



HAL
open science

La transparence et l'obstacle : principes et enjeux des algorithmes d'appariement scolaire

Julien Grenet

► **To cite this version:**

Julien Grenet. La transparence et l'obstacle : principes et enjeux des algorithmes d'appariement scolaire. Melchior Simioni; Philippe Steiner. Comment ça matche? Une sociologie de l'appariement, Presses de Sciences Po, pp.101-138, 2022, collection "Gouvernances", 9782724639001. hal-03686961

HAL Id: hal-03686961

<https://hal.science/hal-03686961>

Submitted on 3 Jun 2022

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

La transparence et l'obstacle : principes et enjeux des algorithmes d'appariement scolaire

Julien Grenet*

Paris School of Economics/CNRS

Juin 2022

Publié dans : Melchior Simioni et Philippe Steiner (éd.), *Comment ça matche ? Une sociologie de l'appariement*, Presses de Sciences Po, collection « Gouvernances », Paris, 2022 (chapitre 3, p. 101-138).

Pour une analyse critique des procédures utilisées en France pour affecter les élèves et les étudiants, voir Grenet J., « Les algorithmes d'affectation dans le système éducatif français », in Simioni M. et Steiner Ph., éd., *Comment ça matche ? Une sociologie de l'appariement*, Presses de Sciences Po, collection « Gouvernances », Paris, 2022 (chapitre 1, p. 21-59). [\[lien\]](#)

Résumé

Ce chapitre propose un tour d'horizon des enjeux théoriques et pratiques de l'utilisation des algorithmes dans la régulation des inscriptions scolaires et universitaires. Il présente le principe et les enjeux de l'appariement ainsi que les principaux algorithmes qui ont été proposés dans la littérature académique pour affecter les élèves et les étudiants aux établissements d'enseignement. Une attention particulière est portée à l'algorithme d'acceptation différée de Gale et Shapley (1962), qui s'est imposé comme le plus à même de résoudre les arbitrages complexes auxquels sont confrontées les procédures d'affectation.

CODES JEL : C78, D47, I23, I24

MOTS CLÉS : *Algorithmes d'appariement, mariages stables, deferred acceptance mechanism, Boston mechanism, top trading cycles.*

*PSE-École d'économie de Paris, 48 boulevard Jourdan, 75014 Paris; e-mail : ju-lien.grenet@psemail.eu.

Introduction

Les mécanismes d'appariement jouent un rôle croissant dans la régulation du « marché » éducatif (Grenet, 2022). À l'origine, ces procédures ont été conçues pour simplifier et améliorer la gestion des inscriptions scolaires et universitaires. Elles présentent en effet de nombreux avantages par rapport aux procédures dites « décentralisées » dans lesquelles les étudiants et les formations interagissent directement sans qu'il y ait de coordination des décisions au moyen d'une plateforme. Les procédures décentralisées se révèlent le plus souvent lentes, inefficaces et inéquitables : du côté des formations, elles créent de l'incertitude sur les effectifs qui s'inscriront effectivement à la rentrée ; du côté des étudiants, elles induisent des phénomènes de congestion et de file d'attente qui les incitent à adopter des comportements stratégiques pénalisant *in fine* les étudiants les moins bien informés. Plus généralement, la décentralisation des inscriptions scolaires souffre d'un manque de transparence qui nuit à l'acceptation des décisions d'affectation. Pour résoudre ces difficultés, les procédures d'inscriptions centralisées s'appuient sur un algorithme qui permet de déléguer les décisions d'affectation à une plateforme à partir de trois types d'informations : les capacités d'accueil des formations, les vœux d'inscription des candidats et le classement des candidats par les formations. Muni de ces informations, l'algorithme d'affectation calcule un appariement qui tient compte de manière optimale des préférences réciproques des candidats et des formations.

Les procédures utilisées en France pour allouer les élèves aux écoles et les étudiants aux formations d'enseignement supérieur font l'objet d'un rejet, qui peut sembler paradoxal. Celui-ci tient pour partie à la nature de la régulation qui s'applique aux marchés scolaires. Contrairement aux marchés traditionnellement analysés par les économistes, les mécanismes de prix ne jouent dans le système éducatif qu'un rôle secondaire pour équilibrer l'offre et la demande. Dans le cadre des procédures centralisées d'affectation, les participants ne sont pas libres d'être appariés avec l'établissement qu'ils souhaitent : ils doivent également être « choisis », selon la formule célèbre employée par Alvin Roth pour décrire les marchés d'appariement (Roth, 2008b). La régulation des inscriptions scolaires et universitaires fait donc intervenir non seulement les préférences des étudiants pour les formations, mais également les « préférences » des formations pour les étudiants, ces préférences pouvant être soit exprimées par les formations elles-mêmes, soit déterminées par

la puissance publique à travers des règles de priorité. Cette forme de régulation crée par nature des conflits de répartition qui nécessitent un arbitrage entre des objectifs de politique éducative souvent difficiles à concilier : garantir une scolarisation proche du domicile, favoriser l'adéquation entre le profil des élèves et les caractéristiques des formations, limiter les inégalités d'accès aux ressources éducatives, etc. La quête de ce point d'équilibre entre choix individuels et contraintes collectives dans la régulation des marchés scolaires constitue le principal défi que tentent de résoudre les algorithmes d'appariement.

Ce chapitre propose un tour d'horizon des enjeux théoriques et pratiques de l'utilisation des algorithmes dans la régulation des inscriptions scolaires et universitaires. Il présente le principe et les enjeux de l'appariement ainsi que les principaux algorithmes qui ont été proposés dans la littérature académique pour affecter les élèves et les étudiants aux établissements d'enseignement. Une attention particulière est portée à l'algorithme d'acceptation différée de Gale et Shapley (1962), qui s'est imposé comme le plus à même de résoudre les arbitrages complexes auxquels sont confrontées les procédures d'affectation. Il constitue ainsi un complément à Grenet (2022) qui propose, à la lumière de la théorie de l'appariement et des travaux empiriques réalisés dans son sillage, une analyse critique des procédures qui sont utilisées en France pour affecter les élèves et étudiants, tant au niveau de l'enseignement secondaire (procédure Affelnet au lycée et expérimentation des secteurs multi-collèges) qu'à l'entrée dans l'enseignement supérieur (plateformes APB et Parcoursup).

1 Le problème de l'appariement scolaire : définitions et aspects normatifs

Issue de la théorie du *mechanism design*, la théorie de l'appariement (ou *matching*) s'intéresse au fonctionnement de « marchés » particuliers, qui ont pour fonction d'apparier deux types d'agents complémentaires, par exemple des élèves et des écoles, des internes en médecine et des hôpitaux, ou encore des donateurs et des receveurs d'organes¹. À la différence des marchés classiques où l'offre et la demande s'équilibrent par l'intermédiaire d'un mécanisme de prix, les marchés d'appariement (*matching markets*) ne font pas intervenir de transferts monétaires – ou seulement de manière secondaire. La caractéristique essen-

1. Pour une présentation synthétique de cette littérature, voir Erdil et Ergin (2013).

tielle de ces marchés est que les transactions y prennent la forme non pas d'un échange anonyme, mais d'un appariement entre agents qui doivent se « choisir » mutuellement.

Les travaux issus de cette branche de l'économie ont contribué à montrer comment ces marchés, longtemps ignorés par la théorie économique, pouvaient être efficacement réorganisés dans le cadre de procédures centralisées mettant en œuvre des algorithmes d'appariement. Ces recherches ont donné lieu à de nombreuses applications pratiques qui ont permis d'améliorer le fonctionnement de « marchés » tels que l'affectation des internes en médecine aux hôpitaux (Roth et Peranson, 1990), les programmes de dons croisés de reins à grande échelle (Roth, Sönmez et Ünver, 2004, 2005; Steiner, 2010) ou encore l'attribution de logements sociaux (Hylland et Zeckhauser, 1979; Abdulkadiroğlu et Sönmez, 1999). On se focalisera ici sur le domaine dans lequel la théorie de l'appariement a le plus profondément bouleversé les pratiques : l'affectation des élèves et des étudiants aux établissements d'enseignement. Après avoir précisé les définitions et les critères normatifs utilisés pour modéliser le problème de l'appariement dans ce contexte particulier, on présentera les principaux algorithmes d'affectation proposés dans la littérature et on décrira leurs propriétés².

1.1 Les procédures de choix scolaire et leurs composantes

Toute procédure de choix scolaire comporte quatre composantes distinctes (Cantillon, 2017) : 1) la définition de la capacité d'accueil de chaque formation ; 2) des règles de participation, qui organisent la manière dont les candidats indiquent leurs préférences pour les formations ; 3) des règles qui déterminent les priorités relatives des candidats, c'est-à-dire la manière dont ils sont classés par chaque formation ; 4) un algorithme (également appelé « mécanisme ») qui détermine la formation d'affectation de chaque candidat, en fonction préférences exprimées par ces derniers et de leurs priorités relatives.

Les règles de priorité visent à répondre à la question suivante : lorsque le nombre de candidats excède la capacité d'accueil d'une formation, lesquels se verront proposer une place et lesquels seront refusés ? Dans la mesure où elles créent des « gagnants » et des « perdants », ces règles sont de nature éminemment politique. Dans le contexte de la régulation des choix scolaires, les règles de priorité les plus fréquemment utilisées sont

2. Le lecteur désireux d'approfondir les aspects théoriques des algorithmes d'appariement dans le contexte dans la régulation des choix scolaires pourra se reporter aux synthèses éclairantes proposées par Pathak (2011) et Abdulkadiroğlu (2021).

la distance entre le domicile de l'élève et l'école, les résultats scolaires, le fait d'avoir un frère ou une soeur déjà scolarisé(e) dans l'établissement demandé et, plus, rarement, des critères sociaux (statut de boursier, revenu des parents).

Les règles de priorité sont généralement structurées selon deux types de schémas. Les priorités sont dites « strictes » lorsqu'elles permettent de classer tous les élèves sans *ex æquo*, à la manière des barèmes qui calculent les priorités à l'aide de points attribués selon différents critères, dont au moins un critère « continu » (comme la moyenne pondérée des notes de l'élève ou la distance entre son domicile et l'école demandée). Les priorités sont dites « larges » lorsque les critères prévus ne produisent pas un classement univoque des élèves mais seulement des classes d'équivalence, c'est-à-dire des « groupes de priorité ». Pour obtenir un ordre de priorité strict, la procédure généralement suivie consiste à départager les élèves d'un même groupe de priorité par tirage au sort ³.

La capacité d'accueil des formations, les vœux des candidats et leurs priorités relatives pour les différentes formations constituent les trois informations clés à partir desquels l'algorithme d'affectation réalise un appariement entre candidats et formations. Pour pouvoir être traités par l'algorithme, les vœux et les priorités sont exprimés sous la forme de classements : le classement des formations par les candidats, d'une part, et le classement des candidats par les formations, d'autre part.

