



Discrimination and public policies

Pierre Deschamps

► To cite this version:

Pierre Deschamps. Discrimination and public policies. Economics and Finance. Institut d'études politiques de paris - Sciences Po, 2018. English. NNT : 2018IEPP0016 . tel-03419457

HAL Id: tel-03419457

<https://theses.hal.science/tel-03419457>

Submitted on 14 Dec 2021

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.

INSTITUT D'ÉTUDES POLITIQUES DE PARIS
ECOLE DOCTORALE DE SCIENCES PO
DÉPARTEMENT D'ÉCONOMIE
THÈSE EN ÉCONOMIE

Discrimination and public policies

Pierre Deschamps

SUPERVISED BY ETIENNE WASMER AND JOSE DE SOUSA

COMMITTEE IN CHARGE:

PROFESSOR GHAZALA AZMAT

PROFESSEUR DES UNIVERSITÉS, SCIENCES PO

PROFESSOR DANIEL HAMERMESH, REFEREE

EMERITUS PROFESSOR, ROYAL HOLLOWAY, UNIVERSITY OF LONDON

PROFESSOR NICOLAS JACQUEMET

PROFESSEUR DES UNIVERSITÉS, PARIS 1 PANTHÉON-SORBONNE

PROFESSOR ROLAND RATHELOT, REFEREE

ASSOCIATE PROFESSOR, UNIVERSITY OF WARWICK

PROFESSOR JOSE DE SOUSA, CO-ADVISOR

PROFESSEUR DES UNIVERSITÉS, UNIVERSITÉ PARIS-SUD

PROFESSOR ETIENNE WASMER, CO-ADVISOR

FULL PROFESSOR NYU-ABU DHABI

DEFENDED ON 6TH JULY 2018

Contents

1	INTRODUCTION	15
1.1	Chapter 1: Labour mobility and racial discrimination	18
1.2	Chapter 2: Gender discrimination in Recruitment Committees	23
1.3	Chapter 3: Homophily and Migration	26
2	LABOR MOBILITY AND RACIAL DISCRIMINATION	31
2.1	The Soccer Market	39
2.2	Theoretical Predictions from a Search Model of Discrimination	42
2.3	Empirical Design	43
2.4	Empirical Results	50
2.5	Conclusion	69
	APPENDIX A LABOR MOBILITY AND RACIAL DISCRIMINATION	71
A.1	Data appendix	71
A.2	A Search Model of Discrimination in Soccer	74
A.3	Further Evidence of Discrimination	83
A.4	Testing for Structural Breaks	87
3	GENDER DISCRIMINATION IN RECRUITMENT COMMITTEES	91
3.1	Model	96
3.2	Data	100
3.3	Estimation	107
3.4	Effect of the quota	114
3.5	Discussion	118
3.6	Conclusion	123
	APPENDIX B GENDER DISCRIMINATION IN RECRUITMENT COMMITTEES	125
B.1	Semiparametric estimation	125
B.2	Differences between the Random Utility Model and the Rank-Ordered Logit . .	127
B.3	Probit and heteroscedasticity	128
B.4	Difference in difference	130
B.5	First Stage results	131
B.6	Full effect of the reform	131

B.7	Additional tables	131
4	HOMOPHILY AND MIGRATION	139
4.1	Introduction	139
4.2	Model	145
4.3	Data	154
4.4	Identification	157
4.5	Estimation	158
4.6	What is the effect of homophily?	166
4.7	Conclusion	173
	APPENDIX C HOMOPHILY AND MIGRATION	175
C.1	Figures	175
C.2	Tables	176
C.3	Perfect substitution between white and black workers	176
C.4	Reduced form effect of the instruments on the dependent variable	181
	REFERENCES	199

Acknowledgments

Il y a maintenant 10 ans, je poussai les portes du 27 Rue Saint-Guillaume pour la première fois. J'avais à l'époque un peu moins de barbe et encore moins de certitudes. Je pensais avoir abandonné pour toujours la modélisation et les mathématiques pour l'histoire, les sciences politiques et la littérature, mais ma première rencontre avec l'économie a tout bouleversé. Dès mes premiers cours, j'ai été fasciné par l'idée que l'on puisse modéliser l'ensemble des comportements humains à partir d'un principe simple: l'optimisation.

Il est probable que je n'aurais jamais fait de thèse sans les cours de microéconomie de première année. J'aimerais donc remercier vivement Étienne Wasmer, qui avant d'être mon directeur de thèse, a été mon premier professeur d'économie. Ses conseils avisés et sa perspicacité m'ont toujours été d'une immense aide. En outre, je lui suis entièrement redevable pour le deuxième chapitre de cette thèse, qui est le prolongement d'une étude sur Sciences-Po qu'il m'avait suggéré de mener.

Je souhaite évidemment remercier très chaleureusement José de Sousa, mon professeur puis co-directeur et co-auteur. À l'époque où nous avons commencé à travailler ensemble, la recherche en économie m'était encore floue. Je souhaitais alors réfléchir à la relation entre commerce international et démocratie. José m'a proposé de m'orienter plutôt vers une question bien plus passionnante: l'effet de l'immigration sur le salaire des natifs, à l'image de la célèbre étude de Card sur les *Marielitos*, mais en utilisant l'arrêt Bosman comme choc exogène. Après six années de travail acharné, et maintes révisions, notre collaboration a donné lieu au premier chapitre de cette thèse. J'ai énormément appris de lui sur le plan intellectuel et sur bien d'autres choses encore comme par exemple prendre les commentaires en séminaire de façon constructive (ceux qui me connaissent savent que ce n'était pas gagné!). Sur chaque nouvel article, je sais ce que j'ai appris à ses côtés et ce que je lui dois.

I would like to thank the other members of the jury, Ghazala Azmat, Daniel Hamermesh, Nicolas Jacquemet and Roland Rathelot. They have all intervened at different stages during my graduate years; Roland Rathelot was my econometrics teacher, Daniel Hamermesh gave excellent feedback on an early version of the first paper at a conference in Porto, Nicolas Jacquemet whose work was an inspiration for the second chapter and Ghazala Azmat was always available to discuss my articles, and her advice improved the second and third articles greatly. I'm delighted that they could be part of my jury, and extremely grateful for the comments and suggestions they have made.

Je remercie aussi les professeurs du département d'économie de Sciences Po. J'aimerais en particulier remercier Guy Laroque, qui encadre notre Reading Group et auprès de qui j'ai beaucoup

appris, d'avoir partagé avec nous ses répliques si cinglantes dont il a le secret et d'avoir démystifié le Saint Graal du "Top 5". J'ai aussi pu bénéficier d'un financement du LIEPP pour ma recherche, au sein duquel j'aimerais remercier Emanuele Ferragina, Morgane Laouënan, Christelle Hoteit et Alexandre Biotteau pour leur aide et leurs commentaires. Je remercie aussi Jean-Claude Légal et Sukkoo Kim, mes professeurs en économie avant le master, pour m'avoir fait aimer "la science lugubre".

J'ai également eu la chance de pouvoir passer un semestre au département de Berkeley. Je remercie Emmanuel Saez pour m'y avoir encadré, et aussi David Card, Pat Kline et Jesse Rothstein pour des discussions très enrichissantes, ainsi que Daphné, Caroline, Louise, Léo, Paul, Élodie et Nick pour les repas à l'I-House et les soirées au Missouri Lounge.

Beaucoup de choses ont changé à Sciences-Po en dix ans, mais certaines traditions perdurent. J'espère que dans dix ans je pourrai toujours retrouver Hugues, Ladislav et Nicolas à l'Indiana Place de Clichy, avant de les battre au Risk. Je remercie aussi Adrien, qui je suis sûr continuera sa grande carrière de chercheur en exil au Canada.

Notre master d'économie à Polytechnique a été particulièrement éprouvant, surtout en février lorsqu'il fallait monter les escaliers verglacés qui montent de Lozère vers le plateau de Saclay. Un grand merci à ceux sans qui il aurait été bien plus difficile de terminer les "problem sets": Caroline, Odran, Richard, Jonas.

J'aimerais remercier mes compagnons d'infortunes sur ces 5 dernières années, et dont les nombreux conseils, suggestions, remarques et pique-niques dans les jardins des Tuileries ont grandement enrichi le travail défendu aujourd'hui. Je pense tout particulièrement à Guillaume W., théoricien et co-auteur hors-pair; Valeria, de loin la meilleure titreuse d'article du département et grand vainqueur du concours Anaface; Clara, ma voisine de bureau partie trop tôt pour la banque de France et devenue aficionado de Marvin Gaye; Jean-Louis, le dieu du dancefloor; Odran, dit "El Loco", virtuose sur le terrain de foot et sur R; Pierre G., le plus grand "Rick-Roller" de sa génération; Pierre C, le référent déontologie du département; Etienne, à la bonne humeur contagieuse, qui sait toujours trouver le bon moment pour organiser une pause salutaire; Arthur, qui après 2 ans de lutte acharnée a enfin reconnu mes 'quartiers de bretonnité'; et aussi Charles, pour qui la modélisation n'est qu'un talent parmi d'autres; le backgammon, les échecs, l'organisation de séminaires. Et aussi Aseem, Andreea, Assia, Camilo, Camille, Clément, Dylan, Edoardo, Elisa G., Elisa M., Emanuele, Eve, Guillaume C., Jamil, Jan, Jean-Benoît, Jean-Baptiste, Joanne, Julien, Léa, Lucas, Ludovic (le SL), Mario, Marion, Maxime, Nicolo, Paul, Quentin, Riccardo, Simon, Sophie et Tyler.

Un grand merci aussi à Anne Boring et Vincent Fertey, grâce à qui j'ai pu enseigner au campus du Havre dans d'excellentes conditions.

J'aimerais remercier toute ma famille et belle-famille pour leur soutien constant. Je souhaite remercier tout particulièrement Victor et Louise, qui me sont tout les deux une grande source d'inspiration; Louise dont l'oeil avisé, l'intelligence et la créativité en feront une excellente réal-

isatrice. J'ai hâte de voir ce que le film de mariage va donner; Victor, dont la musique arrive à se faire déhancher et les écossaises et nos grands-parents. Réussir ce grand écart lui assure un brillant avenir dans la musique. J'espère qu'il ne me tiendra pas trop rigueur de l'avoir poussé à faire de l'économie. Et bien sûr mes parents. Difficile de résumer en quelques lignes tout ce que je leur dois, mais outre un soutien moral sans faille et souvent nécessaire, j'aimerais les remercier pour ce qu'ils m'ont enseigné: la curiosité, le goût de la lecture, et sans doute le plus important, mon amour pour le foot et les Bolton Wanderers.

Mon histoire d'amour pour l'économie n'est rien à côté de celle que je partage avec Émilie, qui est devenue ma femme cet été. Et comme elle reste encore sceptique sur nos théories, je ne ferai pas de blague typique des économistes du style "le partage du surplus de notre appariement respecte la condition d'Hosios". Notre complicité m'amène à toujours donner le meilleur de moi-même et je mesure chaque jour la chance que j'ai de l'avoir à mes côtés.

