

Institut d'études politiques de Paris
ECOLE DOCTORALE DE SCIENCES PO
Programme doctoral en économie
Département d'économie

Doctorat en sciences économiques

Trois essais en économie appliquée (résumé)

Jean-Baptiste VILAIN

Thèse dirigée par Alfred GALICHON et
supervisée par Philippe FÉVRIER
défendue le 15 juin 2018

Jury

- Luc ARRONDEL, Directeur de recherche CNRS, Paris School of Economics
Rapporteur
- Ghazala AZMAT, Professeur des Universités, IEP de Paris
- Philippe FÉVRIER, Professeur associé, Ecole Polytechnique & CEO de Veltys
Responsable scientifique au sein de l'entreprise Veltys (thèse CIFRE)
- Alfred GALICHON, Professor of Economics and of Mathematics, New York University
Directeur de thèse
- Michael VISSER, Directeur de recherche CNRS, CREST
Rapporteur

Résumé en français

Résumé court

Cette thèse porte sur plusieurs questions relatives à l'économie de l'éducation et à l'économie des équipes.

Dans le premier chapitre, basé sur une collaboration avec Laurent Rossignol, nous nous concentrons sur l'orientation scolaire. La plupart des systèmes éducatifs publics regroupent les élèves selon leurs performances académiques. Cette séparation en différents groupes de niveau peut prendre plusieurs formes. Aux Etats-Unis et au Canada, les étudiants sont répartis entre différentes classes de niveau dans la même école, alors que dans la plupart des pays européens, les élèves sont répartis dans différentes écoles avec des programmes scolaires différents dans le secondaire (Betts et al. (2011)). Nous mettons en évidence l'existence de biais d'orientation dans le système éducatif français : l'orientation des élèves ne dépend pas uniquement de leur performance académique mais aussi de leur sexe et de leur origine sociale. Notre contribution principale est de distinguer l'impact des aspirations des élèves de l'impact de la notation des professeurs et de leurs recommandations sur ces biais d'orientation.

Dans les deux chapitres suivants, nous quittons l'économie de l'éducation pour l'économie des équipes. L'organisation du travail en équipe s'est considérablement développée dans la plupart des organisations au cours des dernières décennies (Hamilton, Nickerson & Owan (2003)). Cette tendance soulève de nombreuses questions économiques de premier ordre. Pourquoi le travail en équipe est-il efficace d'un point de vue économique ? Quelles sont les incitations associées au travail en équipe ? Comment peut-on évaluer la contribution individuelle dans une équipe, quand seulement la production collective est observée ? Comment la composition d'une équipe

affecte-t-elle la productivité ? Les équipes sont-elles plus performantes lorsque leurs membres sont homogènes ou hétérogènes ? Dans les chapitres 2 et 3 de cette thèse, nous utilisons des données relatives au sport pour traiter certaines de ces questions. Le sport apparaît en effet comme un excellent laboratoire pour étudier la question de la contribution individuelle en équipe et pour tester des théories économiques sur les incitations et les comportements sur le marché du travail parce que les règles qui s’y appliquent sont clairement définies et relativement simples (Szymanski (2003)). De plus, les données sportives sont souvent accessibles publiquement, ce qui en fait un terrain empirique très attractif. Le second chapitre, coécrit avec Antoine Chapsal, vise à comprendre certaines incitations et effets psychologiques associés au travail en équipe, à partir de données sur les championnats par équipe de squash. Nous montrons que les joueurs valorisent le fait de participer au succès de leur équipe, ce qui explique en partie pourquoi les incitations à l’effort sont plus fortes dans des contextes collectifs que dans des contextes individuels. Le troisième chapitre, issu d’un travail initial avec Rodrigo Lopez-Kolkovsky, a pour but de développer une procédure d’estimation pour mesurer la productivité individuelle en équipe, à partir de données sur le football européen. Nous confrontons ensuite cette mesure à la valeur de marché des joueurs et nous montrons que les joueurs Noirs sont discriminés sur le marché.

Résumé long

Chapitre 1 : L'impact des enseignants sur l'orientation scolaire via la notation et les recommandations

Mon premier chapitre, issu d'une collaboration avec Laurent Rossignol, porte sur l'orientation scolaire, qui peut être définie, dans un sens large, comme la répartition des élèves en différents groupes de niveau.