Les vœux des candidats prennent la forme de listes ordonnées. En utilisant la lettre i pour désigner les candidats et la lettre f pour désigner les formations, la liste ordonnée de vœux d'un candidat i_1 peut s'écrire comme suit :

$$i_1 : (f_2, f_1, f_3)$$

Cette liste indique que le candidat i_1 a classé la formation f_2 en 1^{er} vœu, la formation f_1 en 2^e vœu et la formation f_3 en 3^e vœu.

Les classements des candidats sont déterminés soit par les préférences des formations elles-mêmes, soit par des règles de priorité fixées en amont de la procédure. La nature exogène des priorités des candidats est une condition essentielle au bon fonctionnement des algorithmes décrits dans ce chapitre : elle exclut, en particulier, que les priorités des candidats dépendent de la manière dont ils ont eux-mêmes classé les formations.

3. Sur les propriétés des procédures de tirage au sort qui peuvent être mises en œuvre pour départager les élèves, voir Erdil et Ergin (2008) et Abdulkadiroğlu, Pathak et Roth (2009).

En supposant que la formation f_1 a reçu trois candidatures émanant des candidats i_1 , i_2 et i_3 , et que le candidat i_1 a priorité sur le candidat i_3 , qui a lui-même priorité sur le candidat i_2 , le classement des candidats par la formation f_1 s'écrit comme suit :

$$f_1 : (i_1, i_3, i_2)$$

Muni des classements réciproques des candidats et des formations, l'algorithme d'affectation réalise un appariement (*matching*) en affectant chaque candidat à une seule formation (ou en le laissant sans affectation), en veillant à ce que le nombre de candidats affectés à une formation n'excède pas sa capacité d'accueil, qui est elle-même désignée par la lettre q . L'appariement obtenu, que l'on désigne généralement par la lettre grecque μ , consiste en une série de paires associant chaque candidat avec sa formation d'accueil, par exemple :

$$\mu = \begin{pmatrix} i_1 & i_2 & i_3 & i_4 \\ f_1 & f_2 & f_2 & \emptyset \end{pmatrix}$$

Ici, le candidat i_1 est affecté à la formation f_1 et les candidats i_2 et i_3 sont affectés à la formation f_2 . Le candidat i_4 n'est quant à lui affecté à aucune formation, situation que l'on représente par le symbole \emptyset (ensemble vide).

1.2 Qu'est-ce qu'un « bon » algorithme d'affectation ?

De très nombreux algorithmes ont été proposés dans la littérature pour réguler les choix scolaires. Pour guider le choix entre ces algorithmes, la théorie du *matching* a identifié trois critères normatifs qui permettent d'évaluer et de comparer leurs performances : l'efficacité, l'équité et la non-manipulabilité.

1) *L'efficacité (satisfaction des préférences)*. La performance d'un algorithme d'affectation est d'abord évaluée à l'aune de sa capacité à satisfaire les préférences exprimées par les candidats. Le critère retenu est celui de l'efficacité au sens de Pareto, du nom de l'économiste italien Vilfredo Pareto qui proposa ce critère pour caractériser une allocation des ressources telle qu'on ne peut améliorer la satisfaction d'un individu sans détériorer celle d'au moins un autre. Appliqué à la régulation des choix scolaires, ce critère stipule qu'un appariement entre élèves et formations est plus efficace qu'un second appariement si tous les candidats ont une satisfaction au moins aussi élevée dans le premier que dans

le second et qu'au moins un candidat a une satisfaction strictement plus élevée.

Pour illustrer ce concept, considérons deux candidats, i_1 et i_2 , et deux formations, f_1 et f_2 , qui peuvent chacune accueillir au maximum un candidat. Supposons que les préférences des candidats sont comme suit :

$$i_1 : (f_1, f_2); \quad i_2 : (f_2, f_1)$$

Considérons les deux appariements suivants :

$$\mu_1 = \begin{pmatrix} i_1 & i_2 \\ f_1 & f_2 \end{pmatrix}; \quad \mu_2 = \begin{pmatrix} i_1 & i_2 \\ f_2 & f_1 \end{pmatrix}$$

Dans cet exemple, l'application du critère de Pareto permet de conclure que l'appariement μ_1 est plus efficace que l'appariement μ_2 . En effet, le premier permet à chaque candidat d'obtenir son vœu préféré alors que dans le second appariement, les deux candidats sont affectés à la formation qu'ils ont classée en second vœu.

En vertu de ce critère, est considéré comme efficace un appariement qui n'est dominé par aucun autre appariement au sens de Pareto. Autrement dit, un appariement est efficace s'il n'est pas possible d'améliorer la satisfaction d'un candidat sans diminuer celle d'au moins un autre candidat.

2) *L'équité (respect des priorités)*. Le second critère normatif proposé pour évaluer la performance des algorithmes d'affectation est leur capacité à respecter les priorités des candidats. La notion d'équité utilisée ici est précisément définie : un appariement est considéré comme équitable si aucun candidat i ne s'est vu refuser l'admission dans une formation, notée f , qu'il préfère à sa formation d'affectation, notée f' , alors qu'il bénéficie d'une priorité plus élevée qu'un autre candidat i' admis dans la formation f' . Dans le cas contraire, le candidat est considéré comme ayant de la « jalousie légitime » (*justified envy*) vis-à-vis du candidat i' , dans la mesure où il pourrait légitimement contester le bien-fondé de son affectation au motif qu'elle ne respecte pas les priorités. L'exemple suivant illustre une violation de cette propriété :

Vœux d'affectation	Priorités des candidats	Capacité des formations	Appariement
$i_1 : (f_1, f_2)$ $i_2 : (f_1, f_2)$	$f_1 : (i_1, i_2)$ $f_2 : (i_1, i_2)$	$q_1 = 1$ $q_2 = 1$	$\mu = \begin{pmatrix} i_1 & i_2 \\ f_2 & f_1 \end{pmatrix}$

Dans cet exemple, le candidat i_1 est dans une situation de « jalousie légitime » vis-à-vis du candidat i_2 pour la formation f_1 . En effet, i_1 a été affecté sur son deuxième vœu (f_2) alors qu'il bénéficie d'une priorité plus élevée que le candidat i_2 qui a été admis dans la formation f_1 .

Dans le contexte des procédures de choix scolaire, les notions d'équité et d'élimination de la jalousie légitime correspondent à la propriété plus générale de *stabilité*, qui est le concept d'équilibre central de la théorie de l'appariement. Dans un marché composé de deux types d'agents, un appariement est défini comme stable si 1) aucun agent apparié à un agent de l'autre type ne préférerait rester seul et 2) si on ne peut trouver deux agents de type opposé qui préféreraient être appariés ensemble plutôt que de demeurer dans leur situation actuelle. Appliqué au marché scolaire, où les deux types d'agents sont les candidats et les formations, ce concept est équivalent à celui d'élimination de la jalousie légitime⁴.

Le respect des priorités est généralement considéré comme le critère le plus important par les autorités publiques en charge de la régulation des choix scolaires. En effet, toute violation des priorités est susceptible d'entraîner des recours en justice de la part de candidats qui s'estimeraient injustement lésés par l'algorithme d'affectation.

3) *La non-manipulabilité*. Le troisième critère normatif mis en avant par la théorie du *matching* est celui du caractère non manipulable (*strategy-proofness*) de l'algorithme d'affectation. Un algorithme est non manipulable si la meilleure stratégie pour un candidat consiste toujours à classer ses vœux dans l'ordre réel de ses préférences, quels que soient les vœux soumis par les autres candidats.

Un algorithme vérifiant cette propriété récompense la sincérité des vœux, dans la mesure où il garantit aux candidats qu'ils ne pourront obtenir une meilleure affectation en soumettant une liste de vœux qui ne reflète pas leurs préférences réelles. Cette propriété est considérée comme désirable pour deux raisons. D'une part, elle simplifie grandement le processus de décision : il suffit aux candidats de classer les formations en fonction de leurs seules préférences, sans tenir compte de leurs chances d'admission. D'autre part, la non-manipulabilité de l'algorithme d'affectation permet d'éviter de pénaliser les candidats « sincères » au profit des candidats « stratégiques », en excluant qu'un candidat puisse être avantagé par la manipulation de l'ordre de ses vœux.

4. Cette équivalence est discutée plus en détail dans la section *infra* consacrée à l'algorithme d'acceptation différée de Gale et Shapley.

1.3 L’algorithme idéal existe-t-il ?

Idéalement, tout algorithme d’affectation devrait vérifier les trois propriétés énoncées précédemment. Malheureusement, la littérature théorique a démontré qu’aucun algorithme ne peut satisfaire simultanément les trois critères d’efficacité, d’équité et de non-manipulabilité (Roth, 1982; Kesten, 2010).

Ce résultat d’impossibilité doit cependant être tempéré par un autre résultat, qui ouvre des perspectives plus encourageantes : bien que les trois propriétés précédentes ne puissent être simultanément satisfaites, des compromis optimaux sont possibles. Autrement dit, on peut satisfaire deux des trois critères en faisant « au mieux » sur le troisième. La littérature théorique a identifié deux algorithmes qui offrent de bons compromis : l’algorithme d’acceptation différée (*deferred acceptance*) de Gale et Shapley (1962) et l’algorithme des cycles d’échanges optimaux (*top trading cycles*) de Shapley et Scarf (1974). Le premier est non manipulable et respecte les priorités des candidats, mais ne produit pas nécessairement un appariement efficace. Le second algorithme est également non manipulable mais, à l’inverse du premier, il produit un appariement efficace au prix d’une possible violation des priorités des candidats.

Avant de présenter le fonctionnement et les propriétés de ces deux algorithmes, il est utile de revenir sur les raisons qui ont motivé leur application à la régulation des choix scolaires. En effet, si ces algorithmes ont été conçus il y a près d’un demi-siècle, leur utilisation dans le contexte éducatif est relativement récente. Elle est consécutive aux critiques formulées par les économistes à l’encontre de l’algorithme d’affectation qui demeure le plus utilisé à travers le monde : l’algorithme dit « de Boston » (*Boston mechanism*).

2 Un algorithme déficient : le *Boston mechanism*

Historiquement, la plupart des pays ou des districts scolaires qui mettent en œuvre des procédures de choix régulé se sont appuyés sur un algorithme relativement intuitif pour affecter les élèves aux établissements d’enseignement. Baptisé mécanisme de Boston (*Boston mechanism*) ou « mécanisme d’acceptation immédiate » (*immediate acceptance mechanism*), cet algorithme a été utilisé jusqu’en 2005 pour affecter les élèves aux établissements scolaires de la ville de Boston (d’où son nom). Bien qu’il ait été progressivement remplacé par des algorithmes plus performants, il demeure encore très largement utilisé,

non seulement aux États-Unis (Charlotte, New Haven, Minneapolis), mais également en Chine, en Espagne, ou encore aux Pays-Bas (Biró, 2017).

Le succès de l'algorithme de Boston tient à sa simplicité de mise en œuvre et au caractère *a priori* raisonnable de l'objectif qu'il se donne, à savoir satisfaire au maximum les premiers vœux exprimés par les candidats. L'une des principales contributions de la théorie de l'appariement est d'avoir montré que, en dépit des apparences, cet algorithme présente de très mauvaises propriétés.