List of Figures

1.1	Added-variable plot	22
1.2	Rank-ordered logit on gender of candidates by year	25
1.3	Low-skilled workers	27
1.4	High-skilled workers	27
1.5	Low-skilled workers	27
1.6	High-skilled workers	27
2.1	Share of foreigners in European clubs over time (1981-2008)	34
2.2	Club's Wage Bills	40
2.3	Average Wage Bill and Team Success	41
2.4	Relative Turnover of Black English Players	58
A.1	Average Wage Bill and Team's Quality	74
A.2	Pre-Bosman	83
A.3	Post-Bosman	83
3.1	Rank-ordered logit on gender of candidates by year	120

3.2	Effect of the reform on the ranks of women by discipline. Disciplines most af-	
	ected by the reform are ordered from left to right.	121
B.1	Semi-parametric estimation of the relationship of interest	126
4.1	Low-skilled workers	141
4.2	High-skilled workers	141
4.3	Black workers/Total population per MSA in 1980	143
4.4	Black workers/Total population per MSA in 2000	144
4.5	Low-skilled workers	145
4.6	High-skilled workers	145
4.7	Difference between the true and counterfactual populations with no homophily,	
	as a function of the initial shares of black and white workers.	169
C.1	Low-skilled workers	177
C.2	High-skilled workers	177
C.3	Low-skilled workers	178
C.4	High-skilled workers	178

List of Tables

2.1	Individual Differences in Means: Black vs White English Players	47
2.2	Club Differences in Means	47
2.3	Discrimination Market-test: League Success and Long Span	52
2.4	Discrimination Market-test: Match Success and Short Span	54
2.5	<i>Post-Bosman</i> : Are non-EU Black Players Discriminated? (1996-2008)	56
2.6	Discrimination Market-test: League Success and Long Span - IV and GMM	61
2.7	Share of Black Players and Corporate Control	65
2.8	Discrimination Market-test: Effect of Corporate Control	66
2.9	Discrimination Market-test: Heterogeneity of clubs	68
A.1	Discrimination Market-test: Club's League Success in France and Belgium	85
A.2	Discrimination Market-test: Match Success and Long Span - Within and Within-IV	86
A.3	First Stages: League Success and Long Span	87
A.4	First-Stages: Match Success and Long span	88
A.5	Structural Break: Long Span	89

A.6	Structural Break: Short Span	89
3.1	Some descriptive statistics	102
3.2	Recruitments and gender	102
3.3	Descriptive statistics on hires and applicants	103
3.4	Descriptive statistics on dyads	104
3.5	Descriptive statistics on academic connections	105
3.6	Effect of the reform on the mean share of women jurors and number of female presidents	106
3.7	Effect of the reform on the mean h-index of jurors	107
3.8	Correlation of the gender of jurors on the probability of women being first-ranked	109
3.9	Correlation of the gender of jurors and the rank of candidates	110
3.10	Correlation of the gender of jurors and the probability that a woman is better ranked within a dyad	112
3.11	Correlation of the gender of jurors and the probability that a woman is better ranked within a dyad. Split by subject	113
3.12	Regression of the gender composition of applicant pools on the gender composi- tion of the recruitment committee	114
3.13	Effect of the reform on the candidate pool	115
3.14	Effect of the quota on the rank of female candidates	116

3.15	IV estimate of the effect of the increase in women jurors on the probability that a woman is better ranked within a dyad	117
3.16	Rank-Ordered Logit using the quota: top 3 ranks only	119
3.17	Effect of the reform by gender of the jury president	122
B.1	Power of estimation methods	128
B.2	Estimates from simulations	129
B.3	Estimates from simulations: large parameters	129
B.4	Probability that W is more highly ranked	130
B.5	Probit results, Difference-in-Difference	131
B.6	OLS results, Difference-in-Difference	132
B.7	First stage: IV	133
B.8	Rank-Ordered Logit using the quota	134
B.9	Rank-Ordered Logit using the quota	135
B.10	IV Probit on gender of first-ranked candidate	136
B.11	Rank-Ordered Logit using the quota: Other publication measures	137
4.1	Amenity index	156
4.2	Are the Bartik shocks good instruments?	159
4.3	Labour demand parameters	161
4.4	Transfers and Housing Supply	162

4.5	Sensitivity to wages and rents	163
4.6	Homophily and amenities	164
4.7	Network amenities	164
4.8	Racial wage gap	168
4.9	Racial real wage gap	168
4.10	Response to shocks of the share of white workers in a city.	172
4.11	Response to shocks of the share of black workers in a city.	173
C.1	% of residents who are black in 1980	179
C.2	% of residents who are black in 2000	180
C.3	Estimation of elasticities of substitution	181
C.4	Are the Bartik shocks good instruments?	182

Racial discrimination pervades every aspect of a society in which it is found. It is found above all in attitudes of both races, but also in social relations, in intermarriage, in residential location, and, frequently, in legal barriers. It is also found in levels of economic accomplishment; that is, income, wages, prices paid, and credit extended. This economic dimension hardly appears in general treatments of economics, outside of the specialized literature devoted to it. Nevertheless, it is important not only in itself but as a test of standard theories

Kenneth Arrow

1

Introduction

The quote above, from Arrow's 1998 JEP article, highlights the relevance of discrimination to economic analysis. Arrow pointed out that as overt discrimination has disappeared, it has become a bigger empirical challenge to identify whether certain groups were still facing discrimination. In this dissertation I present three articles that provide new empirical evidence on discrimination and I explain how economic analysis can help inform policy choices in the fight against discrimination.

The first two chapters look at the effect of two different public policies on discrimination, and the last chapter considers how racial wage gaps could be affected by a decrease in the strength of own-group preferences.

Should economists give policy advice on discrimination? Economists have long advised governments, especially in macroeconomics. Our influence is likely to wane in the future, however. The Great Recession of 2009 was not only a disaster for the economy, it also started another crisis for the economics profession as, rightly or wrongly, the entire discipline was chastised for its failure to predict the recession.¹ It seems as though large swathes of voters have taken these critiques on board, and may no longer trust our advice. For instance, in the run-up to the 2016 Brexit vote, the advice of economists was largely disregarded by voters, with the justice secretary Michael Gove (in-?)famously stating that Britons “have had enough of experts”.²

Not only is our macroeconomic advice ignored, but it also appears as though the issues that matter for voters are shifting. Whereas voters’ judgment of how the economy was faring seemed to govern the outcomes of previous elections (see for instance, the famous slogan for Clinton’s 2002 campaign, “it’s the economy, stupid”), during the 2016 election, not only was the consensus of economists ignored³, but concerns about the economy did not seem to be the main driver of

¹See Paul Krugman’s November 6th 2009 op-ed.

²Interview with Sky News, June 3rd 2016. For more substantial evidence, see Sapienza & Zingales (2013), summarized in this article from the Economist, 12th Jan 2013. The US public is much more circumspect than economists on the positive effect of stimulus for instance. Haldane & Turrell (2018), accessible here show that almost no demographic group in the UK trusts economists, not even those that voted remain.

³For instance, the following petition of 794 economists against Trump went unheeded: Economists against Trump.

votes. Rather, identity, race, sexism and immigration seemed to play a large role in determining voting behaviour.¹ It thus becomes crucial that economists are also able to give credible policy advice on these subjects.

Fortunately, the discipline has already started to adjust, with an increasing reliance on empirical work and a move away from purely theoretical models, though this trend predates the Great Recession. This shift also has important consequences on the type of counsel and policy advice economists can give, since macroeconomic models are typically fraught with identification problems. Reliance on empirics may thus change the type of advice that economists can give, away from publically-discredited pronouncements on events that cannot be empirically evaluated, and towards empirically-driven counsel based on narrower policy questions.

In all three chapters of this dissertation, I try to see whether discrimination and own-group preferences exist, in different contexts, and what kind of public policies could mitigate or balance the negative effect of these preferences. In my first chapter, I rely on a famous ruling that changed the monopsony power of firms to see whether, as predicted by Becker, market failures have an impact on wage discrimination. I find that as monopsony power decreases, firms are no longer able to act on their prejudice, and wage discrimination disappears. This result shows that labour market context is essential in evaluating public policies, and that prejudice need not necessarily translate into wage discrimination.

In the second chapter, I analyse the effect of another public policy, a 2015 reform that imposed

¹For instance, see Schaffner et al. (2018) or Becker et al. (2016).

gender quotas in academic recruitment committees. The reasoning of the policymakers was that increasing the share of women evaluators would improve the outcomes for women. I find the opposite instead; women are ranked worse by hiring committees after the reform. However, this result does not show that women discriminate against women. There is some evidence that this result is caused by the reaction of male jurors to the reform, since the negative effect of the reform is found only in committees that are helmed by male jury presidents. This chapter shows that it is necessary to evaluate public policies, lest reforms that are well-meaning in intention turn out to cause more harm than good.

In the third and final chapter, I show two stylised facts: When cities decline, they tend to become more black, and black residents are disproportionately located in cities that pay low wages. One explanation for this could be that living in cities with a larger share of black residents is a positive amenity for black workers. I try to see whether workers have preferences for living in cities that have a larger share of co-ethnic residents, when controlling for wages, rents, transfers and network amenities. I find that these preferences are significant, and then try to see what share of the wage gap these preferences, and the imperfect sorting they imply, could explain.

1.1 CHAPTER 1: LABOUR MOBILITY AND RACIAL DISCRIMINATION

Most political solutions on cultural issues focus on changing preferences or attitudes. By contrast, in economics we mostly take preferences as given, as emphasised by Stigler and Becker in

“De gustibus non est disputandum”. Instead of hoping that individuals who discriminate will change their ways, economists focus on how to minimise the effect of their behaviour. This idea is epitomised by the founding text in the study of discrimination in economics, Becker (1957). In a perfectly competitive economy, Becker argues, discrimination can not be an equilibrium outcome. If some firm-owners refuse to hire employees of a particular group, then the equilibrium wage for these employees will be lower than members of an equally productive, non-discriminated group. The employers of these workers will make extra profits, and should drive or buy out the less productive firms in the long run. In a competitive market, the no-arbitrage condition should sound a death-knell for discriminators.

Empirical evidence, however, shows that discrimination is still present, a fact that escaped neither Becker, nor Arrow (1972). The solution in reconciling empirics with theory is to move away from perfect competition and consider market failures that can allow prejudiced employers to survive. One strand of research has focused on information asymmetries, leading to what has been named “statistical” discrimination in the literature. In this case, employees that have identical expected productivities, but larger uncertainty relative to other groups, are underpaid by their risk-averse employers. This approach has challenged Becker’s, which considered employers who had a psychological disutility or “taste” for discrimination, but whether one has superseded the other in terms of empirical evidence is still up for debate¹. The two market failures held respon-

¹See Guryan & Charles (2013) for a review of the debate on which of the two explanations is the most likely

sible for the presence of taste discrimination are monopsony¹ or monopoly power.

Empirically, there have been many studies that look at the effect of monopoly power on discrimination (see e.g. Ashenfelter & Hannan, 1986), but few that try to gauge the effect of monopsony power. In my first chapter of this dissertation, written with José de Sousa, I contribute to this literature by looking at how changes in the monopsony power of firms can affect wage discrimination. Theoretically, search frictions have been considered as a source of monopsony power, since different firm owners do not have to compete against each other in hiring. We look at the effect of the Bosman Ruling in football, which limited the monopsony power of firms on wage discrimination against black players in the English Premier League.

Before the Bosman Ruling, the European market for footballers was hindered by two different rules: The first was a quota on foreign players per club, which limited intra-European mobility and thus the number of clubs that could potentially bid up a player's wage; the second was a transfer fee, similar to a non-compete clause, that had to be paid by clubs that wanted to hire a player, even if the player's contract had expired. Jean-Marc Bosman, after whom the Bosman ruling is named, fell foul of this second rule, when his club, Liège, only offered 25% of his old wages once his contract expired. Having found another club which could hire him for a higher wage, Liège thwarted this transfer by setting a transfer fee that was too high for the other club to pay. The Bosman ruling removed quotas for EU players, and players can now leave freely for another

¹The term, coined by Joan Robinson, describes markets with only one buyer, such as a labour market with only one hiring firm.