Le point de départ de notre étude est que l'orientation ne dépend pas uniquement du niveau académique des élèves. En effet, de nombreux articles en économie et en sociologie de l'éducation ont montré l'existence de biais d'orientation : les caractéristiques individuelles comme le sexe, l'origine ethnique ou la catégorie socio-économique des parents ont un effet sur l'orientation des élèves, indépendamment de leur niveau scolaire (voir Duru-Bellat, Jarousse & Mingat (1993) pour une analyse du système d'orientation français).

Ces différences d'orientation peuvent être en partie expliquées par des préférences ou des aspirations différentes des élèves (Guyon & Huillery (2016)). Cependant, les professeurs interviennent aussi dans la procédure d'orientation et sont donc susceptibles d'avoir un impact sur l'orientation des élèves. Cet impact peut d'abord être indirect via la notation. Si les professeurs notent différemment les élèves en fonction de leurs caractéristiques individuelles, cela affecte indirectement l'orientation car les notes sont utilisées pour évaluer le potentiel scolaire des élèves et leurs chances de réussite dans les différentes voies de formation. Les enseignants peuvent également avoir un impact direct sur l'orientation via leurs recommandations d'orientation lors des conseils de classe s'ils sont moins enclins à recommander la 2^{de} générale et technologique à certaines catégories d'élèves. Évaluer le rôle des professeurs dans la notation et l'orientation apparaît donc comme une question de recherche essentielle pour mieux comprendre les inégalités

d'orientation documentées dans la littérature.

Nous utilisons des données très riches sur tous les élèves de 3ème des collèges publics de l'académie d'Amiens, qui est l'une des 26 académies en France métropolitaine (N=16864). Nos données, issues du recroisement de bases nationales et académiques contiennent des informations sur les caractéristiques des élèves (PCS, sexe, nationalité), leur performance scolaire (notes obtenues au contrôle continu et à l'examen final du *Diplôme national du brevet*) et sur la procédure d'orientation (choix initial de l'élève et recommandation du conseil de classe).

Notre cadre empirique semble idéal pour étudier le rôle des professeurs sur la notation et l'orientation parce qu'il présente deux caractéristiques importantes. D'une part, les élèves sont évalués à la fois lors de tests passés en classe qui sont préparés et corrigés par leurs propres enseignants (contrôle continu) et lors d'un examen final commun à tous les élèves de l'académie qui est corrigé de façon anonyme. Cela nous permet d'analyser la notation des professeurs. D'autre part, dans la procédure d'orientation, les élèves font un choix initial d'orientation auquel le conseil de classe répond par une recommandation, ce qui nous permet de dissocier l'effet des aspirations des élèves de l'effet des recommandations des professeurs sur l'orientation.

Nous nous intéressons d'abord à la notation des professeurs en comparant les résultats des élèves au contrôle continu et à l'examen final. Notre hypothèse identificatrice est la suivante : la note au contrôle continu reflète le niveau scolaire *observé* ou *perçu* des élèves alors que la note à l'examen final reflète le *vrai* niveau scolaire des élèves. Cette hypothèse est justifiée par plusieurs caractéristiques du système de notation français:¹ (i) L'examen final a lieu à la toute fin de l'année scolaire. Seules les notes au contrôle continu sont donc observées au cours de la procédure d'orientation. (ii) Les tests du contrôle continu sont corrigés par les propres en-

¹La description présentée ici s'applique à l'année scolaire pour laquelle nous avons des données (2013/2014).

seignants des élèves, sans être anonymisés. Ils sont donc soumis à d'éventuels biais de notation, à la différence de l'examen final qui est corrigé de façon anonyme par des professeurs extérieurs à l'établissement. (iii) Les enseignants sont complètement libres en ce qui concerne le format et la difficulté des tests du contrôle continu. Ces derniers sont donc très sensibles à la spécificité contextuelle des classes. Au contraire, l'examen final est une épreuve standardisée commune à tous les élèves de l'Académie. (iv) Les tests du contrôle continu sont typiquement très courts et couvrent seulement la dernière leçon du cours, alors que l'examen final couvre l'ensemble du programme de 3ème et reflète donc mieux les acquis de long-terme des élèves.