2.1 Fonctionnement de l'algorithme de Boston

L'algorithme de Boston vise à maximiser la satisfaction des premiers vœux exprimés par les candidats. Muni de la capacité d'accueil des formations et des classements réciproques des candidats et des formations, il procède de manière automatisée par itérations successives :

- *Itération 1* : On considère uniquement le premier vœu de chaque candidat. Chaque formation considère les candidats qui l'ont classée en premier vœu. Elle accepte *de manière définitive* les mieux classés par ordre de priorité dans la limite de sa capacité d'accueil et rejette les autres.
- *Itération 2* : On considère le deuxième vœu des candidats sans affectation à l'issue de la première itération. Les formations qui n'ont pas saturé leur capacité d'accueil considèrent ces candidats. Elles acceptent *de manière définitive* les mieux classés par ordre de priorité dans la limite des places restantes et rejettent les autres.
- *Itération k* : On considère le vœu de rang k des candidats sans affectation à l'issue de l'itération précédente ($k - 1$). Les formations qui n'ont pas saturé leur capacité d'accueil considèrent ces candidats. Elles acceptent *de manière définitive* les mieux classés par ordre de priorité dans la limite des places restantes et rejettent les autres.

L'appariement est obtenu après un nombre fini d'itérations, lorsque tous les candidats ont été affectés à une formation ou que tous les candidats sans affectation ont été rejetés par toutes les formations de leur liste de vœux.

Exemple (tiré de Roth 1982)

On considère trois candidats, i_1 , i_2 et i_3 , et trois formations, f_1 , f_2 et f_3 , pouvant chacune accueillir au plus un candidat. Les vœux d'affectation et les priorités des candidats sont spécifiés comme suit :

Vœux d'affectation	Priorités des candidats	Capacité des formations
$i_1 : (f_2, f_1, f_3)$	$f_1 : (i_1, i_3, i_2)$	$q_1 = 1$
$i_2 : (f_1, f_2, f_3)$	$f_2 : (i_2, i_1, i_3)$	$q_2 = 1$
$i_3 : (f_1, f_2, f_3)$	$f_3 : (i_2, i_1, i_3)$	$q_3 = 1$

Dans cet exemple, l'algorithme de Boston requiert trois itérations et produit l'appariement suivant⁵ :

$$\mu_{BM} = \begin{pmatrix} i_1 & i_2 & i_3 \\ f_2 & f_3 & f_1 \end{pmatrix}$$

2.2 Les propriétés indésirables de l'algorithme de Boston

Contrairement à ce que sa popularité pourrait laisser entendre, l'algorithme de Boston ne vérifie aucune des trois propriétés désirables décrites plus haut : il est manipulable, ne respecte pas les priorités et conduit généralement à un appariement inefficace.

La principale critique formulée à l'encontre de l'algorithme de Boston est qu'il incite les candidats à adopter des comportements stratégiques et pénalise les candidats « sincères », c'est-à-dire ceux qui classent leurs vœux par ordre de préférence réel. Le caractère manipulable de cet algorithme tient au fait qu'il dissuade les candidats de classer en premier vœu une formation où ils ont peu de chances d'être admis. En effet, si une formation est très demandée, le candidat qui la classe en premier vœu mais qui y est refusé se trouve pénalisé pour les formations classées plus loin dans sa liste de vœux – ces formations n'ayant plus de places disponibles après le premier tour – quand bien même le candidat

5. Détail des itérations de l'algorithme : – *Itération 1* : On considère uniquement le 1er vœu de chaque candidat. Le candidat i_1 postule à la formation f_2 et, en l'absence d'autres candidats, y est affecté de manière définitive. Les candidats i_2 et i_3 postulent à la formation f_1 . Puisqu'il n'y a qu'une place et que cette formation donne priorité à i_3 par rapport à i_2 , i_3 est affecté à f_1 de manière définitive tandis que i_2 est refusé. – *Itération 2* : Le seul candidat sans affectation à l'issue de la première itération est i_2 . On considère son 2^e vœu, qui se porte sur la formation f_2 . Cette formation ayant saturé sa capacité d'accueil à l'étape précédente (elle a admis i_1), elle refuse i_2 . – *Itération 3* : le seul candidat sans affectation à l'issue des deux premières itérations est i_2 . On considère son 3^e vœu, qui se porte sur la formation f_3 . Cette formation ayant une place disponible, elle admet i_2 de manière définitive.

disposerait d'une priorité plus élevée que les candidats précédemment admis dans ces formations. Dans ces conditions, il peut être dans l'intérêt des candidats de « manipuler » l'ordre de leurs vœux en sous-classant certaines formations ou en ne candidatant pas aux formations les plus demandées.

Le caractère manipulable de l'algorithme de Boston peut être illustré en considérant les incitations stratégiques auxquels fait face le candidat i_2 dans l'exemple précédent. Si ce candidat préfère réellement la formation f_1 à la formation f_2 et la formation f_2 à la formation f_3 , il a intérêt à « mentir » sur ses préférences pour améliorer son affectation. On peut aisément vérifier qu'en classant la formation f_2 en vœu 1 plutôt qu'en vœu 2, ce candidat y serait affecté par l'algorithme de Boston. Cette manipulation est profitable, dans la mesure où le candidat i_2 préfère strictement la formation f_2 à la formation f_3 , qu'il obtient s'il classe sincèrement ses vœux (voir l'appariement μ_{BM} de l'exemple précédent). Dans le contexte des affectations scolaires, le fait que l'algorithme de Boston puisse lourdement pénaliser les candidats « sincères » constitue une propriété particulièrement indésirable car contraire au principe d'équité de traitement.

Le second défaut de l'algorithme de Boston est qu'il ne respecte pas les priorités des candidats. Dans l'exemple précédent, l'appariement μ_{BM} crée de la « jalousie justifiée » pour le candidat i_2 : il est affecté sur son troisième vœu (f_3) alors qu'il bénéficie d'une priorité plus élevée que le candidat i_1 , qui a été admis dans la formation que i_2 avait classé en deuxième vœu (f_2).

Enfin, la littérature théorique a montré qu'en raison de son caractère manipulable, l'algorithme de Boston conduit généralement à un appariement inefficace, ce qui signifie qu'il est possible de trouver un appariement qui améliore strictement la satisfaction d'au moins un élève sans réduire la satisfaction d'aucun autre. L'inefficacité de l'algorithme de Boston est un phénomène général, qui ne dépend pas en particulier du degré d'information des candidats sur les préférences et les priorités des autres candidats (Erdil et Ergin, 2006)⁶.

6. Une réévaluation partielle de l'algorithme de Boston a cependant été proposée dans la littérature (Miralles, 2009; Abdulkadiroğlu et al., 2011), au nom de l'idée selon laquelle cet algorithme prendrait plus efficacement en compte l'intensité des préférences des élèves que l'algorithme d'acceptation différée (présenté plus loin). En effet, du fait de la prise de risque qu'il induit, l'algorithme de Boston tend à décourager les candidats qui n'ont qu'une préférence modérée pour une formation à la classer en premier vœu, au profit des candidats qui ont une forte préférence pour cette formation.

2.3 Un algorithme « tenace »

À la lumière de la discussion précédente, il peut paraître *a priori* surprenant qu'un algorithme d'affectation aussi peu performant que celui de Boston soit, aujourd'hui encore, le plus répandu à travers le monde. Cette étonnante résistance tient au fait que cet algorithme donne l'illusion de maximiser la satisfaction des préférences des candidats et qu'il n'y a, par conséquent, pas de raisons d'en changer. Or, comme on vient de le voir, l'algorithme de Boston tend en réalité à « gonfler » artificiellement le taux de satisfaction des premiers vœux en incitant les candidats à ne pas classer en haut de leur liste des formations où ils anticipent de faibles chances d'admission.

Pour démontrer empiriquement ce phénomène, il faudrait pouvoir comparer les préférences *exprimées* par les candidats avec leurs préférences *réelles*. Malheureusement, ces dernières ne sont pas directement observables, ce qui a longtemps limité la portée pratique des arguments théoriques avancés à l'encontre de cet algorithme. Afin de contourner cette difficulté, plusieurs études se sont appuyées sur des méthodes indirectes pour mettre en évidence les comportements stratégiques induits par l'algorithme de Boston (He, 2017; Agarwal et Somaini, 2018; Calsamiglia, Fu et Güell, 2020). Appliquées aux données de la procédure d'affectation utilisée à Boston, ces méthodes ont notamment permis de démontrer que les familles afro-américaines étaient fortement pénalisées par cet algorithme d'affectation en raison de leur défaut d'information sur les stratégies optimales de classement des vœux (Abdulkadiroğlu et al., 2006). Cet argument a joué un rôle décisif pour convaincre les autorités éducatives de la ville de Boston d'abandonner définitivement ce mécanisme au profit d'un algorithme beaucoup plus performant : l'algorithme d'acceptation différée.

3 L'algorithme d'acceptation différée ou algorithme de Gale et Shapley

Les critiques adressées par les économistes à l'encontre de l'algorithme de Boston ont eu un retentissement important aux États-Unis et ont conduit plusieurs districts scolaires à y renoncer au profit un algorithme présentant de meilleures propriétés : l'algorithme

d'acceptation différée de Gale et Shapley⁷. Cet algorithme a été adopté par la ville de New York en 2003, suivie par Boston en 2005, Chicago en 2009, Denver en 2012, La Nouvelle Orléans et Washington DC en 2013 (Pathak, 2019). En Angleterre, l'utilisation de l'algorithme de Boston (connu localement sous le nom de « First Preference First ») pour réguler les choix scolaires a été interdite par la loi en 2007 et a été remplacée dans l'ensemble des districts scolaires par l'algorithme d'acceptation différée (Pathak et Sönmez, 2013). Si l'algorithme de Boston reste utilisé dans plusieurs pays européens (Allemagne, Espagne), il a progressivement laissé la place à l'algorithme de Gale et Shapley, notamment en France (pour les admissions au lycée), en Finlande, en Hongrie, aux Pays-Bas (Amsterdam) et en Suède.

3.1 Origine : la théorie des « mariages stables »

Bien que l'algorithme d'acceptation différée ait été conçu il y a plus d'un demi-siècle, la prise de conscience de son utilité pour la régulation des choix scolaires ne date que du début des années 2000.

L'algorithme d'acceptation trouve son origine dans un court mais lumineux article publié en 1962 par les mathématiciens américains David Gale et Lloyd Shapley. Les auteurs considèrent le problème suivant : étant donné n hommes et n femmes ayant chacun un ordre de préférence strict pour les individus du sexe opposé, est-il possible de former des mariages « stables » ? Il y a instabilité dès lors qu'il existe au moins un homme et une femme qui préféreraient former un nouveau couple plutôt que de rester chacun avec son partenaire respectif (ou de rester célibataire). Dans cette configuration, Gale et Shapley ont démontré qu'il existe toujours au moins un appariement entre hommes et femmes aboutissant à des mariages stables, et ont proposé un algorithme qui permet de construire un appariement vérifiant cette propriété. Cet algorithme, baptisé *deferred-acceptance mechanism* (souvent abrégé en DA), se décline en deux versions, suivant que ce sont les hommes qui font des propositions aux femmes ou le contraire.

