conditionnellement ce qu'ils observent dans

l'échantillon, à des mécanismes à acceptation différée avec troncature des listes à $1, \dots, K$ éléments ou comportant des coûts d'application variant de 0 à des montants substantiels. La troncature ou l'existence de coûts d'application implique que les listes sont plus courtes et cela diminue fortement les coûts de sélection des programmes de master. Néanmoins, cela dégrade aussi la qualité de l'allocation finale en termes de stabilité. He et Magnac [2017] quantifient les arbitrages entre dégradation de qualité et gains en termes de coûts d'évaluation des dossiers d'étudiants et montrent que la troncature modérée des listes ou des coûts d'application modérés semblent être des mécanismes optimaux.

6 Conclusion

Quand j'ai commencé à travailler il y a quelques années sur les données brésiliennes avec mon coauteur José Raimundo Carvalho de l'Université Fédérale du Ceará, peu des outils que j'ai présenté ici existaient déjà à part dans le travail précurseur de Yinghua He et les travaux expérimentaux (Buddish et Cantillon [2012], Calsamiglia, Haeringer et Klijn [2010]) auxquels je n'ai pas pu donner l'importance qu'il se doit dans cette revue. A l'heure d'aujourd'hui, les données provenant de mécanismes d'allocation d'étudiants aux universités deviennent disponibles à travers le monde (voir le Tableau 1 de Agarwal et Somaini [2017], et le site "Matching in practice in Europe") et il est facile de prévoir que ces recherches empiriques continueront à se développer dans un futur proche.

Ce que cette revue a cherché à faire est de montrer la puissance des outils qui ont été mis à la disposition des chercheurs dans les différents travaux que j'ai analysés. On voit par exemple que ces analyses tiennent compte du fait que la révélation d'informations de la part des étudiants est stratégique ce qui en fait un modèle exemplaire pour la prise en compte de tels aspects dans la collecte de données (McFadden [2009]). Cette revue montre aussi leurs zones d'ombre et c'est vers leur éclaircissement que ce domaine va ou pourrait aller. Par exemple:

- Les écoles ont un rôle passif et non stratégique. Or on sait que, par exemple pour un mécanisme d'acceptation différée où les étudiants proposent, les écoles ont des incitations à manipuler les critères d'acceptation – par exemple en omettant les étudiants inatteignables.

- Le choix entre mécanismes peut être étudié dans un cadre d'économie politique puisque certains agents perdent et d'autres gagnent à l'utilisation de tel ou tel mécanisme.
- les mécanismes se basent sur des structures de coûts de recherche, d'application et de sélection qui ne sont pas toujours modélisés mais qui peuvent avoir une influence cruciale sur l'issue de l'allocation.
- les options extérieures ne sont pas toujours présentes et pas toujours modélisées comme il se doit. Pourtant elles sont cruciales en particulier quand elles ont des implications dynamiques.

Nous avons laissé de côté un certain nombre de directions de recherche. Par exemple, Fu [2014] estime un modèle où les préférences des universités et des étudiants jouent un rôle plus symétrique en suivant le modèle théorique de Chade, Lewis et Smith [2015]. Les universités choisissent leurs coûts de scolarité et les étudiants, les universités auxquels ils candidatent ainsi que celle qu'ils choisissent après que les offres sont faites. Pistoletti [2017] analyse des situations, comme dans les universités françaises, où les étudiants apprennent leurs probabilités de réussite dans les filières qu'il pourraient choisir et en quoi cela modifie leurs choix dans la procédure APB. Et il y a beaucoup d'autres exemples d'analyses que je n'ai pu traiter.