Notre analyse montre qu'à performance équivalente à l'examen final, les filles et les élèves favorisés obtiennent des scores au contrôle continu qui sont plus élevés de respectivement 0,16 et 0,1 SD que ceux des garçons et des élèves défavorisés. Au vu de notre hypothèse identificatrice, nous interprétons ces différences comme des biais de notation.

Nous nous concentrons ensuite sur la procédure d'orientation. Nous montrons empiriquement que les aspirations des élèves dépendent fortement de leur origine sociale et de leur sexe. Les élèves favorisés et les filles ont en effet une probabilité plus élevée de choisir initialement la voie générale et technologique que les élèves défavorisés et les garçons de la même classe qui obtiennent les mêmes notes au contrôle continu (respectivement +10,2 points de pourcentage (pp) et + 3,4 pp). Ces résultats confirment la littérature existante (voir par exemple Vrignaud (2016)).

Intégrer notre analyse sur les biais de notation dans notre modèle sur les choix d'orientation des élèves nous permet d'identifier l'effet indirect de la notation des enseignants sur les aspirations des élèves. Lorsque nous utilisons la note à l'examen final plutôt que la note au contrôle continu dans l'estimation, les différences dans les choix initiaux d'orientation deviennent plus marquées en fonction de la PCS et du sexe (respectivement 12,3 pp et 6,7 pp). Ces résultats montrent que

les professeurs influent indirectement sur les aspirations des élèves via leurs biais de notation. L'impact "direct" des enseignants sur l'orientation peut être évalué à partir des recommandations d'orientation du conseil de classe. Comparer à quel point les coefficients associés à la PCS et au genre évoluent lorsque l'on considère la recommandation des enseignants plutôt que le choix initial de l'élève, tout en contrôlant par la note au contrôle continu, permet d'identifier si les enseignants accroissent ou réduisent les différences d'aspirations des élèves. Nous montrons que les enseignants ne "corrige" pas les écarts de choix initiaux entre les filles et les garçons, mais qu'ils réduisent les écarts de choix initiaux entre les élèves favorisés et les élèves défavorisés (le coefficient diminue de 10,2 pp à 5,4 pp).

L'intégration de notre analyse sur les biais de notation dans notre modèle sur les recommandations nous permet de mesurer directement l'impact global des professeurs sur l'orientation, c'est-à-dire la somme de l'effet indirect de leurs biais de notation et de l'effet direct de leurs recommandations sur l'orientation. Nous montrons qu'au global, les enseignants accroissent le différentiel entre filles et garçons dans les choix initiaux d'orientation. Cet impact global est cohérent avec nos résultats précédents : les filles sont mieux notées en classe et les enseignants n'utilisent pas leurs recommandations pour réduire le différentiel initial entre les aspirations des filles et des garçons. L'impact global des enseignants sur l'orientation en fonction de la PCS des élèves est plus contrasté. D'un côté, la notation désavantage les élèves défavorisés, ce qui influe indirectement sur leur orientation. D'un autre côté, les enseignants utilisent leurs recommandations pour réduire le différentiel entre élèves favorisés et élèves défavorisés dans les choix initiaux d'orientation. Nos résultats indiquent que ces deux effets antagonistes s'annulent et qu'au global, les professeurs ne réduisent pas ou n'accentuent pas l'écart socio-économique initial dans les aspirations.

Ainsi, la contribution principale de ce chapitre est de construire un pont entre la notation et l'orientation scolaire, qui sont généralement traitées comme deux questions distinctes. Cette approche nous permet, non seulement d'estimer les biais de notation des enseignants, mais aussi leurs effets indirects sur l'orientation des élèves.