Dans la version où les propositions sont faites par les hommes (*men-proposing DA*), l'algorithme fonctionne de la manière suivante⁸ :

7. Pour une présentation érudite de cet algorithme et de ses applications, voir Roth (2008a).

8. Dans la version décrite ici, on suppose que tous les individus préfèrent toujours se mettre en couple plutôt que de rester célibataires. L'algorithme peut être aisément modifié pour prendre en compte la possibilité que des hommes et des femmes préfèrent rester célibataires plutôt que de se mettre en couple avec certains membres du sexe opposé.

- *Itération 1* : Chaque homme demande en mariage la femme qu'il préfère entre toutes. Chaque femme choisit l'homme qu'elle préfère parmi ceux qui lui ont fait une proposition et est *provisoirement* appariée avec cet homme. Les autres propositions de mariage sont rejetées.
- *Itération k* : Chaque homme dont la proposition a été rejetée à l'itération précédente fait une proposition à la femme qu'il préfère parmi celles qui ne l'ont pas précédemment rejeté. Chaque femme choisit l'homme qu'elle préfère parmi ses prétendants, en considérant *simultanément* les nouvelles propositions et le partenaire avec lequel elle a été éventuellement appariée à l'issue des itérations précédentes, et est *provisoirement* appariée avec le partenaire choisi. Les autres propositions sont rejetées.

Ces itérations sont répétées jusqu'à ce que tous les hommes et toutes les femmes soient appariés. Lorsque c'est le cas, tous les appariements provisoires deviennent *définitifs*.

Cet algorithme produit un appariement unique après un nombre fini d'itérations. Gale et Shapley démontrent que cet appariement « favorise » les hommes dans la mesure où, parmi l'ensemble des appariements qui produisent des couples stables, il s'agit de celui qui est préféré par tous les hommes (*men-optimal*).

Dans la seconde version de l'algorithme, les rôles des hommes et des femmes sont inversés, les propositions de mariage étant faites par les femmes (*women-proposing DA*). Cette version de l'algorithme produit également des couples stables mais l'appariement obtenu favorise cette fois les femmes (*women-optimal*).

Un résultat important démontré par Roth (1982) concerne les manipulations stratégiques auxquelles peut donner lieu l'algorithme de Gale et Shapley. Dans le modèle du mariage stable, Roth démontre que lorsque les propositions sont faites par les hommes, ces derniers n'ont pas intérêt à manipuler l'ordre dans lequel ces propositions sont faites. La meilleure stratégie consiste pour eux à classer les femmes par ordre de préférence et à faire leurs propositions dans cet ordre. En revanche, dans cette version de l'algorithme, il peut exister des manipulations profitables pour des femmes, qui consistent par exemple pour une femme à rejeter la proposition d'un homme au profit d'une proposition moins souhaitée, de manière à influencer l'ordre des propositions qui lui seront faites ultérieurement. De manière symétrique, la version de l'algorithme dans laquelle les propositions sont faites par les femmes n'est pas manipulable par les femmes, mais peut l'être par les

hommes.

3.2 Des mariages stables à la régulation des choix scolaires

Si le problème du mariage stable constitue une étape décisive dans le développement de la théorie de l'appariement, la portée pratique de l'algorithme de Gale et Shapley est restée limitée jusqu'au début des années 2000.

Dès l'origine, cependant, Gale et Shapley avaient imaginé une application possible de leur algorithme à la régulation des inscriptions universitaires. Dans leur article fondateur, ces auteurs ont montré que, par analogie avec le problème des mariages stables, les inscriptions universitaires peuvent être modélisées comme un problème d'appariement impliquant deux groupes d'agents ayant des préférences vis-à-vis des agents de l'autre groupe : les étudiants, d'une part, qui ont des préférences vis-à-vis des universités ; les universités, d'autre part, qui ont des préférences vis-à-vis des étudiants. La seule différence conceptuelle entre le problème des inscriptions universitaires et celui des mariages stables est que l'appariement universitaire ne concerne pas des couples (*one-to-one matching*) mais est multivoque (*many-to-one matching*) car plusieurs étudiants peuvent être appariés avec une même université. Malgré cette différence, Gale et Shapley ont montré que leurs principaux résultats théoriques restaient valables dans le problème des inscriptions universitaires : un appariement « stable » entre étudiants et universités existe toujours et l'algorithme d'acceptation différée (dans une version légèrement modifiée) permet d'obtenir un tel appariement. Ici, la stabilité de l'appariement signifie qu'il n'existe aucun étudiant qui préférerait s'inscrire dans une université autre que celle où il a été admis et qui serait préféré par cette université aux étudiants qu'elle a elle-même admis.

Comme dans son application au marché matrimonial, l'algorithme d'acceptation différée appliqué aux inscriptions universitaires se décline en deux versions, qui l'une et l'autre produisent des appariements stables : la version dans laquelle les propositions sont faites par les étudiants (*student-proposing DA*) favorise les étudiants alors que la version dans laquelle les propositions sont faites par les universités (*college-proposing DA*) favorise ces dernières.

La difficulté soulevée par l'application de l'algorithme de Gale et Shapley aux inscriptions universitaires tient au caractère manipulable de cet algorithme, quelle que soit la version utilisée (Roth, 1985). Lorsque les propositions sont faites par les étudiants, l'algo-

rithme n'est pas manipulable par ces derniers (ils ont toujours intérêt à classer sincèrement leurs vœux d'inscription) mais peut être manipulé par les universités (elles peuvent avoir intérêt à ne pas classer les candidats par ordre de préférence réel). Le problème se pose avec plus d'acuité encore lorsque les propositions sont faites par les universités : dans ce cas, l'algorithme est manipulable tant par les étudiants que par les universités (ce qui n'est pas le cas dans le problème du mariage stable, où les manipulations profitables ne peuvent exister que d'un côté du « marché »).

La réévaluation de l'algorithme de Gale et Shapley comme outil d'affectation scolaire s'inscrit dans un contexte particulier, qui est celui du développement des procédures automatisées d'affectation des élèves, lui-même rendu possible par les progrès réalisés dans le domaine des technologies de l'information. Cette révolution technique a été accompagnée par plusieurs percées théoriques majeures qui, au tournant des années 2000, ont démontré la supériorité de l'algorithme d'acceptation différée sur les algorithmes les plus communément utilisés pour affecter les élèves aux établissements scolaires, en particulier l'algorithme de Boston.

Dans un article influent, Abdulkadiroğlu et Sönmez (2003) ont montré que la régulation des choix scolaires pouvait être modélisée comme une version du problème des inscriptions universitaires analysé par Gale et Shapley (1962) et Roth (1985), avec une différence importante : contrairement aux universités, les écoles publiques n'ont pas à proprement parler de « préférences » vis-à-vis des élèves mais appliquent mécaniquement des règles de priorités qui sont déterminées en amont et qui échappent par conséquent à leur contrôle. La manière dont les écoles classent les candidats ne peut donc pas donner lieu à des manipulations, ce qui résout les problèmes stratégiques posés par l'algorithme de Gale et Shapley lorsque la version utilisée est celle où les propositions sont faites par les élèves.

3.3 Fonctionnement de l'algorithme d'acceptation différée

Comme l'algorithme de Boston, l'algorithme d'acceptation différée détermine un appariement entre élèves et écoles à partir des listes ordonnées de vœux des élèves, de leurs priorités pour chaque école et de la capacité d'accueil de ces dernières. Dans la version où les propositions sont faites par les élèves (*student-proposing DA*), l'algorithme d'acceptation différée fonctionne comme suit :

— *Itération 1* : Chaque école considère les élèves qui l'ont classée en 1^{er} vœu. Elle

accepte *provisoirement* les mieux classés par ordre de priorité dans la limite de sa capacité d'accueil et rejette les autres.

- *Itération k* : Chaque élève dont la candidature a été rejetée à l'étape précédente candidate à l'école qu'il préfère parmi celles qui ne l'ont pas encore rejeté. Chaque école considère *conjointement* les élèves précédemment admis et les nouveaux candidats. Les élèves les mieux classés dans la limite de la capacité d'accueil de l'école sont *provisoirement* admis et les autres sont rejetés.

L'appariement est obtenu après un nombre fini d'itérations, lorsque plus aucun élève n'est rejeté ou que tous les élèves sans affectation ont épuisé leur liste de vœux⁹.

La différence fondamentale qui distingue l'algorithme d'acceptation différée de celui de Boston est qu'à chaque itération, les élèves ne sont que *provisoirement* affectés aux formations. Ainsi, un élève qui a été accepté par une formation lors d'une itération de l'algorithme peut, lors d'une itération ultérieure, être remplacé par un élève disposant d'une priorité plus élevée, ce qui n'est pas permis par l'algorithme de Boston (d'où le nom de *immediate acceptance mechanism* qui lui est parfois donné pour souligner ce contraste).

Le fonctionnement de l'algorithme d'acceptation différée peut être illustré en partant de l'exemple utilisé plus haut pour décrire l'algorithme de Boston :

Vœux d'affectation	Priorités des candidats	Capacité des formations
$i_1 : (f_2, f_1, f_3)$	$f_1 : (i_1, i_3, i_2)$	$q_1 = 1$
$i_2 : (f_1, f_2, f_3)$	$f_2 : (i_2, i_1, i_3)$	$q_2 = 1$
$i_3 : (f_1, f_2, f_3)$	$f_3 : (i_2, i_1, i_3)$	$q_3 = 1$

Dans cet exemple, l'algorithme d'acceptation différée (dans la version *student-proposing*)

9. Dans la version de l'algorithme où les propositions sont faites par les écoles (*student-proposing deferred acceptance*), les rôles des élèves et des écoles sont inversés. – *Itération 1* : chaque école f fait des propositions d'admission aux élèves les mieux classés par ordre de priorité dans la limite de sa capacité d'accueil, notée q_f . Chaque élève ayant reçu une ou plusieurs propositions accepte *provisoirement* celle qui émane de l'école qu'il préfère et rejette les autres propositions. – *Itération k* : chaque école f qui, à l'issue de l'itération précédente a provisoirement admis $q < q_f$ élèves, fait $q_f - q$ propositions d'admission aux élèves les mieux classés parmi ceux auxquels elle n'a pas fait encore fait de proposition. Chaque élève compare les propositions nouvellement reçues avec la proposition provisoirement retenue à l'issue des itérations précédentes. Il accepte *provisoirement* la proposition préférée et rejette les autres. L'appariement est obtenu après un nombre fini d'itérations, lorsque plus aucune des propositions faites par les écoles n'est rejetée ou lorsque les écoles ont épuisé leur liste de candidats. Les affectations provisoires deviennent alors définitives.

nécessite cinq itérations et produit l'appariement suivant¹⁰ :

$$\mu_{DA} = \begin{pmatrix} i_1 & i_2 & i_3 \\ f_1 & f_2 & f_3 \end{pmatrix}$$

3.4 Propriétés de l'algorithme d'acceptation différée

L'exemple précédent permet d'illustrer les bonnes propriétés de l'algorithme d'acceptation différée.