A l'heure où la première version de cet article était écrite, la ministre de l'Enseignement supérieur annonçait que la procédure APB était morte.¹¹ Je formulais le souhait que les travaux théoriques et empiriques, riches d'enseignement, qui ont été développés dans les quinze dernières années servent de base à la réflexion politique sur le nouvel outil qui serait utilisé en France ces prochaines années pour allouer les candidats aux filières universitaires. Cela n'a pas été le cas. La nouvelle version Parcoursup a déjà attiré les critiques de certains économistes comme Lucien Frys, Julien Grenet, Guillaume Haeringer et Vincent Iehlé dans de nombreuses tribunes.¹² En particulier, ceux-ci arguent que la forte décentralisation de la procédure en plusieurs étapes sans ordre précis des dix choix dans la première étape, a de grandes chances d'exacerber les

¹¹<http://www.lejdd.fr/societe/education/frederique-vidal-la-plateforme-apb-ce-sera-terme-lan-prochain-3425705>

¹²Par exemple, http://etudiant.lefigaro.fr/article/parcoursup-avec-dix-voeux-le-risque-de-bacheliers-sans-affectation-augmente-_9d1b8e92-d670-11e7-8428-569ae9712d9b/

problèmes de congestion plus encore que la défunte APB. Le temps qu'il faudra aux étudiants premiers classés pour refuser certaines de leurs offres ralentirait l'affectation. Il ne serait donc pas étonnant qu'à l'heure de publication de cet article, prévue pour septembre 2018, la nouvelle procédure n'ait pas réussi à trouver une place acceptable pour une fraction notable des étudiants.

Bibliographie

Abdulkadiroğlu, A., Agarwal, N., & Pathak, P. A., [2017a], "The Welfare Effects of Congestion in Uncoordinated Assignment: Evidence from the NYC HS Match", *American Economic Review*, 107(12) p. 3635-89..

Abdulkadiroğlu, A., Angrist, J.D., Y. Narita, & Pathak, P. A., [2017b], "Research Design Meets Market Design: Using Centralized Assignment for Impact Evaluation", *Econometrica*, 85(5) p. 1373-1432.

Abdulkadiroğlu, A., Pathak, P., Roth, A. E., & Sonmez, T., [2006], "Changing the Boston school choice mechanism", WP 11965, National Bureau of Economic Research.

Abdulkadiroğlu, A., & Sonmez, T., [1998], "Random serial dictatorship and the core from random endowments in house allocation problems", *Econometrica*, 66(3), p.689-701.

Abdulkadiroğlu, A. & T., Sonmez, [2003], "School Choice: A Mechanism Design Approach", *American Economic Review*, 93(3), p. 729-747

Agarwal, N., [2015], "An empirical model of the medical match", *The American Economic Review*, 105(7), p. 1939-1978.

Agarwal, N., & P., Somaini, [2017], "Demand Analysis using Strategic Reports: An Application to a School Choice Mechanism", forthcoming *Econometrica*.

Ajayi, K. F., [2015], "School Choice and Educational Mobility: Lessons from Secondary School Applications in Ghana", unpublished manuscript.

Akyol, P., & Krishna, K., [2017], "Preferences, selection, and value added: A structural approach", *European Economic Review*, 91, p.89-117.

Arcidiacono, P., [2005], "Affirmative Action in Higher Education: How Do Admission and Financial Aid Rules Affect Future Earnings?", *Econometrica*, 73(5), p. 1477-1524.

Arnosti, N., R. Johari, & Y. Kanoria, [2014] "Managing congestion in dynamic matching markets," In *Proceedings of the fifteenth ACM conference on Economics and computation*, p. 451-451.

Avery, C., Lee, S., & Roth, A. E. [2014], "College Admissions as Non-Price Competition: The Case of South Korea", No. w20774, National Bureau of Economic Research.

Azevedo, E.M., & J.D., Leshno, [2016], "A Supply and Demand Framework for Two-Sided Matching Markets", *Journal of Political Economy*, 124(5), p. 1235-1268.

Balinski M., & T., Sönmez, [1999], "A Tale of Two Mechanisms: Student Placement", *Journal of Economic Theory*, 84, p. 73-94.

Bardella, F. P., & Sotomayor, M. [2014]. "Redesenho e Análise do Mercado de Admissão aos Centros de Pós-Graduação em Economia no Brasil à Luz da Teoria dos Jogos: Um

Experimento Natural em Desenho de Mercados", *Revista Brasileira de economia*, 68(4), p. 425-455.