Chapitre 2 : L'effet de contribution individuelle dans les tournois par équipes

Le point de départ du second chapitre, coécrit avec Antoine Chapsal, est que malgré les incitations de *free-riding* inhérentes aux équipes (Alchian & Demsetz (1972)), l'organisation du travail en équipe est devenue la norme dans la plupart des organisations. L'incitation "négative" de *free-riding*² doit donc être compensée et même dépassée par des incitations "positives" liées au travail en équipe. Dans un article théorique, Kandel & Lazear (1992) montre que la pression des pairs, la culpabilité et la honte limitent le *free-riding*. Dans ce chapitre, nous nous intéressons à une autre incitation positive susceptible de limiter le *free-riding* : "l'effet de contribution individuelle" selon lequel les membres d'une équipe valorisent le fait d'être (au moins partiellement) responsables du succès de leur équipe. Un tel effet pourrait expliquer pourquoi le niveau d'effort peut être plus élevé dans des contextes d'équipes que dans des contextes individuels et donc pourquoi le travail en équipe est devenu si fréquent dans la plupart des organisations.

Comme discuté dans le troisième chapitre, évaluer la contribution individuelle dans la production collective est souvent très difficile. Nous nous concentrons donc sur un contexte d'équipe particulier dans lequel la contribution individuelle est observable : les batailles par paires multi-

²L'expression *free-riding* en anglais peut être traduite par l'expression *phénomène de passager clandestin* en français et renvoie ici au fait de se reposer sur le travail de ses coéquipiers plutôt que de fournir un gros effort soi-même. L'expression française étant assez lourde, nous lui préférons l'expression anglaise dans ce résumé.

ples. Ces dernières se réfèrent à de nombreuses situations dans lesquelles deux équipes s’opposent dans des batailles individuelles (voir Fu, Lu & Pan (2015) pour une analyse théorique de ce type de tournoi). Chaque joueur étant responsable de sa bataille individuelle, l’évaluation des contributions individuelles au succès collectif est directe.

Les tournois de squash par équipe correspondent exactement à des batailles par paires multiples. En effet, dans une confrontation de squash par équipe, deux équipes de trois joueurs s’affrontent dans des matchs individuels qui sont joués de façon séquentielle. Une équipe doit gagner deux matchs individuels pour gagner la confrontation. Dans ce contexte, la contribution individuelle est donc observée puisque le joueur apporte un point à son équipe s’il gagne son match, et zéro sinon.

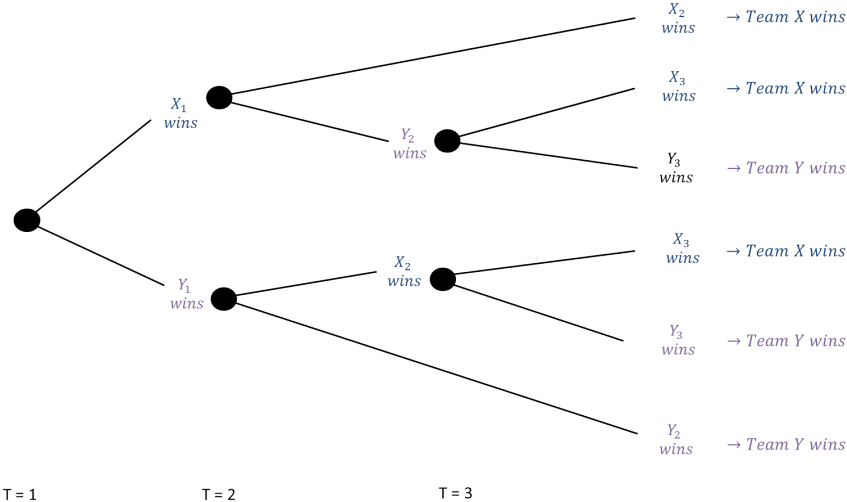


Figure 1: Les matchs de squash par équipe - un exemple type de batailles par paires multiples

En outre, les confrontations de squash par équipe présentent une caractéristique aléatoire qui en font un laboratoire idéal pour étudier les batailles par paires multiples. Au début d’un championnat, chaque équipe doit classer ses joueurs du plus performant au moins performant. Ensuite, au début de chaque confrontation par équipe, l’ordre des matchs individuels est tiré

aléatoirement parmi quatre possibilités : 1-2-3 qui signifie que les meilleurs joueurs de chaque équipe jouent le premier match individuel, que les seconds meilleurs joueurs de chaque équipe jouent le second match et et que les troisièmes meilleurs joueurs de chaque équipe jouent le troisième match, 1-3-2, 2-1-3 and 3-1-2. Ce tirage aléatoire ex-ante garantit que les équipes ne peuvent pas manipuler la séquence de matchs à jouer (par exemple en désignant un joueur ayant un style de jeu particulièrement adapté à son adversaire présumé).