La première de ces propriétés est que, contrairement à l'algorithme de Boston, l'algorithme d'acceptation différée n'est pas manipulable (Dubins et Freedman, 1981; Roth, 1982) : lorsque les élèves peuvent classer l'ensemble des formations, la meilleure stratégie consiste à les classer par ordre de préférence réel. Dans l'exemple, aucun des élèves ne peut espérer obtenir une meilleure affectation en ne classant pas sincèrement ses vœux, et ce, *quelles que soient* les stratégies adoptées par les autres élèves (qu'elles soient sincères ou non). Cette propriété permet de simplifier considérablement le processus de décision, en donnant aux candidats la possibilité de classer leurs vœux sans avoir à tenir compte de leurs chances d'admission et en leur offrant la garantie, s'ils ne sont pas admis sur un vœu, qu'ils ne seront pas pénalisés pour leurs vœux suivants.

La seconde propriété désirable de l'algorithme d'acceptation différée est qu'il respecte les priorités et élimine de ce fait toute « jalousie légitime ». Cette propriété est la transposition au marché scolaire de la propriété de stabilité énoncée par Gale et Shapley à propos du marché matrimonial (Balinski et Sönmez, 1999). De la même manière que les mariages sont « stables » s'il n'existe pas de relation extra-conjugale menaçant les couples formés, l'appariement entre élèves et écoles est « stable » si aucun élève ne peut contester

10. Détails des itérations : – *Itération 1* : On considère uniquement le 1^{er} vœu de chaque élève. L'élève i_1 candidate à l'école f_2 et, en l'absence d'autres candidats, y est provisoirement affecté. Les élèves i_2 et i_3 candidatent à l'école f_1 . Puisqu'il n'y a qu'une place et que cette école donne priorité à i_3 par rapport à i_2 , i_3 est provisoirement affecté à f_1 tandis que i_2 est refusé. – *Itération 2* : L'élève i_2 (rejeté à l'étape précédente) candidate à l'école qu'il a classée en second vœu, f_2 . L'élève i_2 ayant une priorité plus élevée dans cette école que l'élève admis au tour précédent (i_1), l'école f_2 admet provisoirement i_2 et rejette i_1 . – *Itération 3* : L'élève i_1 (rejeté à l'étape précédente) candidate à l'école qu'il a classée en second vœu, f_1 . L'élève i_1 ayant une priorité plus élevée dans cette école que l'élève admis au tour précédent (i_3), l'école f_1 admet provisoirement i_1 et rejette i_3 . – *Itération 4* : L'élève i_3 (rejeté à l'étape précédente) candidate à l'école qu'il a classée en second vœu, f_2 . L'élève i_3 ayant une priorité plus faible dans cette école que l'élève qui y a été provisoirement admis (i_2), l'école f_2 rejette i_3 . – *Itération 5* : L'élève i_3 (rejeté à l'étape précédente) candidate à l'école qu'il a classée en troisième vœu, f_3 . En l'absence d'autres candidats, i_3 est provisoirement affecté à l'école f_3 . Aucun élève n'étant plus rejeté par une école, l'affectation est définitive.

son affectation au motif qu'il aurait été refusé par une formation au profit d'un élève disposant d'une priorité plus faible. Cette propriété implique que l'appariement obtenu par l'algorithme d'acceptation affecte chaque élève à la formation la mieux classée dans sa liste de vœux parmi l'ensemble des formations où il aurait pu être admis.

Le respect des priorités par l'algorithme d'acceptation différée a cependant un coût en matière d'efficacité. En effet, certains élèves pourraient avoir intérêt à échanger leurs affectations de manière mutuellement avantageuse. La raison pour laquelle ces échanges ne sont pas autorisés par l'algorithme est qu'ils entraîneraient une violation des priorités vis-à-vis d'autres élèves. Dans l'exemple précédent, les élèves i_1 et i_2 gagneraient à échanger leurs places : i_1 préfère en effet l'école f_2 à l'école f_1 , tandis que i_2 préfère l'école f_1 à f_2 . Cet échange mutuellement bénéfique n'est pas permis par l'algorithme car il violerait les priorités de l'élève i_3 : cet élève serait refusé par l'école f_1 alors qu'il est prioritaire par rapport à l'élève f_2 pour cette école.

En vertu du « triangle d'impossibilité » décrit plus haut, les trois propriétés d'efficacité, d'équité (stabilité) et de non-manipulabilité ne peuvent être simultanément vérifiées par aucun algorithme d'affectation scolaire. Sous cette contrainte, l'algorithme d'acceptation différée offre un compromis optimal : il s'agit du seul algorithme non manipulable qui produit un appariement stable (Alcalde et Barberà, 1994) et, dans la classe des algorithmes qui vérifient la propriété de stabilité, c'est celui qui offre la satisfaction maximale des préférences des élèves. L'appariement qu'il produit est en effet préféré par chaque élève à tous les appariements qui ne violent pas les priorités. Cet algorithme peut donc être qualifié d'« efficace sous contrainte » (*constrained efficient*).

3.5 Les conséquences indésirables de la limitation du nombre de vœux

Une question d'importance pratique considérable concerne les conséquences de la limitation du nombre de vœux sur les propriétés de l'algorithme de Gale et Shapley. Les bonnes propriétés de cet algorithme supposent en effet que les candidats soient autorisés à classer par ordre de préférence l'ensemble des formations où ils souhaiteraient être admis. Or la plupart des procédures d'affectation centralisées utilisées à travers le monde, notamment en France, limitent le nombre de vœux que les candidats peuvent soumettre. Quelles sont les conséquences de cette restriction ?

La limitation du nombre de vœux est une pratique que la littérature théorique recommande de proscrire car elle compromet les bonnes propriétés de l’algorithme d’acceptation différée. Une première conséquence dommageable de cette limitation est qu’elle réintroduit des considérations stratégiques que la version non contrainte de l’algorithme permettait justement d’éliminer. En effet, si les candidats ne sont autorisés qu’à classer K formations parmi les N qui leur sont proposées, il n’est pas nécessairement dans leur intérêt de classer par ordre de préférence les K formations qu’ils préfèrent, dès lors qu’une telle stratégie pourrait les conduire à se retrouver sans affectation. Pour se prémunir contre ce risque, le choix des formations à inclure dans la liste de vœux doit tenir compte des chances objectives d’admission, même s’il reste dans l’intérêt des candidats de classer les formations *sélectionnées* par ordre de préférence réel (Haeringer et Klijn, 2009). Le corollaire de ces considérations stratégiques est que, lorsque le nombre de vœux est limité, les candidats peuvent être conduits à omettre de leur liste des formations dans lesquelles ils auraient en réalité pu être admis et, par conséquent, à regretter *a posteriori* de ne pas y avoir candidaté. La possibilité que des candidats éprouvent ce type de regret constitue une violation directe de la seconde propriété désirable de l’algorithme de Gale et Shapley, qui est de produire un appariement stable entre les candidats et les formations, c’est-à-dire un appariement tel que chaque candidat est affecté à la formation qu’il préfère parmi l’ensemble des formations où il aurait pu être admis. Lorsque le nombre de vœux est limité, cette propriété n’est plus assurée¹¹.

Les conséquences de la limitation du nombre de vœux sur le bon fonctionnement des algorithmes d’appariement sont suffisamment sérieuses pour recommander, sinon l’abandon de cette pratique, du moins son assouplissement. Permettre aux candidats de soumettre un nombre de vœux plus élevé que ce qui est généralement permis par les procédures d’affectation se justifie d’autant plus qu’il n’existe aucun obstacle technique au relâchement de cette contrainte – le temps d’exécution de l’algorithme de Gale et Shapley ne s’en trouve allongé que de quelques secondes.

11. Sous certaines conditions, la stabilité de l’appariement obtenu par l’algorithme d’acceptation différée peut néanmoins être satisfaite de manière approximative, même lorsque le nombre de vœux est limité. C’est notamment le cas lorsque la publication des « barres d’admission » observées au cours des années précédentes permet aux candidats d’évaluer avec un degré de précision suffisant leurs chances d’admission dans les différentes formations (Fack, Grenet et He, 2019).

4 L’algorithme des cycles d’échanges optimaux

L’algorithme d’acceptation différée n’est pas l’unique compromis envisageable pour réguler les inscriptions scolaires. Abdulkadiroğlu et Sönmez (2003) ont en effet montré qu’au prix d’adaptations mineures, un autre algorithme proposé dans la littérature – l’algorithme des cycles d’échanges optimaux (*top trading cycles* ou TTC) – permet de réaliser un compromis différent entre les trois propriétés désirables énoncées plus haut, en privilégiant les critères d’efficacité et de non-manipulabilité par rapport au critère d’équité (ou stabilité).

4.1 Fonctionnement de l’algorithme des cycles d’échanges optimaux

Conçu par David Gale et publié par Shapley et Scarf (1974) pour analyser le problème des échanges de biens indivisibles (comme des maisons occupées par des propriétaires), l’algorithme des cycles d’échanges optimaux a été adapté au problème de la régulation des choix scolaires par Abdulkadiroğlu et Sönmez (2003).

Muni des informations habituelles (capacité des formations, classement des formations par les candidats et classement des candidats par les formations), cet algorithme fonctionne comme suit :

- *Itération 0* : Un « compteur » est mis en place pour conserver la trace du nombre de places disponibles dans chaque école. Ce compteur prend comme valeurs initiales les capacités d’accueil des différentes écoles.
- *Itération 1* : On construit un graphe orienté en considérant les élèves et les écoles comme les « sommets » du graphe et en construisant les « arrêtes » (c’est-à-dire des flèches pointant d’un élève vers une école ou d’une école vers un élève) comme suit : chaque élève « pointe » vers l’école qu’il a classée en 1^{er} vœu et chaque école « pointe » vers l’élève qui bénéficie de la priorité la plus élevée parmi ceux qui ont candidaté à cette école. Le graphe orienté ainsi construit possède au moins un « cycle », c’est-à-dire une succession d’arrêtes (flèches) qui, partant d’un élève donné, permettent de revenir vers cet élève. Chacun des élèves appartenant à un cycle de ce type est définitivement affecté à l’école vers laquelle il pointe et est retiré du graphe.

Pour chacune des écoles concernées, le compteur des places restantes est diminué de un. Si la capacité d'une école est saturée, celle-ci est retirée du graphe.

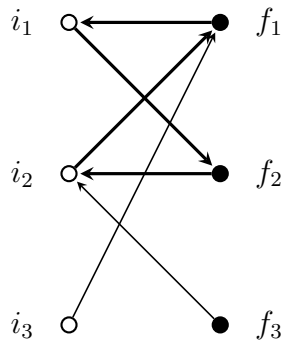
- *Itération k* : les opérations précédentes sont répétées pour les élèves non affectés et pour les écoles dont la capacité n'est pas saturée.