Berry, S. T., & Haile, P. A., [2014], "Identification in differentiated products markets using market level data", *Econometrica*, 82(5), p1749-1797.

Budish, E., & Cantillon, E. [2012], "The multi-unit assignment problem: Theory and evidence from course allocation at Harvard", *American Economic Review*, 102(5), p. 2237-71.

Burgess, S., Greaves, E., Vignoles, A., & Wilson, D., [2015], "What parents want: School preferences and school choice", *The Economic Journal*, 125(587), p. 1262-1289.

Calsamiglia, C., C., Fu & M.Güell, [2014], "Structural Estimation of a Model of School Choices: the Boston Mechanism vs its alternatives", WP 2014-21, Universidad Autonoma, Barcelona.

Calsamiglia, C., Haeringer, G., & Klijn, F. [2010], "Constrained school choice: An experimental study", *American Economic Review*, 100(4), p. 1860-74.

Carvalho, J. R., Magnac, T., & Xiong, Q., [2017], "College Choice, Selection and Allocation Mechanisms: A Structural Empirical Analysis", unpublished manuscript.

Chade, H., Lewis, G., & Smith, L., [2014], "Student portfolios and the college admissions problem", *The Review of Economic Studies*, 81(3), 971-1002.

Chade, H., & Smith, L., [2006], "Simultaneous search", *Econometrica*, 74(5), p. 1293-1307.

Che, Y.K., & Y. Koh, [2016], "Decentralized College Admissions", *Journal of Political Economy*, 124, p. 1295-1338.

Chen, Y., & Kesten, O., [2017], "Chinese college admissions and school choice reforms: A theoretical analysis", *Journal of Political Economy*, 125(1), p. 99-139.

Chen, Y., & Sönmez, T. [2006], "School choice: an experimental study", *Journal of Economic theory*, 127(1), p. 202-231.

Diamond, W., and Agarwal, N. [2017]. Latent indices in assortative matching models. *Quantitative Economics*, 8(3), p. 685-728.

Dubey, P., O. Haimanko and A., Zapechelnuyk, [2006], "Strategic Complements and Substitutes and Potential Games", *Games and Economic Behavior*, p. 54:77-94.

Epple, D., R. Romano & H. Sieg, [2006], "Admission, Tuition, and Financial Aid Policies in the Market for Higher Education", *Econometrica*, 74(4), p. 885-928.

Ergin, H., & Sönmez, T., [2006], "Games of school choice under the Boston mechanism", *Journal of Public Economics*, 90(1), p. 215-237.

Fack, G., & Grenet, J., [2010], "Que peut-on attendre de la réforme de la sectorisation en France? Quelques enseignements des politiques de choix scolaire", *Revue d'économie politique*,

120(5), p. 709-737.

Fack, G., J., Grenet & Y., He, [2015], "Estimating Preferences in School Choice Mechanisms", unpublished manuscript.

Forges, F., Haeringer, G., & Iehlé, V., [2013], "Appariement: des modèles de Lloyd Shapley à la conception de marchés d'Alvin Roth", *Revue d'économie politique*, 123(5), p. 663-696.

Fox, J. T., [2010], "Identification in matching games", *Quantitative Economics*, 1(2), p. 203-254.

Fu C., [2014], "Equilibrium tuition, applications, admissions, and enrollment in the college market", *Journal of Political Economy*, 122(2), p. 225-281.

Gale, David E. & Lloyd S. Shapley, [1962], "College Admissions and the Stability of Marriage," *American Mathematical Monthly*, 69(1), p. 9-15.

Galichon, A., & Salanié, B., [2015], "Cupid's invisible hand: Social surplus and identification in matching models", unpublished manuscript.

Gazmuri, A.M., [2016], "School Segregation in the Presence of Student Sorting and Cream-Skimming: Evidence from a School Voucher Reform", mimeo TSE.

Haeringer, G., & Iehlé, V., [2010], "Enjeux stratégiques du concours de recrutement des enseignants-chercheurs", *Revue économique*, 61(4), p. 697-721.