Nous avons collecté à partir d'un site de squash spécialisé des données sur les résultats des matchs de squash par équipe dans les tournois mondiaux, européens et asiatiques sur la période 1998 à 2016. Nous avons également récupéré les classements mondiaux mensuels de l'ensemble des joueurs jouant au moins un match par équipe dans ces tournois. Le classement individuel des joueurs est basé uniquement sur leurs performances dans les tournois individuels.

D'après l'effet de contribution individuelle, l'utilité d'un joueur augmente quand il contribue au succès de son équipe, c'est-à-dire lorsque qu'il gagne son match individuel et que son équipe remporte la rencontre. Pour tester cet effet sur les données de squash, nous étudions si la probabilité de victoire d'un joueur dans le second match individuel est plus élevée lorsque son coéquipier a gagné le premier match individuel. L'intuition derrière ce test est que dans le second match, il y a une asymétrie entre les deux joueurs. Le joueur dans l'équipe qui mène sait que s'il gagne son match individuel, alors son équipe gagnera la confrontation et il obtiendra le payoff de contribution. Au contraire, le joueur de l'équipe menée n'obtiendra le payoff de contribution que s'il gagne son match individuel et si son coéquipier gagne le troisième match. Les deux joueurs ont donc des incitations différentes en ce qui concerne leur contribution individuelle, ce qui augmente la probabilité de victoire du joueur qui est dans l'équipe qui mène. Afin de tester ce mécanisme empiriquement, nous commençons par nous concentrer sur le sous-échantillon de rencontres pour lesquelles les joueurs impliqués dans le deuxième match ont des classements très

similaires. Nous montrons que le joueur qui est dans l'équipe qui mène gagne dans 60% des cas, ce qui confirme l'effet de contribution. Dans d'autres spécifications, nous utilisons l'intégralité de notre échantillon et nous introduisons le ratio des classements ou des modalités de classement pour contrôler le niveau relatif des joueurs qui s'affrontent. Dans ces estimations, nous trouvons que le joueur dans l'équipe qui mène a environ 55% de chance de gagner, ce qui plaide également en faveur de l'effet de contribution.

Nous étudions également l'impact du résultat attendu du troisième match individuel sur le résultat du second match individuel. Dans le second match, si le joueur qui est dans l'équipe menée est sûr que son coéquipier perdra le troisième match, l'asymétrie des incitations de contribution individuelle atteint son maximum, puisque le joueur dans l'équipe menée sait qu'il n'obtiendra pas le payoff de contribution, quel que soit son résultat individuel. Au contraire, si le joueur dans l'équipe menée est sûr que son partenaire gagnera le troisième match, il n'y a plus d'asymétrie car le second match devient décisif pour les deux joueurs. D'après l'effet de contribution individuelle, la probabilité de gagner le second match augmente donc avec la probabilité attendue de gagner le troisième match. Nous montrons empiriquement, en contrôlant par le classement des joueurs qui s'affrontent, que la probabilité de gagner le second match augmente quand le coéquipier engagé dans le troisième match est mieux classé que son adversaire.

Dans un autre test, nous nous concentrons sur les troisièmes matchs triviaux, c'est-à-dire les matchs pour lesquels l'équipe victorieuse est déjà connue à l'issue des deux premiers matchs individuels. L'effet de contribution individuelle pourrait aussi jouer, dans une moindre mesure, pour ces matchs. Le joueur dans l'équipe qui a remporté les deux premiers matchs a en effet plus d'incitations à gagner que son adversaire parce qu'il peut contribuer au succès de son équipe. Au contraire, le joueur dans l'équipe qui a perdu les deux premiers matchs ne peut pas contribuer au succès de son équipe car celle-ci a déjà perdu. Nous montrons empiriquement, à partir de

l'échantillon de batailles triviales, que le joueur dans l'équipe qui a gagné les deux premiers matchs a plus de chance de gagner que son adversaire, en contrôlant par le classement relatif des deux joueurs. Ce résultat confirme l'effet de contribution individuelle. Nous effectuons enfin plusieurs tests de robustesse pour tester d'autres mécanismes pouvant expliquer pourquoi gagner le premier match augmente la probabilité de gagner le second match. Cela nous permet de montrer que l'aversion à la culpabilité, l'altruisme, le momentum psychologique ou l'étouffement sous la pression ne sont pas les mécanismes à l'œuvre.