La procédure se termine après un nombre fini d'itérations, lorsqu'il n'existe plus de cycles dans le graphe orienté.

Le fonctionnement de l'algorithme des cycles d'échanges optimaux peut être illustré en partant de l'exemple précédemment utilisé pour décrire le fonctionnement des algorithmes de Boston et de Gale-Shapley.

Vœux d'affectation	Priorités des candidats	Capacité des formations
$i_1 : (f_2, f_1, f_3)$	$f_1 : (i_1, i_3, i_2)$	$q_1 = 1$
$i_2 : (f_1, f_2, f_3)$	$f_2 : (i_2, i_1, i_3)$	$q_2 = 1$
$i_3 : (f_1, f_2, f_3)$	$f_3 : (i_2, i_1, i_3)$	$q_3 = 1$

À l'étape 1, le graphe orienté construit à partir de cet exemple est :



Dans ce graphe, chaque élève pointe vers son école préférée et chaque école pointe vers l'élève ayant la priorité la plus élevée. Chaque école pouvant accueillir au plus un élève, le compteur des places restantes est initialement fixé à 1 pour chaque école.

Ce graphe ne comporte qu'un cycle (indiqué par les flèches en gras) : $i_1 - f_2 - i_2 - f_1 - i_1$. Les deux élèves appartenant à ce cycle sont affectés aux écoles vers lesquelles ils « pointent » respectivement : i_1 est affecté à f_2 et i_2 est affecté à f_1 . Ces deux élèves sont retirés du graphe, ainsi que les deux écoles f_1 et f_2 qui n'ont plus de places restantes.

Le graphe résultant de ces modifications est simplement :



Il ne reste plus qu'un cycle, $i_3 - f_3 - i_3$, qui conduit à affecter l'élève i_3 à l'école f_3 .

Dans cet exemple, l'algorithme des cycles d'échanges optimaux ne nécessite donc que deux itérations et produit l'appariement suivant :

$$\mu_{TTC} = \begin{pmatrix} i_1 & i_2 & i_3 \\ f_2 & f_1 & f_3 \end{pmatrix}$$

4.2 Propriétés de l'algorithme des cycles d'échanges optimaux

Comme l'algorithme d'acceptation différée, l'algorithme des cycles d'échanges optimaux est non manipulable : quels que soient les vœux soumis par les participants, la meilleure stratégie consiste toujours pour un élève à classer ses vœux par ordre de préférence.

Les deux algorithmes se distinguent cependant par le compromis qu'ils proposent entre les critères normatifs d'efficacité et d'équité. Alors que l'algorithme d'acceptation différée privilégie l'équité au prix d'une perte potentielle d'efficacité, l'algorithme des cycles d'échanges fait le choix inverse. Cet algorithme autorise en effet deux élèves à « échanger » leurs priorités pour différentes écoles si cela leur permet d'obtenir une affectation mutuellement avantageuse. Dans l'exemple précédent, la première itération de l'algorithme conduit les élèves i_1 et i_2 à échanger leurs priorités pour les écoles f_1 et f_2 de manière à ce que chacun obtienne son école préférée (f_2 pour i_1 et f_1 pour i_2), alors que l'algorithme d'acceptation différée affecte ces deux élèves sur leur deuxième vœu. L'appariement obtenu par l'algorithme des cycles d'échanges optimaux domine, au sens du critère de Pareto, celui obtenu par l'algorithme d'acceptation différée : les élèves i_1 et i_2 obtiennent en effet une satisfaction strictement plus grande sans que cela n'altère la situation de l'élève i_3 (qui dans les deux cas est affecté à l'école f_3).

Ce gain d'efficacité n'est possible que parce que l'algorithme des cycles d'échanges optimaux tolère une violation des priorités. Dans l'exemple précédent, l'appariement μ_{TTC} crée une situation de « jalousie légitime » pour l'élève i_3 . Bien que sa propre affectation soit inchangée par rapport à celle qu'il aurait obtenue par l'algorithme d'acceptation différé, cet élève voit désormais sa priorité violée pour l'école f_1 , l'élève i_2 y ayant été affecté alors que sa priorité est inférieure à celle de l'élève i_3 .

4.3 Une portée pratique limitée

Lorsque les autorités éducatives de la ville de Boston décidèrent de renoncer définitivement à l’algorithme de Boston, il leur fallut choisir entre les deux algorithmes proposés par les économistes dont l’avis avait été sollicité : acceptation différée ou cycles d’échanges optimaux (Abdulkadiroğlu et al., 2006). Le choix des autorités se porta sans trop d’hésitation sur le premier, pour deux raisons essentielles. D’une part, le caractère indéniablement plus complexe de l’algorithme des cycles d’échanges optimaux était perçu comme plus à même de perpétuer chez les parents l’idée selon laquelle il était dans leur intérêt de manipuler l’ordre de leurs vœux. D’autre part, les violations des priorités induites par cet algorithme présentaient le risque de favoriser les recours en justice formés à l’encontre des décisions d’affectation.

Cette « aversion » à l’encontre de l’algorithme des cycles d’échanges optimaux (Pathak, 2019) s’est manifestée à chaque fois que celui-ci a été proposé pour réguler les affectations scolaires. À ce jour, il n’a été utilisé que par un seul district scolaire aux États-Unis, le Recovery School District de la Nouvelle-Orléans. Expérimenté en 2012, il fut abandonné dès l’année suivante au profit de l’algorithme d’acceptation différée (Abdulkadiroğlu et al., 2017). Les raisons de cet abandon tiennent aux difficultés rencontrées pour expliquer le fonctionnement de l’algorithme aux parents d’élèves et, plus encore, pour leur faire comprendre que les violations des priorités d’un élève (tel que i_3 dans l’exemple précédent) n’étaient tolérées que pour permettre à d’autres élèves d’obtenir une plus grande satisfaction sans que cela ne détériore la situation du premier élève par rapport à la situation où cet échange n’aurait pas été autorisé. Ces considérations pratiques expliquent que, en dépit de son attrait théorique, l’algorithme des cycles d’échanges optimaux n’ait pas connu le même succès que l’algorithme d’acceptation différée dans la régulation des inscriptions scolaires ¹².

5 Algorithmes d’affectation et objectifs de diversité

Dans l’histoire récente de la théorie de l’appariement, la question de la compatibilité des algorithmes d’affectation avec l’objectif de promotion de la diversité dans le recru-

12. Cet algorithme a en revanche eu une portée pratique considérable dans d’autres applications, telles que l’organisation centralisée des dons croisés de reins (Roth, Sönmez et Ünver 2004, 2005 ; Combe et al. 2022).

tement des établissements scolaires et des formations d'enseignement supérieur a donné lieu à de nombreux développements. Ces derniers ont permis de mieux cerner les forces et les faiblesses des outils mobilisables pour mettre en œuvre cet objectif dans le cadre des procédures d'affectation centralisées.

5.1 La diversité du recrutement : un objectif aux déclinaisons multiples

L'objectif d'accroître la diversité de recrutement des établissements scolaires par l'intermédiaire des procédures d'affectation se décline de manière différente selon les pays. Aux États-Unis, les politiques de déségrégation scolaire mises en œuvre après l'arrêt de la Cour suprême *Brown v. Board of Education* de 1954 ont généralement pris la forme de quotas raciaux imposant aux écoles de réserver certaines places exclusivement aux élèves afro-américains. Déclaré inconstitutionnel par la Cour suprême dans son arrêt *Parents Involved in Community School District No. 1* de 2007, le recours aux quotas raciaux dans les affectations scolaires fut progressivement abandonné aux États-Unis au profit d'autres politiques de discrimination positive ne faisant pas directement intervenir le critère racial, telles que la mise en place de quotas en faveur d'élèves issus de quartiers défavorisés (Ellison et Pathak, 2021).

En Europe, l'objectif de diversité dans le recrutement des écoles est davantage appréhendé sous l'angle de la mixité sociale ou scolaire que de la diversité raciale ou ethnique. Dans certains pays, comme la Belgique, les procédures de choix scolaires utilisées dans l'enseignement primaire et secondaire font intervenir des quotas réservés aux élèves de milieux sociaux défavorisés dans chaque école (Cantillon et Gothelf, 2009; Cantillon, 2013). Dans d'autres pays, la mise en œuvre de l'objectif de mixité sociale passe par l'attribution de bonifications spécifiques aux élèves boursiers sur critères sociaux, à l'image des procédures de choix scolaire utilisées en Espagne pour affecter les élèves aux écoles élémentaires (Cal-samiglia et Güell, 2018) ou de la procédure Affelnet post-troisième qui régule en France les affectations au lycée (voir Grenet, 2022, Section 1, « Affelnet et la régulation du choix scolaire au lycée »). Au Royaume-Uni, l'objectif de diversité met davantage l'accent sur la mixité scolaire, à travers la mise en place d'un système dit de *banding*, qui consiste dans certains districts scolaires à équilibrer les populations des écoles sur le plan de l'aptitude des élèves, le plus souvent sous la forme de quotas réservés aux élèves « forts », « moyens »

et « faibles » dans chaque établissement (Noden, West et Hind, 2014).

5.2 Diversité et stabilité : des objectifs potentiellement conflictuels

Un résultat important de la théorie de l'appariement est qu'il existe une tension fondamentale entre l'objectif de diversité dans le recrutement des établissements scolaires et l'objectif de stabilité de l'appariement, entendu au sens d'élimination de la « jalousie légitime ».

L'exemple suivant, issu de Echenique et Yenmez (2015), illustre cette tension fondamentale :

Vœux d'affectation	Priorités des candidats	Capacité des formations
$i_1 : (f_1, f_2)$	$f_1 : (\{i_1 + i_2\}, \emptyset)$	$q_1 = 2$
$i_2 : (f_2, f_1)$	$f_2 : (i_1, i_2)$	$q_2 = 1$

On suppose ici que les élèves i_1 et i_2 incarnent des « types » différents (par exemple, l'un est socialement favorisé alors que l'autre est défavorisé). Les préférences de la formation f_1 traduisent une préférence pour la diversité dans le recrutement. Cette formation, qui peut accueillir deux élèves, préfère en effet accueillir un élève de chaque type plutôt que l'un ou l'autre séparément. À défaut de ne pouvoir accueillir les deux, elle préfère n'admettre personne, ce que l'on désigne par la notation $(\{i_1 + i_2\}, \emptyset)$ pour exprimer ses préférences.

Il est aisé de montrer que, dans cette configuration, il n'existe aucun appariement stable, c'est-à-dire aucun appariement qui élimine la « jalousie légitime » : si i_1 et i_2 sont affectés à l'école f_1 , i_2 préférerait être admis dans l'école f_2 où il reste une place vacante ; si i_2 est affecté à l'école f_2 , il ne peut pas être affecté à l'école f_1 et est dans une situation de jalousie légitime vis-à-vis de i_1 pour l'école f_2 (puisque l'on a priorité sur cet élève dans cette école) ; enfin, si i_1 est affecté à l'école f_2 , i_2 ne peut être affecté à l'école f_1 et se retrouve sans affectation ; dans ce cas, les deux élèves préféreraient être admis dans l'école f_1 , qui pourrait accueillir l'un et l'autre.