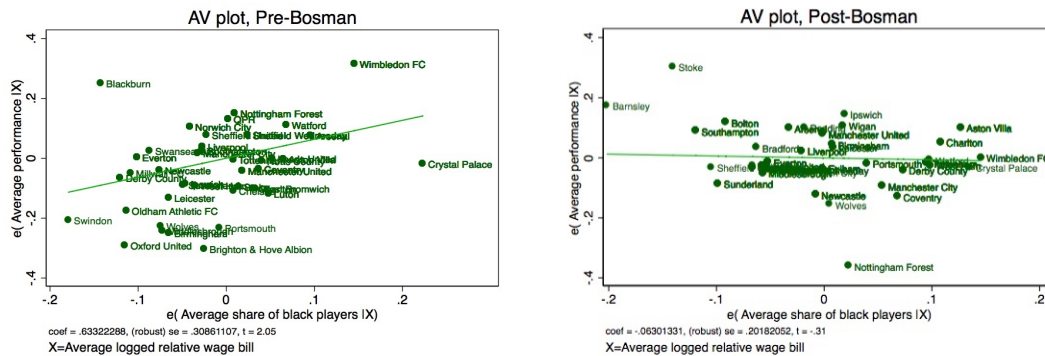
club once their contracts have expired, ending what seemed to be a textbook case of monopsony power.

We use this ruling change as a quasi-experimental shock to mobility constraints and monopsony power, and analyse the effect of the shock on wage discrimination. In order to measure wage discrimination in football, we rely on the following idea: If every player's productivity is rewarded fairly, then the club's wage bill should perfectly explain the team's performance. In fact a regression of average performance on average wage bill has an R^2 of 0.9. If we add another variable in the regression, the share of black players that are hired by the club, this variable should be insignificant. If it is not, it implies that for the same wage bill, a club can get a higher ranking by hiring more black players, and hence that black players are not compensated at the same wage rate as white players for their performance. This method has been called the market-test by Szymanski (2000), and solves two important problems that we may have with Mincer-style wage regressions at the individual level. The first is that there is no clear measure of individual productivity in a team sport such as football, and the second is that there is likely to be omitted variable biases that will be mitigated when doing the analysis at the team level.

In line with a search and matching model that we use to illustrate's Becker's ideas, we find wage discrimination before the Bosman ruling, and none afterwards. This result holds over short and long time-windows. Our results can be illustrated in the graph below, that plots the effect of having more black footballers in a team, when controlling for the club's wage bill,. We can see

that having more black players increases the ranking before 1995, and has no effect after the ruling.

Figure 1.1: Added-variable plot



We provide additional evidence that this decrease is due to the ruling by applying the market-test to non-EU players in England and other countries, since these players still face restrictive work permit rules, and accordingly we find evidence of wage discrimination against these players. We try to rule out alternative explanations for our finding, such as the rise of corporate control. Teams under corporate control are not more likely to hire black players before Bosman.

This chapter provides us with concrete guidance on how to reduce wage discrimination. Though acting to reduce prejudice is also an essential policy goal, it is one where we have little idea what works. If we cannot solve the root causes of prejudice however, we can mitigate its effects by making labour markets more competitive.

1.2 CHAPTER 2: GENDER DISCRIMINATION IN RECRUITMENT COMMITTEES

Assessing what works in reducing discrimination is important, since if we misidentify the true causes of discrimination, policies that seek to redress imbalances may end up causing more harm than good. For instance, one subject to which a lot of ink has been devoted is the under-representation of women in academia, especially in STEM. That women are under-represented in academia is not in doubt: EU data shows that women make up only 21% of full professors, despite making up between 40-60% of PHD students.¹

One policy to redress this imbalance could be to increase female representation in academic recruitment or promotion committees. Implicitly, this supposes that one reason for the lack of female professors is bias that male professors have in recruitment against female professors. Yet, this is one out of many possible explanations for this structural imbalance. For instance, the unequal sharing of home production, especially in time spent taking care of children, which diminishes the time that women can work on their research. Another hypothesis is that women have different attitudes to competition. Imposing gender quotas in recruitment committees may succeed in fulfilling some alternative requirements, but will only improve the career prospects of women if discrimination is the main cause of under-representation, and not any of the alternative hypotheses presented above. It may also increase the administrative workload of female professors relative to their male colleagues.

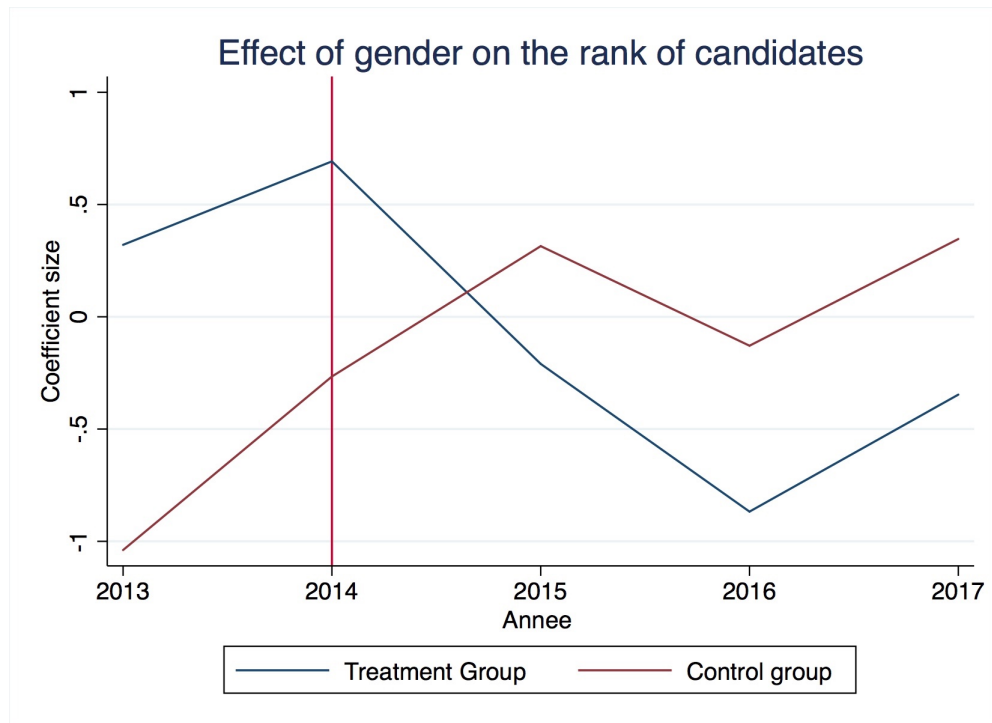
¹Data from the SHE figures

In my second chapter, I try to see whether a French reform that imposed gender quotas in recruitment committees had a positive effect on the outcomes of women. This quota fixed the minimum gender representation of both genders at 40%. All fields were not equally affected by the reform however, since fields with more male professors (like Political Science or Economics) were more likely to be affected by the reform than fields with a more equal representation (such as Biology or Languages). I analyse ranked data from academic recruitment committees in France, combined with data on citations and publications, as well as data on the theses of candidates scraped from the Internet, and try to estimate the effect of having more women in recruitment committees on the rank of women by the committee, using the differential impact of the reform by field as an exogenous source of variation. I find a very large, significant and negative effect of the feminisation of committees on the rank of women, in accordance with another paper by Bagues et al. (2017). My paper goes one step further by directly estimating the effect of a gender quota, whereas their article identifies gender bias from random assignment of committee members to candidates. I illustrate the results in Figure 1.2 below, which shows the effect of gender on rank before and after the reform, within the treatment and control group. It is easily seen that the effect of gender on rank worsens after the reform in the treatment group, but improves in the control group.

According to the estimates, increasing the number of women by 10% leads to a decrease in the probability that a woman is higher ranked akin to a 25 unit drop in h-index. The large size

of these estimates imply that it might not be the case that men are more likely to rank women higher, but that they may also react to the quotas by giving lower ranks to women candidates. I investigate this concern by splitting the sample into committees with a male president and those with a female president. Jury presidents are very powerful, since they have a deciding vote in case of a tie, and also have a large role in choosing the other members of the committee. I find a negative effect of the reform only for committees that have a male president, which could be an indication that it is male jurors that react to the reform by downvoting women, rather than women who have exophilic preferences.

Figure 1.2: Rank-ordered logit on gender of candidates by year



Nevertheless, the result of this study indicates that imposing gender quotas in recruitment committees may not be a good solution to solve the problem of female under-representation. This chapter shows that evaluating public policies is paramount; many well-meaning policies may end up harming the ones they are supposed to help.

1.3 CHAPTER 3: HOMOPHILY AND MIGRATION

Though the expert consensus stresses the positive economic effects of immigration,¹ recent election results and surveys suggest that a large share of voters are unconvinced by the evidence put forward. In fact, there may be benefits and costs from immigration above and beyond wages, but these costs are difficult to quantify, and thus to estimate. For instance, if preferences for co-ethnic residents exist, then the costs of immigration for natives may be underestimated. Ignoring these costs may lead us to underestimate the effect of immigration on natives, and thus the amount of redistribution that could be necessary to compensate natives for their losses.

In my third chapter, I try to see whether preferences for co-ethnic residents have an effect on segregation and location choices at the city level, while controlling for most other determinants of city choice. I present two stylised facts that seem to support the hypothesis that own-group preferences matter for location choice. In Figure 1.3, I plot the share of black workers in a city against the residual wages paid to black workers. I find a negative relationship, implying that black

¹For instance, this economists' survey: <http://www.igmchicago.org/surveys/low-skilled-immigrants>. In the economic literature there are contrasting results (see Peri et al. (2015) or Borjas & Monras, 2017), but a recent article by Dustmann et al. (2016) shows that many of these disagreements can be explained by a difference in the parameters being estimated, between the negative relative effects of immigration and the positive total effects.

workers are located in cities that pay them low wages. In Figure 1.5, I compare the decadal change in the share of black workers with the decadal change in total population. Again the relationship is negative; as cities decline, the share of black workers increases. In both cases, these relationships hold for both low- and high-skilled workers. One answer to why black workers locate in declining cities that pay them low wages could be that being in cities with large shares of black workers has an amenity value for them, the size of which this paper tries to estimate.