En résumé, nous utilisons des données sur les confrontations de squash par équipe pour montrer un lien de dépendance entre les matchs individuels dans les contextes de batailles par paires multiples. Nous montrons que cette dynamique vient de l'effet de contribution individuelle : les joueurs valorisent le fait de contribuer au succès de leur équipe. Cet effet est susceptible de limiter les incitations au free-riding en équipe, et d'être l'un des facteurs expliquant pourquoi le travail en équipe est si développé.

Chapitre 3 : Estimation de la productivité individuelle au football

Mon troisième chapitre, issu d'un travail initial avec Rodrigo Lopez-Kolkovsky, vise à mesurer le niveau individuel des footballeurs professionnels. Le football apparaît comme un laboratoire idéal pour explorer la question de la productivité individuelle en équipe pour trois raisons. D'abord, c'est un exemple typique d'environnement où seule la production collective peut être observée directement. Alors que le nombre de buts ou de points marqués par une équipe est connu, il est difficile d'évaluer dans quelle mesure chacun des joueurs y contribue. Ensuite, les règles sont clairement définies, ce qui en fait un environnement "sous contrôle". Enfin, l'explosion des données disponibles rend possible l'analyse statistique du sport le plus populaire du monde.

Deux approches peuvent être envisagées pour estimer la productivité individuelle des joueurs. La première approche est basée sur les faits de jeu individuels comme les passes et les tirs. Après avoir identifié les faits de jeu qui font le succès des équipes, la performance des joueurs peut être évaluée à partir de leurs participations à ces faits de jeu (voir par exemple Berri & Schmidt (2010)). Adopter cette approche dans le cadre du football ne semble pas forcément approprié parce que les faits de jeu pertinents ne sont pas clairement identifiés. La seconde approche est basée sur l'estimation d'effets fixes joueurs : la productivité d'un joueur dépend de sa participation aux différents matchs. En contrôlant par le niveau de ses coéquipiers et de ses adversaires, plus son équipe est performante quand il est présent, plus sa productivité est élevée. Dans ce type d'estimation, la variable dépendante correspond à la performance de l'équipe et les variables explicatives à des indicatrices précisant si les différents joueurs se trouvent ou non sur le terrain. Cette approche, qui s'appuie sur le *plus minus* développé aux Etats-Unis pour le hockey sur glace dans les années 1960 et sur ses versions ajustées (voir Rosenbaum (2004) et Ilardi & Barzilai (2008)), requiert très peu de données puisque seuls les scores finaux et les feuilles de matchs sont nécessaires à l'estimation. De plus, elle permet de considérer la productivité individuelle dans toutes ses dimensions, y compris celles qui ne sont pas observables dans les données comme le charisme des joueurs.

Nous avons collecté des données sur les feuilles de matchs et les scores finaux de tous les matches opposant deux équipes des cinq grands championnats européens entre les saisons 2007/2008 et 2014/2015 à partir de quatre sites internet différents.

L'identification des effets fixes joueurs est possible grâce à la variabilité des feuilles de matchs au cours du temps, due aux stratégies de coaching, aux transferts et aux blessures. Toutefois, l'estimation est compliquée car le nombre de coefficients à estimer est très grand : dans notre base de données finales, le ratio entre le nombre d'observations et le nombre de variables est

de seulement 8. En outre, la colinéarité est très forte dans la mesure où les joueurs jouent souvent avec les mêmes coéquipiers. Ces deux éléments impliquent qu'une estimation standard des effets fixes joueurs souffre de sur-apprentissage statistique : si les joueurs A et B sont toujours ensemble sur le terrain, sauf lors d'un match où seul le joueur A est aligné et où l'équipe est très performante, alors le joueur A aura un coefficient beaucoup plus élevé que le joueur B. Cela n'est pas désirable car le différentiel de performance entre les deux joueurs sera basé sur seulement un match. Une estimation standard va donc mener à des coefficients avec une variance très grande et à de mauvaises prédictions sur des données externes. Par conséquent, nous avons besoin d'ajuster la procédure d'estimation pour qu'elle aboutisse à des coefficients différents pour deux coéquipiers si et seulement s'il y a suffisamment de matchs dans lesquels les deux joueurs ne sont pas alignés simultanément.