Cet exemple montre que les préférences pour la diversité ne sont pas toujours compatibles avec le critère normatif de stabilité de l'appariement. Ce résultat, de portée très générale, implique en particulier que l'objectif de diversité ne peut être mis en œuvre par l'algorithme d'acceptation différée (qui produit toujours un appariement stable) que si

certaines restrictions sont imposées à la structure des priorités. Dans l'exemple précédent, la raison pour laquelle les préférences de l'école f_1 sont incompatibles avec le critère de stabilité de l'appariement, et ne peuvent pas par conséquent être satisfaites par l'algorithme d'acceptation différée, est que ces préférences imposent une forme de complémentarité entre les élèves i_1 et i_2 : chacun d'eux ne peut être admis dans l'école f_1 que si l'autre y est admis.

Le « conflit » entre diversité et stabilité ne peut être résolu qu'en adoptant des règles de priorité qui préservent la substituabilité entre élèves, c'est-à-dire des priorités qui soient définies sur des individus (par exemple, « l'élève i_1 a priorité sur l'élève i_2 ») plutôt que sur des ensembles d'individus (par exemple, « les élèves i_1 et i_2 ne peuvent être admis qu'ensemble »).

5.3 Quotas et bonifications

En pratique, deux types de mécanismes permettent de mettre en œuvre un objectif de diversité dans le cadre des procédures d'affectation centralisées : l'instauration de quotas et la mise en place de bonifications dans un système de priorités par points.

Les quotas peuvent être de deux types : « rigides » ou « souples ». Pour illustrer leur fonctionnement, on considérera que les élèves se répartissent entre un groupe majoritaire (par exemple, les élèves non boursiers), noté M , et d'un groupe minoritaire, noté m (par exemple, les élèves boursiers). Ce modèle peut être aisément généralisé pour tenir compte de l'existence de plus de deux groupes au sein de la population.

Le système des quotas « rigides » (*majority quotas*) consiste à réserver, dans chaque école, des places qui ne peuvent être occupées que par les élèves du groupe minoritaire, les autres places pouvant être occupées indifféremment par les élèves de l'un ou l'autre groupe. Pour mettre en œuvre ce type de quotas dans le cadre d'une procédure s'appuyant sur l'algorithme d'acceptation différée, il suffit : 1) de scinder chaque école en deux « mini-écoles » (*mini-schools*), la première ne comprenant que les places réservées aux élèves du groupe minoritaire et la seconde les places ouvertes à l'ensemble des élèves, et 2) de modifier les vœux des élèves comme suit : les vœux des élèves du groupe minoritaire sont « dédoublés » de manière à ce que, pour un vœu donné, ils candidatent d'abord aux places qui leur sont exclusivement réservées, puis aux places ouvertes à tous ; les vœux des élèves du groupe majoritaire sont également modifiés de manière à ce que ces élèves

ne candidatent que sur les places qui sont ouvertes à tous. Les priorités des élèves restent quant à elles inchangées. Appliqué à ces vœux et priorités modifiés, l'algorithme d'acceptation différée permet d'obtenir un appariement qui vérifie simultanément les propriétés de stabilité et de respect des quotas.

Le système des quotas « souples » (*minority reserves*) se distingue du précédent en ce que les quotas attribués aux élèves du groupe minoritaire ne leur sont pas *exclusivement* réservés (Hafalir, Yenmez et Yildirim, 2013). Ces élèves sont seulement *prioritaires* pour ces places qui, si elles ne sont pas toutes occupées par des élèves du groupe minoritaire, peuvent être attribuées à des élèves du groupe majoritaire. La mise en œuvre de ces quotas procède de la même manière que pour les quotas rigides, avec deux modifications supplémentaires : d'une part, les vœux des élèves du groupe majoritaire sont « dédoublés » de manière à ce que, pour un vœu donné, ils candidatent d'abord aux places ouvertes à tous, puis aux places réservées en priorité aux élèves du groupe minoritaire ; d'autre part, les priorités des élèves candidatant aux places réservées au groupe minoritaire sont modifiées pour faire en sorte que les élèves de ce groupe soient tous classés au-dessus des élèves du groupe majoritaire.

Dans les procédures de choix régulé où les priorités sont calculées au moyen d'un barème par points, une manière alternative de mettre en œuvre un objectif de diversité consiste à attribuer des bonifications spécifiques aux élèves dont on souhaite favoriser une meilleure représentation dans les établissements scolaires les plus demandés. À la différence des quotas, les bonifications accordées au nom d'un objectif de diversité ne requièrent pas d'adaptation particulière de l'algorithme d'affectation différée, puisqu'elles se contentent de modifier l'ordre de priorité des candidats. En France, ce type de mécanisme est utilisé dans le cadre de la procédure Affelnet post-troisième, à travers la mise en place d'un bonus spécifique en faveur des élèves boursiers sur critères sociaux (voir Grenet, 2022, Section 1, « Affelnet et la régulation du choix scolaire au lycée »).

5.4 Le choix des instruments : un arbitrage délicat

Si les procédures de choix scolaire peuvent être aisément adaptées pour mettre en œuvre un objectif de diversité dans le recrutement des établissements scolaires, le choix de l'instrument à privilégier demeure une question épineuse.

Ce choix met en jeu un arbitrage entre deux conceptions distinctes des politiques de

diversité : la première, plus proche du sens originel du terme, met l'accent sur la nécessité de renforcer la cohésion sociale en favorisant les interactions entre membres de groupes définis en fonction de critères d'origine sociale, ethnique ou culturelle, alors que la seconde, plus proche de la philosophie de l'*affirmative action*, vise à favoriser certaines catégories d'individus en vue de corriger des inégalités de fait.

Appliquées aux procédures de choix scolaire, ces deux conceptions se distinguent par l'importance qu'elles accordent au critère de satisfaction des préférences individuelles par rapport au critère de diversité dans le recrutement. Alors que la première approche s'efforce d'obtenir une composition des établissements scolaires qui s'éloigne le moins possible d'une répartition « idéale » entre les différentes catégories d'élèves (favorisés/défavorisés, blancs/noirs, bons/moyens/faibles), la seconde se donne pour objectif de satisfaire au maximum les préférences des catégories d'élèves que l'on souhaite avantager.

La littérature théorique a mis en évidence l'existence d'une tension fondamentale entre ces deux approches de la diversité, qui sous-tend un résultat d'impossibilité démontré par Kojima (2012) : les instruments mobilisables pour mettre en œuvre un objectif de diversité dans les procédures de choix scolaire (quotas rigides ou souples, bonifications) ne peuvent jamais totalement éliminer le risque de pénaliser *in fine* les élèves que l'on souhaite pourtant avantager. Ce paradoxe peut être illustré au moyen de l'exemple suivant :

Vœux d'affectation	Priorités des élèves	Capacité des écoles
$M_1 : (f_1, f_2)$	$f_1 : (M_1, M_2, m_3)$	$q_1 = 2$
$M_2 : (f_1, f_1)$	$f_2 : (M_2, m_3, M_1)$	$q_2 = 1$
$m_3 : (f_2, f_1)$		

Dans cet exemple, les deux élèves M_1 et M_2 appartiennent au groupe majoritaire alors que l'élève m_3 appartient au groupe minoritaire. En l'absence de quotas, l'appariement obtenu par l'algorithme d'acceptation différée conduit à une ségrégation totale entre les écoles :

$$\mu_{DA} = \begin{pmatrix} M_1 & M_2 & m_3 \\ f_1 & f_1 & f_2 \end{pmatrix}$$

Tous les élèves du groupe majoritaire sont en effet regroupés dans l'école f_1 alors que l'école f_2 n'admet qu'un élève du groupe minoritaire.

Supposons que, pour diversifier le recrutement de l'école f_1 , on impose à cette école

de n'admettre au maximum qu'un élève du groupe majoritaire, ce qui peut être obtenu en réservant l'une des deux places de cette école aux élèves du groupe minoritaire exclusivement. En tenant compte de ce quota, l'algorithme d'acceptation différée produit l'appariement suivant :

$$\mu_{DA}^m = \begin{pmatrix} M_1 & M_2 & m_3 \\ f_1 & f_2 & f_1 \end{pmatrix}$$

Bien que ce nouvel appariement permette de réduire la ségrégation entre les deux écoles, l'élève du groupe minoritaire (m_3) voit sa situation se *dégrader* du fait de l'application du quota : il est désormais affecté sur son second vœu alors qu'en l'absence de quota, il aurait obtenu son premier vœu. Cette conséquence indésirable provient de ce qu'en limitant l'accès des élèves du groupe majoritaire à l'école f_1 , le quota reporte la demande des élèves du groupe majoritaire sur l'école f_2 . Ce phénomène pénalise en définitive l'élève du groupe minoritaire (m_3) : en concurrence avec un élève du groupe majoritaire (M_2) pour l'école f_2 , il se voit refuser l'accès à cette école qu'il a classée en premier vœu et est finalement admis dans l'école f_1 .

Cet exemple montre que l'objectif de diversité peut, dans certaines configurations, entrer en conflit avec la satisfaction des préférences des élèves que l'on souhaitait avantager : si, comme dans l'exemple, les préférences des élèves sont très polarisées en fonction de leur type (origine sociale ou ethnique, notamment), l'utilisation d'instruments visant à diversifier le recrutement des écoles peut aller à l'encontre des préférences des élèves non seulement du groupe majoritaire, mais également du groupe minoritaire.

Dans ce contexte, le choix entre les différents types d'instruments (quotas rigides, quotas souples, bonifications) dépend de l'arbitrage réalisé entre les objectifs de diversité et de respect des préférences (Echenique et Yenmez, 2015). Si l'accent est mis sur la diversité (*diversity first*), le choix se portera sur les quotas rigides, que les priorités des élèves soient exprimées en points ou en classes d'équivalence. Ce système est en effet celui qui permet de se rapprocher le plus possible de la répartition souhaitée dans chaque école ; au contraire, si l'accent est mis sur la satisfaction des préférences des catégories d'élèves que l'on souhaite avantager (*preferences first*), le choix se portera sur un système de quotas souples ou sur un système de bonification (dans le cas où les priorités sont exprimées selon un barème en points). Ces mécanismes permettent de respecter davantage les préférences

des élèves mais sans garantir que les différentes catégories d'élèves se répartissent de manière équilibrée au sein des établissements scolaires¹³.

6 Conclusion : des outils efficaces mais perfectibles

Le climat de suspicion qui s'est propagé en France à propos des algorithmes d'affectation des élèves et des étudiants ne rend pas justice aux progrès considérables qui ont été accomplis par ce champ de la recherche depuis les travaux fondateurs de Gale et Shapley. Partant du constat que les procédures les plus communément utilisées à travers le monde pour réguler les inscriptions scolaires et universitaires sont non seulement inefficaces, mais aussi porteuses d'inégalités, les travaux des théoriciens de l'appariement ont permis d'identifier des procédures d'affectation centralisées qui, à défaut de posséder toutes les propriétés attendues, réalisent des compromis optimaux entre la satisfaction des préférences, le respect des règles de priorité et la simplicité stratégique du processus d'appariement.