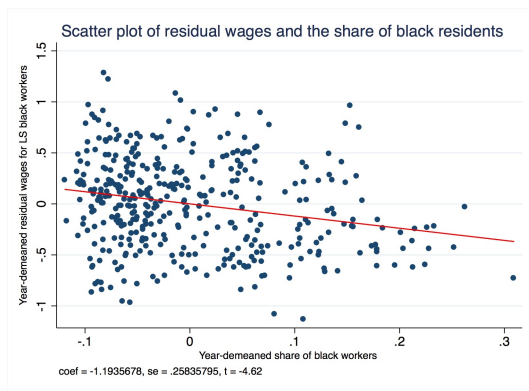


Figure 1.3: Low-skilled workers

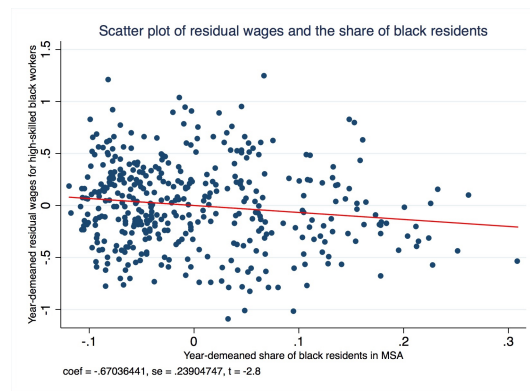


Figure 1.4: High-skilled workers

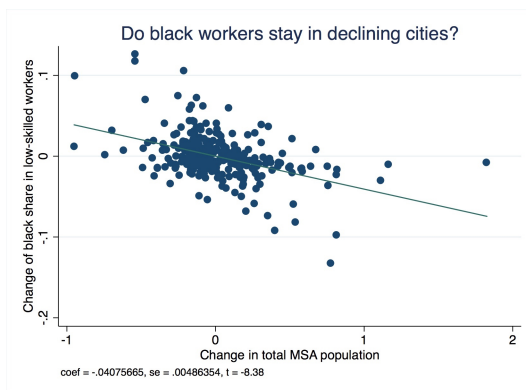


Figure 1.5: Low-skilled workers

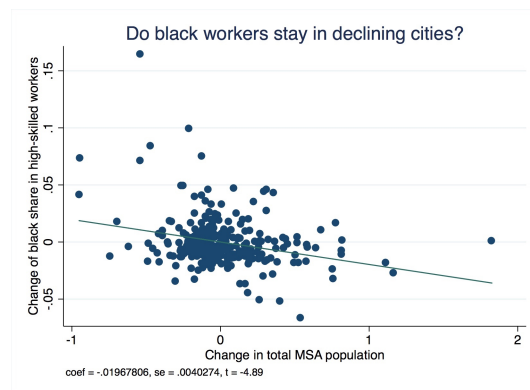


Figure 1.6: High-skilled workers

Although there is a large literature on these preferences, both theoretical (Schelling, 1971) and empirical (for instance Card et al., 2008), these articles focus on preferences at the neighbourhood level. I look at the effect of these preferences at the city level instead, where segregation is also high and persistent. This approach is relevant since choices within-city could be constrained by redlining, which would make an explanation of segregation based on preferences difficult to support.

In practice, estimating own-group preferences is extremely complicated since shares of white or black residents could be correlated with any number of other determinants of city choice, such as wages or rents. I thus estimate an n -equation spatial equilibrium model that takes into account wages, education, rents, transfers, networks, amenities and own-group preferences. All equations in the model are estimated simultaneously through GMM, using US census data from 1980 to 2000. A second problem is that most of the variables are simultaneously determined by supply and demand. Lower rents may induce more workers to move to a city, but as these workers arrive rents will increase. A regression of rents on housing demand may not be informative. To solve this problem, all variables in the model are instrumented by Bartik shocks, interacted with housing supply parameters and hispanic and asian immigration. I find that preferences for living in cities with a higher share of co-ethnic residents are significant, and comparable in magnitude to preferences for real wages.

In the final part of the paper, I use simulations to test how own-group preferences amplify the

effect of labour demand shocks on segregation. Using predicted variation in wages from Bartik shocks, I find that white workers are more reactive to skill-neutral Bartik shocks. This initial response is then significantly amplified by own-group preferences, increasing segregation. In a second simulation, I try to see what a world without own-group preferences would look like, and in particular its effect on the racial wage gap. I find that reducing spatial misallocation decreases the racial wage gap, especially once rents are taken into account, but this decrease is very small in magnitude, since the within-city wage gap is bigger than the between-city one.

Despite the model's complexity, there are still two important dimensions that need to be taken into account to estimate these own-group preferences with greater accuracy. The first is heterogeneity in taste, since it is likely that own-group preferences differ among individuals, and this could have important consequences for aggregate sorting. In this chapter, there is only taste heterogeneity by skill. The second is that the model only integrates between city choice and not within city. A richer model could integrate both these dimensions, within a nested logit framework for instance. I intend to pursue these leads in future work.

2

Labor Mobility and Racial Discrimination

We examine the effect of labor mobility constraints on racial wage discrimination. In a competitive market, taste-based discrimination cannot survive in equilibrium, since firms that discriminate make lower profits than those who do not, and are consequently driven out of the market (Becker, 1957). Racial wage gaps are thus explained through some kind of market failure: imperfect information leading to statistical discrimination, or market power either on the product

market or on the labor market. In this paper, we estimate the effect of a change in labor mobility on wage discrimination in a monopsonistic labor market.

Our theory builds on search models with discrimination (Black, 1995, Bowlus & Eckstein, 2002). We show that wage gaps due to taste-based discrimination could persist in a monopsonistic labor market. If job-to-job transition is constrained, and employers can set wages to some degree, then employers exhibiting racial biases can persist. Indeed, because the distribution of outside wages sets the outside wage of workers, the presence of taste-based discrimination in even a few potential employers can generate lower wages for minorities at all firms. In contrast, when constraints on job-to-job transition are low, we show that the monopsony power of prejudiced firms is low and discrimination disappears.

Our contribution is to bring the above predictions to the data. We estimate the effect of labor mobility constraints on racial wage discrimination using data on soccer players by ethnicity in the English Premier League. This monopsonistic labor market offers various advantages detailed below. In particular, we exploit an exogenous mobility shock: the Bosman ruling by the European Court of Justice in 1995 lifted restrictions on player mobility in the European soccer labor market. Using a simple market test for the presence of discrimination, we show that discrimination existed before the Bosman ruling but disappeared afterwards, as theory would predict.

The soccer market offers three important advantages to study the link between discrimination and labor mobility constraints. First, following the Bosman ruling, we observe large variation

in labor mobility within the European Union (EU) soccer market. This variation is expected because the pre-Bosman era had two important restrictions on job-to-job mobility: (1) transfer fees needed to be paid for out-of-contract players¹ and (2) the number of foreigners per team was restricted to 5 by a quota system. The ruling removed the quota barriers for EU nationals and the obligation to pay a fee for out-of-contract players.² Figure 2.1 illustrates the intensified mobility of the European soccer market in the wake of the Bosman ruling. After 1995, the ratio of foreigners in squads skyrocketed. Players can now field offers from potentially any country in the EU. This policy change creates a compelling quasi-experimental variation to identify a plausibly causal effect of mobility constraints on racial discrimination.

Second, extensive data on the career paths of professional players can be gathered for most countries over long time periods (Kleven et al., 2013). We match this individual data with information on the skin-color of players. Third, the soccer market offers a simple test for racial discrimination in salary setting (Szymanski, 2000), which can be used to test the effect of our mobility shock on discrimination. Soccer is considered an efficient market,³ where team success

¹This restriction was similar to the 'reserve clause', frequent in North American sports, that bound players to the teams that held their contracts, irrespective of whether their contract had expired. This meant the player was not free to enter into another contract with another team. The reserve clause was abolished in baseball in 1975. Cymrot (1985) looks at different categories of players in baseball after the reserve clause, those who could become free agents, and those who could not (because they had played professional baseball less than 6 years), a test similar to our test on non-EU players. He finds no discrimination for players under free agency, and some evidence of discrimination for those who were not yet eligible for free agency.

²Specifically, in June 1990, at the end of his contract, the Belgian player Jean-Marc Bosman refused Liege's offer of a contract extension at only 25% of his last wage, and accepted a contract from the French club Dunkirk instead. Because Dunkirk refused to meet the transfer fee demand, Liege refused to let him go. Bosman decided to take his case to the courts and won. As a result, the European Commission applied European law on worker mobility to the soccer labor market.

³Unlike professional sport labor markets in the US, there are no collective bargaining agreements, salary caps, or

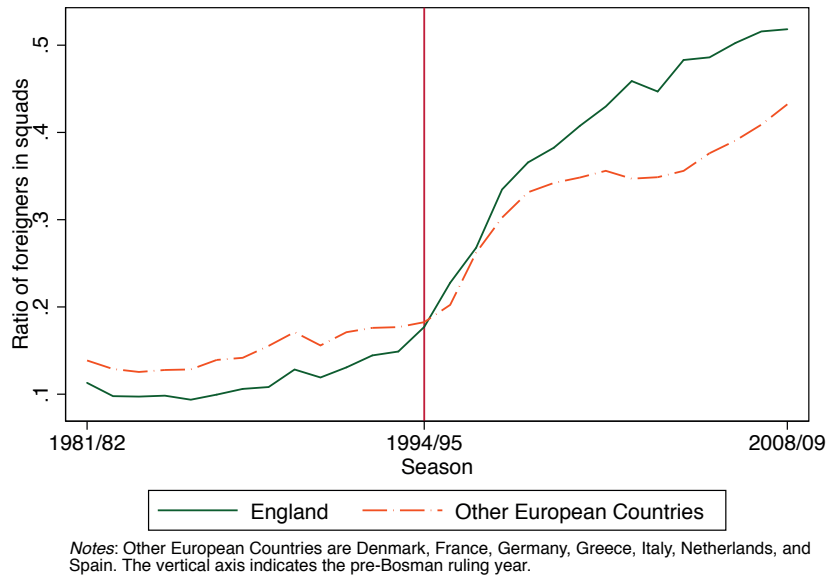


Figure 2.1: Share of foreigners in European clubs over time (1981-2008)

depends on the wage bills of clubs that reward the players' talent. Discrimination can then be said to exist if, for a given wage bill, clubs fielding an above-average proportion of black players systematically outperform clubs with a below-average proportion of black players. This implies that black players are being paid less than their talent would warrant. Szymanski finds evidence of discrimination while performing his test on a panel dataset of professional English clubs *before* the Bosman ruling between 1978 and 1993. Relative to this work, we exploit the Bosman ruling shock and provide empirical evidence that discrimination disappears with low mobility constraints.

The evidence that wage discrimination disappeared in the post-Bosman period is robust to the use of (1) various estimators (OLS, within, IV and GMM), (2) alternative dependent variables

draft picks to maintain a competitive balance between teams.

(either league or match success), and (3) different groups of players according to their ethnicity or citizenship. We show that the decrease in discrimination appears to be extremely fast: we find evidence of wage discrimination against black English players in the 5 years before the ruling, but not in the 5 years afterwards.

Our main result compares black to white English players and shows that wage discrimination against black English players disappeared after the Bosman ruling. However, if prejudiced attitudes persist we should observe post-Bosman discrimination against Black players who do not benefit from the Bosman ruling. In particular, non-EU players, participating in the Premier League, still face high mobility constraints compared to EU players. Accordingly, we find evidence that *black non-EU* players face some wage discrimination in the post-Bosman era. By contrast no discrimination is found against white non-EU players.¹

A natural concern when estimating the impact of a quasi-experimental design is the occurrence of events potentially contemporaneous to our shock, the Bosman ruling. First, Palacios-Huerta (2014)'s intuition is that the emergence of a market for corporate control of English professional clubs has increased the competitiveness of English soccer, which would have driven discriminating firms out of the market. However, some issues regarding the impact of this insightful hypothesis on discrimination in the Premier League could be raised. First, the emergence of

¹There is also no evidence of post-Bosman discrimination against black and white EU players. Our comparison mimics a difference in differences strategy by comparing black to white English players before and after the ruling to non-EU players which still face high mobility constraints. However, a limitation is that there are few (black) non-EU players in the Premier League before the Bosman ruling.

a market for corporate control of English clubs started in the early 1980s, i.e., ten to fifteen years *before* the Bosman ruling. Then, a club experiencing very poor performance simply moves down from one division to another and is not necessarily driven out of the market.¹ Relegation can lead to promotion the following season without disciplining the club's discriminatory behavior. Finally, there is no clear evidence that the market for corporate control has increased the competitiveness of English soccer. Only a third of the Premier League clubs during the post-Bosman period have been listed on the stock market, and few of the best ones. Despite these concerns, we directly test for the corporate control effect and find no evidence that clubs on the stock market discriminate less than privately owned-clubs. Clubs under corporate control did not hire more black players. Our result is that the labor market became more competitive primarily because of relaxed constraints on mobility following the Bosman ruling.