Haeringer, G. & F.Klijn, [2009], "Constrained School Choice," *Journal of Economic Theory*, 144(5), p. 1921-1947.

Hafalir, I.E., R., Hakimov, D. Kübler & M. Kurino, [2014], "College Admissions with Entrance Exams: Centralized versus Decentralized, WP 2014-208, WZB, Berlin.

Hastings, J., T. J. Kane, & D. O. Staiger, [2009], "Heterogenous Preferences and the Efficacy of Public School Choice," Working paper, Yale University.

He, Y., [2015], "Gaming the School Choice Mechanism in Beijing", unpublished manuscript, Toulouse School of Economics.

He, Y. & T., Magnac, [2017], "A Pigouvian Approach to Congestion in Matching Markets", unpublished manuscript.

Hiller, V. et O.Tercieux, [2014], "Choix d'écoles en France", *Revue économique*, 65(3), p. 619-656.

Hwang, I. M., [2015], "A robust redesign of High School Match", unpublished manuscript.

Kapor, A., C.A.Neilson, S.D., Zimmerman, [2017], "Heterogeneous Beliefs and School Choice Mechanisms", mimeo, Princeton.

Lewbel, A., [2000], "Semiparametric qualitative response model estimation with unknown heteroscedasticity or instrumental variables", *Journal of Econometrics*, 97(1), p. 145-177.

- Lufade, M.**, [2018], "The Value of Information in Centralized School Choice Systems", mimeo Duke.
- Machado, C., & Szerman, C.**, [2015], "The Effects of a Centralized College Admission Mechanism on Migration and College Enrollment: Evidence from Brazil", unpublished manuscript.
- Magnac, T., & Maurin, E.**, [2007], "Identification and information in monotone binary models", *Journal of Econometrics*, 139(1), p. 76-104.
- Manski, C. F.**, [1988], "Identification of binary response models", *Journal of the American Statistical Association*, 83(403), p. 729-738.
- Matzkin, R. L.**, [1993], "Nonparametric identification and estimation of polychotomous choice models", *Journal of Econometrics*, 58(1), p. 137-168.
- McFadden, D.**, [2009], "The human side of mechanism design: a tribute to Leo Hurwicz and Jean-Jacques Laffont", *Review of Economic Design*, 13(1), p. 77-100.
- McKinney, C. N., Niederle, M., & Roth, A. E.**, [2005], "The Collapse of a Medical Labor Clearinghouse (and Why Such Failures Are Rare)", *American Economic Review*, 95(3), p. 878-889.
- Menzel, K.** [2015], "Large Matching Markets as Two-Sided Demand Systems", *Econometrica*, 83(3), p. 897-941.
- Pathak P.A., & T., Sonmez**, [2013], "Leveling the Playing Field: Sincere and Sophisticated Players in the Boston Mechanism," *American Economic Review*, 98(4), p. 1636–1652.
- Pistolesi, N.**, [2017], "Advising Students on their Field of Study", *Labour Economics*, 44, p. 106–121
- Roth, Alvin E.**, [1982], "The Economics of Matching: Stability and Incentives," *Mathematics of Operations Research*, 7 (4), p. 617-628.
- Roth, A.E**, [2008], "Deferred acceptance algorithms: history, theory, practice, and open questions", *International Journal of Game Theory*, 36, p.537–569
- Roth, A. E., & Sotomayor, M. A. O.**, [1992], *Two-sided matching: A study in game-theoretic modeling and analysis* (No. 18). Cambridge: Cambridge University Press.
- Roth, A. E., & Xing, X.**, [1994], "Jumping the gun: Imperfections and institutions related to the timing of market transactions", *The American Economic Review*, 84(4), p. 992-1044.
- Sönmez, T., & Ünver, M. U.**, [2011], "Matching, allocation, and exchange of discrete resources", *Handbook of Social Economics*, 1, p. 781-852.
- Wu, B., & Zhong, X.**, [2014], "Matching mechanisms and matching quality: Evidence from a top university in China", *Games and Economic Behavior*, 84, p. 196-215.