Le Ridge est une méthode de pénalisation particulièrement adaptée à cet objectif. L'intuition derrière le Ridge est que l'estimation arbitre entre la qualité de l'ajustement (comme dans une estimation standard) et la variance des coefficients, en ajoutant un terme de pénalisation correspondant à la somme des coefficients estimés au carré dans le programme d'optimisation. Il en résulte que le Ridge va attribuer des coefficients élevés aux joueurs si et seulement s'il dispose de suffisamment d'observations et de variabilité pour être sûr que ces joueurs sont très performants. Ainsi, le Ridge permet de surmonter le problème de sur-apprentissage.

Nous choisissons le degré optimal de pénalisation λ du Ridge par validation croisée. Nous obtenons des résultats cohérents avec les avis des spécialistes du football. Lionel Messi et Cristiano Ronaldo, qui sont largement reconnus comme étant les deux meilleurs joueurs sur la période, sont en tête de notre classement des effets fixes joueurs. Tous les autres joueurs de notre top 30 sont des joueurs de classe internationale. Ces résultats sont très satisfaisants car le modèle est complètement agnostique sur ce qui fait qu'un joueur est productif ou non. Pour

tester la robustesse de notre estimation, nous utilisons plusieurs tests en échantillon et hors échantillon au niveau des équipes. L'intuition derrière ces tests est que les joueurs sont souvent transférés d'un club à un autre. Par conséquent, si le niveau des équipes est bien prédit, alors le niveau individuel des joueurs est également bien prédit car la compensation des coefficients des joueurs est peu probable.

Table 1: Classement des joueurs : top 10

Classement	Joueur	Total	Coefficient offensif	Coefficient défensif
1	Lionel Messi	.351	.351	
2	Cristiano Ronaldo	.333	.333	
3	Cesc Fabregas	.302	.243	.059
4	Frank Lampard	.274	.205	.069
5	Yaya Toure	.255	.138	.117
6	Karim Benzema	.253	.253	
7	Bastian Schweinsteiger	.246	.171	.075
8	Xavi	.241	.175	.066
9	Arturo Vidal	.238	.128	.11
10	Franck Ribery	.237	.237	

Seasons 2007/2008 to 2014/2015.

Notre mesure de la productivité individuelle des joueurs présente plusieurs limites. D'abord, la pénalisation implique que l'estimation ne fournit que les coefficients associés aux effets fixes, et pas leurs écarts-types. Cela nous empêche de tester si deux joueurs ont une productivité significativement différente. Ensuite, les coefficients sont difficiles à interpréter. Pour remédier à cela, nous les convertissons en effets marginaux, mais cette transformation ignore le biais introduit par le Ridge. Enfin, nous imposons des restrictions sur la dynamique des coefficients pour avoir suffisamment de variabilité dans nos données : la productivité des joueurs est considérée fixe au cours du temps; la productivité ne dépend pas des interactions avec les coéquipiers; les défenseurs n'ont pas d'effets fixes offensifs et les attaquants n'ont pas d'effets fixes défensifs.

Malgré ces limites, obtenir des résultats qui font autant de sens montre qu'une approche qui combine des effets fixes avec une pénalisation Ridge est une stratégie d'estimation particulièrement adaptée à l'évaluation de la productivité individuelle en équipe. Les effets fixes nous permettent de considérer la productivité dans toutes ses dimensions (y compris celles qui ne sont pas directement observables) et le Ridge nous permet d'identifier la contribution individuelle quand les joueurs jouent souvent avec les mêmes coéquipiers. Ainsi, notre procédure d'estimation pourrait être facilement adaptée à d'autres contextes collectifs que le football. Les seules conditions sont que les travailleurs restent suffisamment longtemps dans l'entreprise (nombre d'observations) et que les compositions d'équipes changent fréquemment (variabilité).