Parmi les algorithmes proposés dans la littérature, le mécanisme d'acceptation différée de Gale et Shapley s'est imposé comme la meilleure option pour remplacer les algorithmes d'affectation qui, à l'image du mécanisme de Boston, sont déficients car manipulables. Parce qu'il incite à la révélation sincère des préférences et maximise leur satisfaction sous la contrainte de respecter les règles de priorité, l'algorithme d'acceptation différée est aujourd'hui considéré comme celui qui offre les meilleures garanties de lisibilité et d'équité du processus d'affectation. À l'inverse, le peu de succès rencontré par l'algorithme des cycles d'échanges optimaux, qui privilégie la satisfaction des préférences par rapport au respect des priorités, est révélateur de la profonde aversion que suscite toute entorse au principe d'équité dans les procédures d'affectation.

Malgré les progrès accomplis dans la réflexion théorique sur les algorithmes d'appariement, il reste encore beaucoup à apprendre sur leur fonctionnement dans le monde réel, à partir notamment des données très riches que les procédures d'affectations scolaires et universitaires ont permis de collecter. Les déterminants des préférences des élèves et des

13. En vertu du résultat d'impossibilité énoncé plus haut, aucun des dispositifs présentés ici ne permet de garantir que *tous* les élèves ciblés par les quotas ou les bonifications soient *effectivement* avantagés dans leur affectation. Dans certaines configurations, les quotas ou bonifications peuvent en effet conduire une partie de ces élèves à être affectés sur un voeu de rang inférieur à celui qu'ils obtiendraient sans ces mesures, en raison du phénomène d'éviction mis en évidence dans l'exemple plus haut.

étudiants, la transcription des objectifs de politique éducative dans les règles de priorité ou encore la manière dont ces procédures interagissent avec les autres « marchés » scolaires (l'enseignement privé, en particulier) sont autant de questions cruciales qui appellent des réponses empiriques. Entre le rejet en bloc des algorithmes d'affectation et la foi aveugle en leur toute-puissance, la démarche scientifique propose une voie médiane : les considérer comme des outils efficaces mais perfectibles.

Références

- Abdulkadiroğlu A. (2021).** « School Choice », in Vulkan N., Roth A., Neeman Z., éd., *Handbook of Market Design*, Oxford University Press, Oxford, p. 138-169.
- Abdulkadiroğlu A., Che Y.-K., Pathak P. A., Roth A. E., Tercieux O. (2017).** « Minimizing Justified Envy in School Choice : The Design of New Orleans' OneApp », NBER Working Paper No. 23265.
- Abdulkadiroğlu A., Che Y.-K., Yasuda Y. (2011).** « Resolving Conflicting Interests in School Choice. The "Boston Mechanism" Reconsidered », *American Economic Review*, 101(1), p. 399-410.
- Abdulkadiroğlu A., Pathak P. A., Roth A. E. (2009).** « Strategy-proofness versus Efficiency in Matching with Indifferences : Redesigning the NYC High School Match », *American Economic Review*, 99(5), p. 1954-1978.
- Abdulkadiroğlu A., Pathak P. A., Roth A. E., Sönmez T. (2006).** « Changing the Boston School Choice Mechanism », NBER Working Paper No. 11965.
- Abdulkadiroğlu A., Sönmez T. (1999).** « House Allocation with Existing Tenants », *Journal of Economic Theory*, 88(2), p. 233-260.
- Abdulkadiroğlu A., Sönmez T. (2003).** « School Choice : A Mechanism Design Approach », *American Economic Review*, 93(3), p. 729-747.
- Agarwal N., Somaini P. (2018).** « Demand Analysis Using Strategic Reports : An Application to a School Choice Mechanism », *Econometrica*, 86(2), p. 391-444.
- Alcalde J., Barberà S. (1994).** « Top Dominance and the Possibility of Strategy-proof Stable Solutions to Matching Problems », *Economic Theory*, 4, p. 417-435.
- Balinski M., Sönmez T. (1999).** « A Tale of Two Mechanisms : Student Placement », *Journal of Economic Theory*, 84, p. 73-94.
- Biró P. (2017).** « Applications of Matching Models under Preferences », in Endriss U., éd., *Trends in Computational Social Choice*, AI Access, El Segundo (Calif.), p. 345-373.
- Calsamiglia C., Fu C., Güell M. (2020).** « Structural Estimation of a Model of School Choices : The Boston Mechanism versus Its Alternatives », *Journal of Political Economy*, 128(2), p. 642-680.
- Calsamiglia C., Güell M. (2018).** « Priorities in School Choice : The Case of the Boston Mechanism in Barcelona », *Journal of Public Economics*, 163, p. 20-36.
- Cantillon E. (2013).** « Mixité sociale : le rôle des procédures d'inscriptions scolaires », in Maystadt Ph. et al., éd., *Actes du 20^e Congrès des économistes belges de langue française*, Centre interuniversitaire de formation permanente, Charleroi.
- Cantillon E. (2017).** « Broadening the Market Design Approach to School Choice », *Oxford Review of Economic Policy*, 33(4), p. 613-634.
- Cantillon E., Gothelf N. (2009).** « Applications of Matching Models under Preferences », in de Callataÿ, E., et al., éd., *Actes du 18^e Congrès des économistes belges de langue française*, Centre interuniversitaire de formation permanente, Charleroi.
- Combe J., Hiller V., Tercieux O., Benoît A., He Y., Jacquelinet C., Macher M.-A. (2022).** « Perspectives sur le programme de dons croisés de reins en France », Note IPP n°41, Institut des politiques publiques.
- Dubins L. E., Freedman D. A. (1981).** « Machiavelli and the Gale-Shapley Algorithm », *American Mathematical Monthly*, 88(7), p. 485-494.
- Echenique F., Yenmez M. B. (2015).** « How to Control Controlled School Choice », *American Economic Review*, 105(8), p. 2679-2694.

- Ellison G., Pathak P. A. (2021).** « The Efficiency of Race-Neutral Alternatives to Race-Based Affirmative Action : Evidence from Chicago’s Exam Schools », *American Economic Review*, 111(3), p. 943-975.
- Erdil A., Ergin H. (2006).** « Games of School Choice under the Boston Mechanism », *Journal of Public Economics*, 90(1-2), p. 215-237.
- Erdil A., Ergin H. (2008).** « What’s the Matter with Tie-Breaking? Improving Efficiency in School Choice », *American Economic Review*, 98(3), p. 669-698.
- Erdil A., Ergin H. (2013).** « Appariement : des modèles de Lloyd Shapley à la conception des marchés d’Alvin Roth », *Revue d’économie politique*, 123(5), p. 663-696.
- Fack G., Grenet J., He Y. (2019).** « Preference Estimation with Centralized School Choice and College Admissions », *American Economic Review*, 109(4), p. 1486-1529.
- Gale D., Shapley L. (1962).** « College Admissions and the Stability of Marriage », *The American Mathematical Monthly*, 69(1), p. 9-15.
- Grenet J. (2022).** « Les algorithmes d’affectation dans le système éducatif français », in Simioni M., Steiner P., éd., *Comment ça matche? Une sociologie de l’appariement*, chap. 1, Presses de Sciences Po, collection « Gouvernances », Paris. Url : https://www.parisschoolofeconomics.com/grenet-julien/Matching/Grenet_2022_chap1.pdf, p. 21-59.
- Haeringer G., Klijn F. (2009).** « Constrained School Choice », *Journal of Economic Theory*, 144, p. 1921-1947.
- Hafalir I. E., Yenmez M. B., Yildirim M. A. (2013).** « Effective Affirmative Action in School Choice », *Theoretical Economics*, 8, p. 325-363.
- He Y. (2017).** « Gaming the Boston School Choice Mechanism in Beijing », Toulouse School of Economics, Working Paper no 15-551.
- Hylland A., Zeckhauser R. (1979).** « The Efficient Allocation of Individuals to Positions », *Journal of Political Economy*, 87(2), p. 293-314.
- Kesten O. (2010).** « School Choice with Consent », *The Quarterly Journal of Economics*, 125(2), p. 1297-1348.
- Kojima F. (2012).** « School Choice : Impossibilities for Affirmative Action », *Games and Economic Behavior*, 75, p. 685-693.
- Miralles A. (2009).** « School Choice : The Case for the Boston Mechanism », in Das S., Ostrovky M., Pennock D., Szymanski B., éd., *Auctions, Market Mechanisms and Their Applications*, Springer, Boston (Mass.), p. 58-60.
- Noden P., West A., Hind A. (2014).** *Banding and Ballots. Secondary School Admissions in England : Admissions in 2012/13 and the Impact of Growth of Academies*, The Sutton Trust, Londres.
- Pathak P. A. (2011).** « The Mechanism Design Approach to Student Assignment », *Annual Review of Economics*, 3, p. 513-536.
- Pathak P. A. (2019).** « What Really Matters in Designing School Choice Mechanisms », in Honoré B., Pakes A., Piazzesi M., Samuelson L., éd., *Advances in Economics and Econometrics, Eleventh World Congress, vol. 1*, Cambridge University Press, Cambridge, p. 176-214.
- Pathak P. A., Sönmez T. (2013).** « School Admissions Reform in Chicago and England : Comparing Mechanisms by their Vulnerability to Manipulation », *American Economic Review*, 103(1), p. 80-106.
- Roth A. E. (1982).** « The Economics of Matching : Stability and Incentives », *Mathematics of Operations Research*, 7(4), p. 617-628.

- Roth A. E. (1985).** « The College Admissions Problem is not Equivalent to the Marriage Problem », *Journal of Economic Theory*, 36(2), p. 277-288.
- Roth A. E. (2008).** « Deferred Acceptance Algorithms : History, Theory, Practice, and Open Questions », *International Journal of Game Theory*, 36(3), p. 537-569.
- Roth A. E. (2008).** « The Theory and Practice of Market Design », Nobel Lecture, 8 décembre 2012.
- Roth A. E., Peranson E. (1990).** « The Redesign of the Matching Market for American Physicians : Some Engineering Aspects of Economic Design », *American Economic Review*, 89(4), p. 748-780.
- Roth A. E., Sönmez T., Ünver U. M. (2004).** « Kidney Exchange », *The Quarterly Journal of Economics*, 119(2), p. 457-488.
- Roth A. E., Sönmez T., Ünver U. M. (2005).** « A Kidney Exchange Clearinghouse in New England », *American Economic Review*, 95(2), p. 376-380.
- Shapley L., Scarf H. (1974).** « On Cores and Indivisibility », *Journal of Mathematical Economics*, 1(1), p. 23-37.
- Steiner P. (2010).** *La Transplantation d'organes. Un commerce nouveau entre les êtres humains*, Gallimard, Paris.