Second, what we are measuring may not truly be the effect of the Bosman ruling but instead a general change in attitudes towards racism and racial discrimination. The data reject, however, the idea of a general and dramatic reduction in prejudiced views since 1995. Sport is not immune against racism, as shown by recent episodes of racial bias in basketball courts (Price & Wolfers, 2010), baseball (Parsons et al., 2011) and soccer fields. Racist incidents in soccer, whether from fellow players, owners, managers or supporters continue to make the headlines of English newspapers today.² These incidents suggest first that racist attitudes are still present at all levels of

¹Note also that some firms may operate under a soft budget constraint (see Andreff, 2014) relaxing the impact of a very poor performance.

²In an online appendix, we show strong evidence of racism in English soccer before and after the Bosman ruling.

English soccer, and, second, that the decrease in discrimination is more likely to be caused by a decrease in job-to-job mobility constraints, in the wake of the Bosman ruling, rather than a dramatic change in attitudes or prejudiced views in 1995. Moreover, the post-Bosman persistence of wage discrimination against black non-EU players and the evidence that black players change clubs relatively more often than white after Bosman suggest persistent problems of prejudice in some clubs.

A different objection to our findings could be related to the potential mismeasurement of wage bills. We identify discrimination through the simple idea that if black players' talent is underpaid, the overall talent of the team is not fully accounted for by the club's wage bill and team success is partly explained by the share of black players. One might be worried that the extent of discrimination could potentially be biased in case of systematic mismeasurement of wage bills in prejudiced clubs. Even if such a systematic pattern is unlikely, we use an instrumental variable approach, which confirms that racial wage discrimination disappeared after the Bosman ruling.

Even though there is a large theoretical literature on how labor market frictions enable firms to discriminate against their employees, there is little empirical work on this question.¹ Exceptions are Biddle & Hamermesh (2013) and Baert *et al.* (2014). They show that employers discriminate

As a recent example, in August 2014, it came to light that the Cardiff City manager Malky Mackay shared racist e-mails and texts with the director of soccer in charge of transfers, Iain Moody. The Daily Mail reports that on August 16, 2012, a list of players proposed by a French agent is forwarded, stating to Mackay that "he needs to rename his agency the All Blacks." A separate text in reference to a list of French players states that "Not many white faces amongst that lot but worth considering."

¹In contrast, there is some empirical evidence that intensified product market competition reduces race-based differentials caused by prejudice, see e.g., Ashenfelter & Hannan (1986) or Levine *et al.* (2008).

less in labor markets with a small number of job seekers relative to vacancies.¹ With regards to these articles, our mobility shock, by offering more job opportunities, can be interpreted as an exogenous decrease in the ratio of employed job seekers to job offers. This situation discourages employers from indulging in discriminatory tendencies.

This paper also contributes to the literature on the effects of labor mobility constraints on wages. For instance, Naidu (2010) and Naidu *et al.* (2016) document that wages rise when labor mobility constraints decrease.² We show that decreasing these frictions also reduces racial wage discrimination.

The results could be important for public policy. It is indeed necessary to understand the root causes of discrimination before we can craft policies to diminish it (List & Rasul, 2011). Removing constraints on worker mobility, such as quotas, work permits, or restrictive contracting rules,³ may improve the capacity of workers to move from prejudiced to unprejudiced firms and reduce discrimination. When job-to-job mobility is constrained, a firm is able to act on its prejudice

¹Biddle & Hamermesh (2013) establish that gender discrimination is lower when the ratio of job seekers to vacancies decreases. The evidence for racial discrimination is less conclusive, however. Baert *et al.* (2014) sent resumes to firms in industries with different labor market tightness. In sectors with few available workers and a large number of vacancies, the difference in call-back rates between resumes with Flemish-sounding names and those with Turkish-sounding names was almost zero, whereas it was significantly lower in “loose” labor markets.

²Naidu (2010) and Naidu *et al.* (2016) study the effect of lifting mobility constraints on the wages of black Americans in the post-Bellum south and migrant workers in the United Arab Emirates, respectively.

³Examples of restrictive contracting rules are the use of non-compete rules or tacit agreements between firms to restrain labor competition. One recent example is the “Techtopus” case, where major Silicon Valley firms such as Google, Apple or Intel stand accused of passing tacit agreements not to hire each other’s employees (see NY Times, April 24th 2014). Another worrying development is the rise of non-compete clauses in all types of jobs, such as summer camp counselors or hair stylists (see NY Times, June 8th 2014). Alan Krueger describes non-competes rules as part of a ‘rigged’ labor market in which employers “undoubtedly restrict worker mobility”. According to Starr *et al.* (2015) 18% of American workers are currently constrained by non-compete clauses, while nearly 40% have signed non-compete agreements at some point in their careers.

because of the low cost of doing so.

The paper is organized as follows. Section 2.1 describes the context of our analysis and the competitive soccer market. Section 4.2 outlines what our search model with prejudice would predict about the effects of labor mobility constraints on discrimination. Section 2.3 presents our identification strategy and the specifications of the market test for discrimination. Our empirical results on discrimination in the English soccer league are presented in section 2.4. The most important result is that discrimination disappears when constraints on worker mobility are lowered. Section 3.6 concludes.

2.1 THE SOCCER MARKET

We have already discussed the three important advantages offered by the soccer market to the study of labor mobility and discrimination: (1) an exogenous change in mobility constraints following the Bosman ruling, (2) an extensive collection of individual data on the career paths of professional soccer players,¹ that can be matched with skin-color information, and (3) a simple test for racial discrimination in salary setting (Szymanski, 2000).

We document four facts about the soccer market that guide our empirical and theoretical analysis.

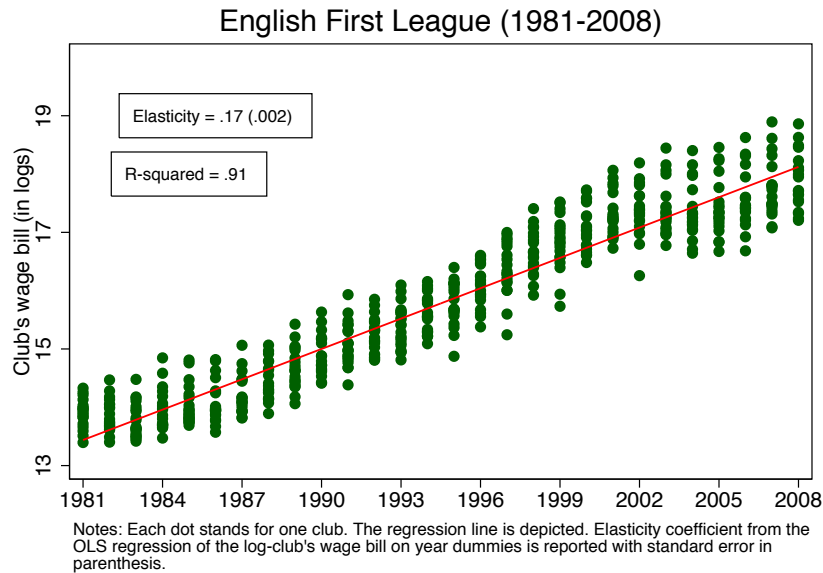
Fact 1. League competition is hierarchical. Each year, approximately 20 teams participate in the English first league. Competition is focused on league rankings without play-offs. Teams are

¹See Appendix (A.1) for further information on variable definitions and data sources.

ranked by total points, then goal difference, and then goals scored. At the end of each season the worst-performing teams swap places with the highest-ranked teams in the second league. There are no collective bargaining agreements, salary caps, or draft picks to maintain a competitive balance between teams.

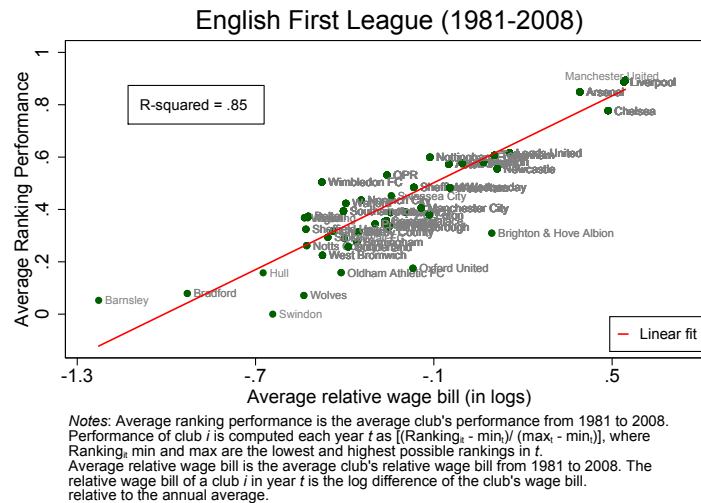
Fact 2. Clubs are heterogeneous in wage bills. Figure 2.2 reports the log of the clubs' wage bills in the English first league from 1981 to 2008. Each year, we observe a dispersion of wage bills supporting club heterogeneity. It is also worth mentioning that wage bills are linearly increasing over the years for all clubs, without being affected by the Bosman ruling shock. A common explanation for the increasing trend in wage bills is the increasing price of talent. Players earn wages proportional to their talent and demand for talent is increasing, which raises its price.

Figure 2.2: Club's Wage Bills



Fact 3. Team success depends on sporting talent. A team's sporting talent is highly correlated with its success and performance. Thus, the higher the club's sporting talent, the higher its wage bill, and the higher its ranking.¹ In Figure 2.3, we observe that wage expenditures and success are heavily correlated in the English Premier League between 1981 and 2008.

Figure 2.3: Average Wage Bill and Team Success



Fact 4. Soccer is not immune to prejudice. Racial wage gaps persist in many different occupations, and evidence suggests that at least some of this is due to discrimination from prejudiced firms.² English soccer is no exception. Racist incidents, such as the English national team captain aiming a racial slur at another player, continue to occur with alarming regularity.³

¹In Appendix A.1.1, we construct a proxy for a team's sporting talent, based on a crude measure of players' quality. Figure A.1 depicts a clear linear correlation between wage bill and team quality, with an r-squared of 0.81.

²See Lang & Lehmann (2012) or Charles & Guryan (2011) for recent literature reviews documenting the persistence of racial wage gaps in the US.

³See our online appendix for a more detailed compendium.

2.2 THEORETICAL PREDICTIONS FROM A SEARCH MODEL OF DISCRIMINATION

In Appendix A.2, we build a simple labor market search model in the style of Burdett & Mortensen (1998) and Bowlus & Eckstein (2002). This model presents three features. First, it is specific to the economics of soccer, relying on the four stylized facts that emerge from Section 2.1. Second, it links race-based wage discrimination to job-to-job mobility constraints. Finally, it offers a micro-foundation for our discrimination market test. Given that sporting success depends on the quantity of sporting talent, the model generates the following predictions:¹

- When job-to-job mobility is constrained,
 - the team’s sporting talent depends on the club’s wage bill and its share of black players;
 - disliked players change club less often than preferred ones;
- When job-to-job mobility is unconstrained,
 - the team’s sporting talent depends on the club’s wage bill *but not* in its share of black players;
 - disliked players move *more often* than preferred ones.

The story behind the model is straightforward.² Clubs face a trade-off between offering individual wages that perfectly reward players’ talent, ensuring they do not search on the job, or

¹These predictions correspond to Propositions A.2.2 and A.2.3 shown in the Appendix.