Une fois notre mesure de la productivité obtenue, nous l'utilisons pour tester l'existence de discrimination raciale sur le marché des joueurs. Nous comparons notre mesure à la valeur de marché des joueurs d'après le site allemand Transfermarkt. Les deux mesures sont fortement corrélées, ce qui confirme que notre procédure d'estimation mène à des résultats globalement en accord avec le marché. Toutefois, la tendance est loin d'être parfaite et nous pouvons identifier des joueurs sous-évalués et sur-évalués sur le marché. Nous comparons la proportion de joueurs Noirs dans les deux groupes et nous montrons que ces derniers sont sur-représentés parmi les joueurs sous-évalués. Notre approche est ainsi complémentaire de la littérature existante sur la discrimination dans le sport, qui est basé sur des tests empiriques à l'échelle des équipes (voir en particulier Szymanski (2000)).

Annexe : Projet de statistiques appliquées en cardiologie

Pendant mon doctorat, j'ai également participé à programme de recherche impliquant des cardiologues du CHU de Nancy³ et la Ligue de Football Professionnel⁴. Notre objectif était d'établir le profil cardiaque des footballeurs professionnels, qui est évidemment très différent de celui d'une population "normale".

J'étais en charge de la gestion des données et des analyses statistiques et j'ai coécrit un article sur le profil électrocardiographique des joueurs qui a été publié dans la revue *Archives of Cardiovascular Diseases*⁵. Nous sommes actuellement en train de travailler sur d'autres articles qui devraient être publiés en 2019.

³Service de cardiologie, institut Lorrain du cœur et des vaisseaux Louis-Mathieu, centre hospitalier universitaire de Nancy, 4, rue du Morvan, 54500 Vandœuvre-lès-Nancy, France

⁴Ligue de Football Professionnel, 6 rue Léo Delibes, 75116 Paris

⁵Huttin, O., Selton-Suty, C., Venner, C., Vilain, J. B., Rochecongar, P., & Aliot, E. (2017). Electrocardiographic patterns and long-term training-induced time changes in 2484 elite football players. *Archives of cardiovascular diseases*.

Bibliographie

- Alchian, A. A. & Demsetz, H. (1972), ‘Production, Information Costs, and Economic Organization’, *American Economic Review* **62**(5), 777–795.
- Berri, D. & Schmidt, M. (2010), *Stumbling on wins in basketball*, Pearson Education.
- Betts, J. R. et al. (2011), ‘The economics of tracking in education’, *Handbook of the Economics of Education* **3**(341-381), 4.
- Duru-Bellat, M., Jarousse, J.-P. & Mingat, A. (1993), ‘Les scolarités de la maternelle au lycée: étapes et processus dans la production des inégalités sociales’, *Revue française de sociologie* pp. 43–60.
- Fu, Q., Lu, J. & Pan, Y. (2015), ‘Team Contests with Multiple Pairwise Battles’, *American Economic Review* **105**(7), 2120–40.
- Guyon, N. & Huillery, E. (2016), ‘Biased aspirations and social inequality at school: Evidence from french teenagers’, *Sciences Po Working Paper* .
- Hamilton, B. H., Nickerson, J. A. & Owan, H. (2003), ‘Team incentives and worker heterogeneity: An empirical analysis of the impact of teams on productivity and participation’, *Journal of political Economy* **111**(3), 465–497.
- Ilardi, S. & Barzilai, A. (2008), ‘Adjusted plus-minus ratings: New and improved for 2007-2008’, <http://www.82games.com/ilardi2.htm> .

- Kandel, E. & Lazear, E. P. (1992), 'Peer Pressure and Partnerships', *Journal of Political Economy* **100**(4), 801–817.
- Rosenbaum, D. T. (2004), 'Measuring how nba players help their teams win', *82Games.com* (<http://www.82games.com/comm30.htm>) pp. 4–30.
- Szymanski, S. (2000), 'A market test for discrimination in the english professional soccer leagues', *Journal of political Economy* **108**(3), 590–603.
- Szymanski, S. (2003), 'The assessment: the economics of sport', *Oxford Review of Economic Policy* **19**(4), 467–477.
- Vrignaud, P. (2016), 'L'évolution des intentions d'orientation et du choix professionnel au cours du college : l'impact du genre et de l'origine sociale', *CNESCO report* .