²Note that our model does not distinguish consumer from employer discrimination. However, previous research shows that discrimination in soccer is unlikely to be consumer-driven. Preston & Szymanski (2000) find no evidence that having more black players decreases attendance or club revenue, and no correlation between measures of prejudice at the local level and the number of black players hired.

alternatively offering lower wages, but facing costly turnovers if their players are poached by another club. If job-to-job mobility is constrained, and clubs can set wages to some degree, then clubs exhibiting racial biases can persist, since discriminated players find it hard to find new employers. By contrast, when constraints on job-to-job mobility are low, players can relocate or credibly threaten to leave, which weakens the monopsony power of prejudiced firms. The team's sporting talent is then independent of the club's share of black players. Appendix A.2 presents the model and the propositions.

2.3 EMPIRICAL DESIGN

In moving from theory to estimation we face one issue. The levels of mobility constraints in the European soccer market are unfortunately not readily quantifiable. In lieu of quantifiable measures, we use the Bosman ruling shock to infer information on the level of constraints. The pre-Bosman period was characterized by high constraints on mobility, which have been removed for EU players in the post-Bosman era. Recall that EU players may now receive job offers from all the EU clubs without quota limitations and out-of-contract restrictions. We expect the post-Bosman constraints to be sufficiently low to eliminate the pre-Bosman wage discrimination.

2.3.1 THE MARKET-BASED TEST FOR DISCRIMINATION

Our empirical analysis uses Szymanski's market-based test. This test is a clever way out of the omitted variable bias that plagued conventional approaches to testing for racial discrimination

using earning functions (Palacios-Huerta, 2014). A conventional approach would be to run a Mincerian individual wage equation, controlling for individual ability and adding a dummy for the player's skin color. However, controlling for individual ability/productivity has proven difficult since there is no clear way to measure individual ability/productivity in teamwork, such as soccer. Moreover, individual wage data for soccer players are not publicly available for all players on a large time-scale.

The intuition behind the market test is simple: if all individual talent is perfectly rewarded, the team's performance (a function of talent) should be perfectly explained by the team's wage bill. Crucially, this performance should be independent of the team's racial composition. By contrast, for a given wage bill, if teams fielding an above-average proportion of black players systematically outperform clubs with a below-average proportion of black players, then the labor market may be unfair towards black players (i.e., their talent is not fully rewarded and they face wage discrimination). This test is compatible with our theoretical framework (see subsection A.2.2).

The market test for discrimination requires information on the skin-color of players that we can match with extensive individual data, as well as data on club characteristics (wage bill, team performance). Appendix A.1 explains how we collect and code the data.

2.3.2 THE BOSMAN SHOCK AS A SOURCE OF IDENTIFICATION

We apply the market test for discrimination to a panel of all English clubs in the first league from 1981 to 2008, and we explore the Bosman ruling as an exogenous mobility shock to the European

soccer labor market. The Bosman ruling was decided on December 15, 1995, by the European Court of Justice. This important decision lifted restrictions on soccer player mobility based on the European Community Treaty of the free movement of labor (article 39). This decision had a profound effect on transfers in the European soccer market by banning restrictions on EU players in the EU's national leagues and by allowing players in the EU to move to another club at the end of their contract without a transfer fee being paid.

Though this decision came into force in December 1995, it could have been anticipated. Indeed, the Bosman case had been submitted to the Court two years before, on October 6, 1993. Thus, in December 1993, the European Union of Football Associations amended the regulations governing the *Status and Transfer of Football Players*. This amendment provided that a player may enter into a contract with a new club when the contract between him and his club has expired, has been rescinded or will expire within six months. However, the two clubs were still forced to agree on a transfer fee with the case going to a tribunal in case of disagreement. To prevent any contamination of the results caused by a possible anticipation, we omit the 1994-1995 and 1995-1996 seasons. Then, to identify a plausibly causal effect of intensified mobility on racial discrimination, we compare long and short time intervals before and after the Bosman ruling: 1981-1993 vs 1995-2008 in the longer case or 1989-1993 vs 1996-2001 in the shorter one. We look at shorter time intervals to make sure that our results are not just explained by other factors leading to different trends in discrimination.

Exploiting the ruling and using our extensive player data, we construct informative descriptive statistics (see Appendices A.1 and A.1.1 for data details). We focus here on the longer pre- and post-Bosman time intervals, 1981-1993 and 1996-2008. We have information on 3,788 first league players of which 70% are English.¹ The number of English players was higher before the Bosman ruling (93.2%) than after it (52.6%). In total, 10.7% of players are English and black; this number is fairly stable before (9.9%) and after (11.2%) Bosman. However, the number of black non-English players has skyrocketed from 0.8% before the ruling to 15% after it. Overall the number of white English players decreased from 83.3% before the ruling to 41.3% after the ruling.

To avoid comparing players with different nationalities, table 2.1 reports average individual characteristics on English players only. We thus compare black to white English players. We observe in the pre-Bosman period that they played a fairly similar number of matches. However, black players were slightly more qualified, one year younger and a bit more tenured. In the post-Bosman era, we observe notable changes: the quality difference becomes statistically not significant and black players are now *less* tenured than white players. Interestingly, all the differences in differences are significant between the two periods (reported in the last column of Table 2.1). In the post-Bosman era, players on average played less matches, are qualitatively better, more than 6 months older and more tenured.

The difference between the pre- and post-Bosman eras is also observed in Table 2.2 using *club*

¹ We use the term English but a minority of those players are Irish, Scottish or Welsh. Considering these players as English players makes sense in our context because they were not considered foreigners in the English soccer market and thus were not subject to the foreign quotas.

Table 2.1: Individual Differences in Means: Black vs White English Players

Variable	Pre-Bosman (1981-1993)			Post-Bosman (1996-2008)			Overall D. in D.
	Black	White	Diff.	Black	White	Diff.	
Player's number of matches	20.79 [9.3]	20.46 [9.6]	.33 (.37)	18.29 [8.3]	17.70 [8.6]	.59 ^c (.30)	2.68 ^a (.18)
Player's quality level	2.42 [1.9]	2.24 [2.1]	.18 ^b (.08)	2.88 [1.9]	2.78 [1.9]	.10 (.07)	-.54 ^a (.04)
Age of players	23.79 [4.0]	24.87 [4.4]	-1.08 ^a (.16)	24.92 [4.7]	25.59 [5.0]	-0.67 ^a (.17)	-.71 ^a (.09)
Tenure (in years)	2.50 [1.2]	2.38 [1.2]	.12 ^a (.05)	2.4 [1.1]	2.72 [1.8]	-.32 ^a (.07)	-.25 ^a (.03)
Observations	737	5152		980	3531		

Notes: This table reports average characteristics for black and white English players in the first league. 'Diff.' means difference in means between blacks and whites; 'D. in D.' means difference in difference between pre- and post-Bosman period. Standard deviations are in brackets; standard errors are reported in parentheses, with ^a, ^b and ^c denoting significance at the 1%, 5% and 10% level, respectively.

statistics. All differences in terms of wage bill, transfer fee record and stadium attendance are highly significant both economically and statistically. The only exception is the share of black English players, which is constant in both eras. These club variables are used as control variables or instruments and will be discussed later on.

Table 2.2: Club Differences in Means

Variable	Pre-Bosman (1981-1993)	Post-Bosman (1996-2008)	Differences in Difference
Wage bill (in millions of pounds)	2,9 [2,4]	35,5 [25,5]	32,6 ^a (1,5)
Transfer fee record (in millions)	1,31 [1,3]	8,96 [6,5]	7,65 ^a (0,38)
Stadium attendance (in thousands)	21,5 [8,9]	33,2 [12,0]	11,7 ^a (0,9)
Share of black English players	0.13 [.1]	0.14 [.1]	0.01 (0.01)
Observations	304	258	

Notes: This table reports average characteristics for English first league clubs. 'Difference' means difference in means between pre- and post-Bosman period. Standard deviations are in brackets; standard errors are reported in parentheses, with ^a denoting significance at the 1% level. Wage bills and transfer fee records are in nominal values.

2.3.3 ESTIMATED EQUATIONS

To apply the market test, relating a club's success to its wage bill and its share of black English players, we use two different measures of team success. The first one is based on league rankings. This measure allows us first to replicate Szymanski's discrimination result in the pre-Bosman period and then to show post-Bosman that discrimination disappears. The second measure is based on match results and offers us a larger number of observations to run the market test on two shorter time periods, 5 years before and 5 years after the Bosman ruling. Using shorter periods makes sure that the earlier results are not just explained by other factors leading to different "pre" and "post" trends in discrimination. Estimations on match results corroborate the finding of apparent wage discrimination in the *pre*-Bosman era, which disappears *post*-Bosman. The two measures of the team's success determine two different specifications.

The league success specification is as follows

$$\begin{aligned} \text{League Success}_{it} = & \alpha_i + \beta_1(\text{WageBill}_{it} - \overline{\text{WageBill}_t}) \\ & + \beta_2(\text{PlayersNb}_{it} - \overline{\text{PlayersNb}_t}) \\ & + \beta_3(\text{Shareblack}_{it} - \overline{\text{Shareblack}_t}) + \epsilon_{it}, \end{aligned} \quad (2.1)$$

where $\text{League success}_{it}$, is computed in relative terms as $\left(\frac{\text{Ranking}_{it} - \min_t}{\max_t - \min_t} \right)$. Ranking_{it} is the final ranking of team i at the end of season t ; and min and max are the lowest and highest possible

rankings each season. This relative success accounts for a different number of teams in the first league across years. α_i is a club i fixed effect, capturing permanent team-specific characteristics affecting success, such as a location effect,¹ and ϵ_{it} is the usual error term. The team's wage bill ($\text{WageBill}_{it} - \overline{\text{WageBill}_t}$) is measured as the log difference of the club wage bill relative to the annual average. The relative number of players ($\text{PlayerNb}_{it} - \overline{\text{PlayerNb}_t}$) is computed as the difference between the number of players used in club i in season t relative to the annual average. When controlling for club fixed effects and wage bills, the relative number of players captures “bad luck,” as high turnover typically reflects a high level of injuries sustained. Finally, the relative share of black players ($\text{Shareblack}_{it} - \overline{\text{Shareblack}_t}$) is measured as the share of black players' appearances for a team i in a given season t relative to the annual average. For now, we compute this variable based only on the share of matches played by Black *English* players,² since foreign players are differently impacted by the Bosman Ruling. We later discuss results comparing English and non-English players.

The coefficient of interest to us is β_3 , the effect of the share of black English players on success.

In case of race-based wage discrimination, we expect β_3 to be positive.

¹Some players could prefer to live in London rather than Liverpool for instance, which would allow London-based clubs to have better performances for lower wage bills.

²English players include Irish, Scottish and Welsh players - see footnote 1.

The match success specification is as follows

$$\begin{aligned}
\text{Match Success}_{ijt} = & \xi_{ij} + \beta_1 \log(\text{WageBill}_{it}/\text{WageBill}_{jt}) \\
& + \beta_2(\text{PlayersNb}_{it} - \text{PlayersNb}_{jt}) \\
& + \beta_3(\text{Shareblack}_{it} - \text{Shareblack}_{jt}) + e_{ijt}, \tag{2.2}
\end{aligned}$$

where $\text{Match Success}_{ijt}$ is the goal difference between the home team i and the away team j in year t . We expect that the larger the log difference in wage bills, $\log(\text{WageBill}_{it}/\text{WageBill}_{jt})$, the larger the difference in success (i.e., in goals scored). This specification also allows us to run the market-based test by including the difference in relative share of matches played by black English players. We also add as a control the difference in the number of players used ($\text{PlayersNb}_{it} - \text{PlayersNb}_{jt}$), a match fixed effect (ξ_{ij}), and the error term (e_{ijt}).

2.4 EMPIRICAL RESULTS

2.4.1 DISCRIMINATION MARKET TEST ON LEAGUE SUCCESS AND LONG TIME PERIODS

Table 2.3 shows the discrimination market test on league success (equation 2.1) before (col. 1 & 2) and after (col. 3 & 4) the Bosman ruling. We use relatively long time intervals here before turning to shorter ones in the next section. An alternative way to test for the presence of a structural break would be to interact the share of black players with a dummy variable for the post-Bosman

period, a procedure similar to a Chow-test. Results using this methodology are qualitatively and quantitatively similar, and are presented in Appendix A.4.

In Table 2.3, we use two different estimators, the ordinary least squares (OLS) in columns 1 and 3, and the fixed-effects “within” estimator in columns 2 and 4 to eliminate the individual effect (α_i). Standard errors are robust to club clustering to allow for a correlation of the error within clubs across years. In all columns, the relative wage bill variable has an expected positive effect on team success, which is economically and statistically significant. This effect is unsurprisingly larger if we do not control for club fixed effects. The relative number of players used in a season exhibits a negative effect in line with a “bad luck” interpretation, because high turnover typically reflects a high level of injuries sustained.

In the pre-Bosman era (columns 1 and 2), we find that success depends significantly on the team’s racial composition. Teams fielding an above-average proportion of black players outperform clubs with a below-average proportion of black players. This suggests race-based wage discrimination: because the talent of black English players is underpaid, the club’s wage bill does not fully explain team success. In other words, the club’s wage bill does not fully account for the team’s talent. In contrast, after Bosman (columns 3 and 4), the apparent wage discrimination disappears. Success is now independent of the team’s racial composition. This result is consistent with our theoretical predictions on the effect of relaxed mobility constraints on wage discrimination. The decrease in frictions associated with the ruling allows players to escape discrimination

by relocating or threatening to leave.¹

Table 2.3: Discrimination Market-test: League Success and Long Span

Dependent Variable:	League Success			
Sample:	<i>Pre-Bosman</i> (1981-1993)		<i>Post-Bosman</i> (1996-2008)	
Estimator:	OLS	Within	OLS	Within
	(1)	(2)	(3)	(4)
Relative log wage bill	0.509 ^a (0.05)	0.396 ^a (0.12)	0.474 ^a (0.04)	0.157 ^b (0.08)
Relative number of players used	-0.025 ^a (0.01)	-0.031 ^a (0.01)	-0.026 ^a (0.01)	-0.027 ^a (0.01)
Relative share of black English players employed	0.577 ^a (0.18)	0.491 ^b (0.25)	-0.130 (0.15)	-0.000 (0.20)
Observations	262	259	258	251
R^2	0.438		0.606	
Club fixed effect	No	Yes	No	Yes

Notes: the dependent variable is computed in relative terms as $\left(\frac{Ranking_{it} - min_t}{max_t - min_t}\right)$, where $Ranking_{it}$ is the final ranking of team i at the end of season t ; and min and max are the lowest and highest possible rankings each season. Robust standard errors are in parentheses, clustered by club, with ^a, ^b denoting significance at the 1% and 5% level respectively.

How economically meaningful is the estimate of discrimination in the pre-Bosman period?

Let us compare the 1993 situation of two clubs that are identical except in their share of black English players: (1) a club that does not employ black players, and (2) a club that employs 3.7 black players (the 1993 average number). Based on our estimates, we find that the club that does not employ black players should pay about 930,000 pounds more than the other club to employ equally talented white players and to obtain the same success. This overpayment amounts to

¹One potential issue is that black and white players play in somewhat different positions. Black players are less likely to be goalkeepers and more likely to be forwards. We could thus be picking up unconditional differences in wages between different positions. We run the same test as in equation 2.2 using the share of matches played by black goalkeepers, defenders, midfielders or forwards. We find similar results. Having more black players as defenders, midfielders or forwards increases team's performance before Bosman. However, after Bosman, there is no effect of having more black players in any position. Hence our results do not seem to be caused by relative positional differences between black and white players. These tests are available upon request.

18.5% of the average wage bill in 1993.¹

2.4.2 DISCRIMINATION MARKET-TEST ON MATCH SUCCESS AND SHORT TIME PERIODS

The match success measure enables us to run the market test (equation 2.2) on 5-year time periods before and after the Bosman ruling. Using shorter periods makes sure that the earlier results are not just explained by other factors leading to different “pre” and “post” trends in discrimination. We measure the effect of a difference in racial composition on the success of two teams competing against each other, conditioning on a match fixed effect. Standard errors are two-way clustered at the match and the receiving club level to allow for a correlation of the error within receiving clubs across years (e.g., any specific trend at home), and within matches across years. We consider that the match_{*ij*} between team *i* and team *j* is in the same cluster as *ji* because a common trend may influence the score of some club pairings independently of which club is receiving or visiting. Note, however, that the results hold for different ways of clustering the standard errors.

We check the extent of discrimination on two scenarios and 5-year time windows to stress how dramatic and quick the change is following the Bosman ruling: (1) 1989-1993 vs. 1996-2000 and (2) 1990-1994 vs. 1995-1999.² The results are reported in Table 2.4 using the within estimator.

¹Using the notation of Equation (2.1), the change in performance from having an average share of black players to a zero share is the change (Δ) in the relative share of black English players times β_3 . This change has to be compensated by a wage bill variation of β_1 times the change in the relative log wage bill to get an equivalent performance and hire equally-talented white players. Using the OLS estimates of Table 2.3 (col. 1), the average wage bill in 1993 in our data (i.e., around 5 million pounds), and the average share of black English players (i.e., about 15%), the 930,000 pounds are computed as $\exp[(\Delta \text{shareblack} * \hat{\beta}_3 / \hat{\beta}_1) + \ln(\text{mean wage bill})] - (\text{mean wage bill})$.

²The results using the match success specification on the longer pre-Bosman (1981-1993) and post-Bosman (1996-2008) periods are reported in Table A.2 in Appendix A.3.3. The structural break results on match success are pre-

In the first scenario (col. 1 and 2), we drop the years 1994 and 1995 to avoid any anticipation effect (see section 2.3.2), compared to the second scenario (col. 3 and 4).

Table 2.4: Discrimination Market-test: Match Success and Short Span

Dependent Variable:	Match Success			
Sample:	<i>Pre-Bosman</i>	<i>Post-Bosman</i>	<i>Pre-Bosman</i>	<i>Post-Bosman</i>
	(1989-1993)	(1995-1999)	(1990-1994)	(1996-2000)
	(1)	(3)	(2)	(4)
Difference in log wage bill	0.517 ^c (0.29)	0.434 ^c (0.26)	0.717 ^b (0.33)	0.390 ^a (0.13)
Difference in number of players used	-0.035 ^a (0.01)	-0.039 ^a (0.01)	-0.035 ^a (0.01)	-0.024 ^b (0.01)
Difference in black English players employed	1.409 ^a (0.45)	-0.819 (0.58)	1.203 ^b (0.55)	-0.588 (0.70)
Observations	1898	1558	1726	1571
Club-pair fixed effect	Yes	Yes	Yes	Yes

Notes: the Match Success dependent variable is the goal difference in the match. Robust standard errors in parentheses with ^a, ^b, and ^c denoting significance at the 1%, 5%, and 10% percent level, respectively. Standard errors are two-way clustered at the match and the receiving club level (see text for details).

In the pre-Bosman period (col. 1 and 3), the estimates of the share of black English players are positive and significant. Then, as expected, in the 5 years following the Bosman ruling the estimates become statistically insignificant (col. 2 and 4). This sharp difference before and after the ruling supports the idea that only a profound change in labor market conditions could have brought down discrimination this quickly.

The change in discrimination is remarkable when we consider that contracts for soccer players last on average around 3 to 4 years. Some players were thus not able to leave immediately after the Bosman ruling.¹ However, an alternative to changing club is to renegotiate the contract, to avoid

sented in Appendix A.4.

¹We may therefore expect the discrimination coefficient to decrease gradually. We tested for this possibility by interacting the share of black English players with a year trend in our Post-Bosman regression, and found a negative

the player being poached by other clubs. This renegotiation is quite frequent post-Bosman and typically leads to wage increases that may have favored discriminated players.

2.4.3 ARE NON-EU BLACK PLAYERS DISCRIMINATED POST-BOSMAN?

Although the ruling lifted constraints on EU players, non-EU players in the English Premier League are still subject to restrictive contracting conditions. For instance, to obtain a UK work permit, non-EU players must fulfill a set of stringent conditions.¹ As a consequence, despite a general decrease in frictions, mobility is relatively more constrained for non-EU players, even after the Bosman ruling. If prejudiced attitudes persist over time, black non-EU players could thus face wage discrimination. Some firms may take advantage of the mobility constraints to exert their monopsony power and act on their prejudice. This test relies on the heterogeneous impact of the Bosman ruling on different groups of players.²

We find evidence of wage discrimination against black non-EU players by performing the market test after Bosman. Results are reported in Table 2.5, where we decompose different shares of players: black English, black EU, black non-EU, and white non-EU.³ We find that the estimates

trend that was only significant at the 10% level (results are available upon request).

¹The conditions are such that a non-EU player must have played at least 75% of his national team's competitive matches over the last two years and his national team must be in the top 70 countries in the world. The appeals process allows for some flexibility in the rules, but the non-EU nationals playing in the Premier League are still expected to be of high quality.

²Unfortunately, this test cannot rely on temporal variation to create a difference in differences strategy because there are almost no black EU or non-EU players in the Premier League before 1995.

³Non-EU players are non-member players of the common market or the European free trade association zone. We do not report results for the pre-Bosman period, as there were only few non-English black players playing in England (see section 2.3.2).

of the share of non-EU black players are significant and positive after Bosman, even if statistical significance is lower when club fixed effects are introduced (col. 3 and 4). Those estimates suggest that wage discrimination could still be present in the English Premier League post-Bosman, but only for players who could face both prejudiced owners and mobility constraints, i.e., for black non-EU players.

Table 2.5: Post-Bosman: Are non-EU Black Players Discriminated? (1996-2008)

Dependent Variable:	League Success			
Estimator:	OLS	IV	Within	Within-IV
	(1)	(2)	(3)	(4)
Relative wage bill	0.461 ^a (0.03)	0.475 ^a (0.03)	0.144 ^b (0.07)	0.111 (0.13)
Relative number of players used	-0.026 ^a (0.01)	-0.025 ^a (0.01)	-0.026 ^a (0.01)	-0.026 ^a (0.01)
Relative share of black English players employed	-0.060 (0.13)	-0.051 (0.13)	0.095 (0.20)	0.096 (0.20)
Relative share of black EU players employed	0.056 (0.20)	0.029 (0.20)	0.125 (0.31)	0.174 (0.32)
Relative share of white non-EU players employed	0.082 (0.17)	0.073 (0.17)	0.256 (0.17)	0.267 (0.16)
Relative share of <i>black non-EU</i> players employed	0.566 ^a (0.20)	0.580 ^a (0.20)	0.587 ^c (0.34)	0.613 ^c (0.34)
Observations	258	257	251	249
Club fixed effect	No	No	Yes	Yes

Notes: the dependent variable is computed in relative terms as $\left(\frac{Ranking_{it} - min_t}{max_t - min_t}\right)$, where $Ranking_{it}$ is the final ranking of team i at the end of season t ; and min and max are the lowest and highest possible rankings each season. Robust standard errors in parentheses, with ^a, ^b, and ^c denoting significance at the 1%, 5%, and 10% percent level, respectively. Standard errors are clustered by club. We instrument the relative wage bill in columns 2 and 4 with the same instruments as in Table 2.6. The first-stages for the Within-IV (col. 4) is quite comparable with the one reported in Table (A.3) of Appendix (A.3.4) and available upon request with the corresponding usual statistics.

We present some additional results of the same market test in other EU countries (France and Belgium) in the post-Bosman period (see Table A.1 in the appendix). The results are in line with those found in the English Premier League post-Bosman: no evidence of discrimination against

black EU players, and some evidence of discrimination against black non-EU players, who still face some labor mobility constraints in France and Belgium.

2.4.4 ARE BLACK ENGLISH PLAYERS MORE MOBILE POST-BOSMAN?

As emphasized in our theory (see Proposition A.2.2 in the Appendix), we should expect a racially differentiated change in job turnover to accompany a change in job-to-job mobility. In our empirical context, we define job turnover as club transfers, i.e., moving from one club to another during a given season. Figure 2.4 contrasts the turnover of black (B) and white (A) English players by comparing their share in the total number of transfers per year with respect to their share in the total population. The variable analyzed is the following:

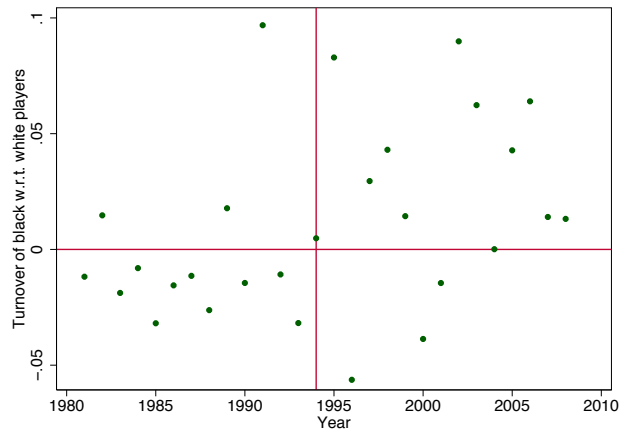
$$\text{Turnover} = \frac{\text{Share in transfers}_B}{\text{Share in population}_B} - \frac{\text{Share in transfers}_A}{\text{Share in population}_A}.$$

This variable is positive if black players change clubs more often in a given year than their white colleagues and negative if they do not change more often.

As shown in Figure 2.4, before the Bosman ruling, white players tended to change clubs more often than black players, but this tendency was somewhat reversed after Bosman.¹ This reversed trend suggests that black players “voted with their feet” when mobility constraints were removed.

¹It is unclear, however, when precisely this trend reversed. This may be because we imperfectly observe club transfers. We have data on the first division only, so what we are measuring are transfers within the first division. This may explain our outlying data point for the season 1991-1992, a season where four clubs were promoted and whose transfers were not perfectly captured.

Figure 2.4: Relative Turnover of Black English Players



This is in line with the reduction in job tenure observed for black players in Table (2.2) after the ruling. We also find evidence, presented in Appendix (A.3.1), that young black players took advantage of the new ruling to change clubs when they could have been discriminated against.

2.4.5 ARE OUR RESULTS PLAGUED BY WAGE BILL ENDOGENEITY?

We identify discrimination through the simple idea that if black players' talent is underpaid, the club's wage bill does not fully explain team success. However, two possible issues are worth mentioning: (1) the potential mismeasurement of the wage bills, and (2) the fact that bonuses result in reverse causation since a higher performance may induce higher bonuses and thus a higher wage bill (if salary is incentive based). As a result one might be worried that the extent of discrimination could be potentially biased in case of a systematic mismeasurement of wage bills in prejudiced clubs. Even if such a systematic pattern is unlikely, we address this concern first with

an instrumental variable (IV) estimator and then with a generalized method of moments (GMM) estimator.

Regarding the IV approach, we instrument the wage bill with the lagged performance in cups, lagged stadium attendance, and relative record transfer fees. The key identifying assumption is that these variables do not affect team success in the league, except through their effect on the club's wage bill.

We briefly discuss the relevance of our three excluded instruments before presenting the way they are constructed and then discussing the results. Regarding our first two instruments (performance in cups, and stadium attendance), we know that wages are set by clubs before the season starts. Higher attendances and better performance in cups in the previous season generate higher revenues, which allow clubs to recruit better players in the following season, and increases the wage bill. A detailed breakdown of club revenues is not usually available in the audited accounts from which we build our dataset. One exception is the analysis of data from the 1981/82 season, in which we find that match revenues coming from attendance made up 62% of total revenues. Nottingham Forest's accounts in that season provides a more precise picture. Reaching the quarter finals of the League Cup increased their turnover by about 10% compared to an early elimination. In 1982, most of the clubs' revenue came from stadium attendance, and performances in cups. This has changed in the intervening years, as the revenue generated by broadcasting and merchandising has increased. However, match revenues still accounted for 37% of total revenues

in the season 2006/07 (a more recent season where we also observe a partial breakdown of club revenues).¹ Regarding the third instrument, we use record transfer fees to capture potential buy-outs by rich owners. When buying out a club, new owners often break the club's transfer fee record, which increases the team's wage bill because of a positive correlation between transfer fees and wages. However, many record purchases prove to be poor value for money with on average very little effect on the final team performance.²

The instruments are constructed for each season as follows.³ Relative lagged stadium attendance in thousands of people is measured as the one-year lag of a club's log deviation from the lagged annual average. Performance in cups is measured as the one-year lag club performance in the Football Association Challenge Cup and the League Cup, using a linear performance index for each additional round reached. The relative transfer fee record variable is measured as the difference in a club's transfer fee record in a given season relative to the annual average.

The first-stage results of the IV are reported in Table A.3 (see Appendix A.3.4). The instruments have a significant effect on the relative log wage bill, except for the lagged cup performance in the post-Bosman era. The second-stage results of the within-IV approach are presented in columns 1 and 2 of Table 2.6. They confirm the Bosman ruling effect on wage discrimination. Whereas the coefficient for the share of black players employed is positive and significant before

¹Note that a good performance in cups also boosts merchandising and broadcasting revenues.

²It is worth noting that conditional on club fixed effects and wage bills, record transfer fees are not related to performance in the Premier League. A reason could be that most of the club's owners in the English Premier League are rich billionaires.

³See Appendix (A.1) for data sources.

the Bosman ruling (col. 1), it is insignificant post-Bosman (col. 2).

Table 2.6: Discrimination Market-test: League Success and Long Span - IV and GMM

Dependent Variable:	League Success			
Sample:	<i>Pre-Bosman</i>	<i>Post-Bosman</i>	<i>Pre-Bosman</i>	<i>Post-Bosman</i>
Estimator:	Within-IV		Arellano-Bond	
	(1)	(2)	(3)	(4)
Relative log wage bill	0.335 (0.25)	0.200 ^c (0.12)	0.576 ^a (0.10)	0.472 ^a (0.04)
Relative number of players used	-0.030 ^a (0.01)	-0.027 ^a (0.01)	-0.036 ^c (0.02)	-0.010 (0.01)
Relative share of black English players employed	0.501 ^b (0.25)	-0.111 (0.18)	1.364 ^b (0.54)	-0.116 (0.48)
Observations	259	247	262	257
Number of clubs	38	31	41	39
Number of instruments	3	3	28	30
AR1 p-value			0	0
AR2 p-value			0.70	0.37
Hansen p-value	0.72	0.15	0.47	0.48

Notes: the dependent variable is computed in relative terms as $\left(\frac{Ranking_{it} - min_t}{max_t - min_t}\right)$, where $Ranking_{it}$ is the final ranking of team i at the end of season t ; and min and max are the lowest and highest possible rankings each season. Robust standard errors are in parentheses, clustered by club, with ^a, ^b, and ^c denoting significance at the 1%, 5%, and 10% percent level, respectively.

The IV results could be affected by a weak instrument problem. If the instruments correlate only weakly with the wage bill, then statements of statistical significance may be misleading. However, the first stage F-statistics on the excluded instruments are above the recommended threshold of 10 (see Table A.3 in appendix A.3.4). It is also reassuring that the standard errors on the second-stage estimates are not much larger than those in the within model of Table 2.3. Moreover, the instruments pass standard validity assessments. The F-test of joint significance of the excluded exogenous variables is rejected at the 1% level. The test of overidentifying restrictions for the excluded instruments is also passed and the Angrist-Pischke first-stage chi-squared statistics

reject the null of underidentification (Angrist & Pischke, 2009).¹

In columns 3 and 4 of Table 2.6, we use the two-step GMM approach of Arellano & Bond (1991). This estimator differences away time-invariant club specific effects. It relies on the dynamic structure of the model for identification by using lagged levels of the independent variables as instruments for current differences. A problem with GMM estimators is that their validity is subject to the use of a relatively small or large number of instruments. A large number generates implausibly low values of Hansen tests of instruments exogeneity (Roodman, 2009), while using too few instruments is likely to generate a weak instruments problem and to deliver inaccurate estimates. Following Roodman's (2009) rule of thumb, we use a number of instruments that is strictly lower than the number of clubs (groups) in the sample. This strikes a balance between estimate consistency and test validity. The diagnostic tests (Hansen and first and second order autocorrelation) presented at the bottom of the table reveal no evidence against the validity of the instruments used by the GMM estimator.

The GMM estimates of the share of black English players employed confirm our main results: wage discrimination appears to be significant before the ruling (col. 3) but not afterwards (col. 4). Again, these findings are consistent with our theoretical predictions on the effect of relaxed mobility constraints on racial wage discrimination.

¹Results using the within-IV approach on the match specification are qualitatively similar and available upon request.

2.4.6 DOES CORPORATE CONTROL DECREASE DISCRIMINATION?

The rise of a market for corporate control of English professional clubs could potentially affect clubs' discriminatory behavior by increasing the competitiveness of English soccer (Palacios-Huerta, 2014). In a competitive market, discrimination would be competed away since nondiscriminatory clubs would employ more black players, given their lower cost at any given level of talent. We would then expect Arrow (1973)'s predictions to be realized. Black players' wages would increase until they became independent of race and discriminating clubs would be driven out of the market since they make lower profits than nondiscriminating clubs.

Some issues regarding the impact of this insightful hypothesis on discrimination in the Premier League could be raised, however. First, the emergence of a market for corporate control of English clubs started in the early 1980s, i.e., ten to fifteen years *before* the Bosman ruling.

The second concern is that a club experiencing very poor performance simply moves down from one division to another and is not necessarily driven out of the market. Relegation can lead to promotion the following season. So, competition per se may not discipline discriminatory behavior in our context.

The third concern is that there is no clear evidence that the market for corporate control has increased the competitiveness of English soccer. Only a third of the Premier League clubs during the post-Bosman period have been listed on the stock market.¹ Among the "Big Four", singled

¹See Andreff (2015) for the detailed list of English clubs that have been floated on the stock market, as well as their initial public offering and delisting date. Furthermore, Andreff states that governance changes very little with stock

out in the upper right corner of Figure A.1, only Chelsea¹ and Manchester United have been listed but not Arsenal and Liverpool. Interestingly enough, most clubs floated on the English stock exchange were unprofitable and all were eventually delisted, which would lead to an increase in discrimination today if corporate control is the causal mechanism driving the decrease in racial discrimination.

Despite these concerns, we directly test for the corporate control effect on discrimination. We benefit from the early listings on the stock market to test the impact of corporate control on the share of black players employed, as well as on club's performance and racial discrimination. To begin with we start with the assumption that the market for corporate control may lead to more black players being employed. We test this crucial hypothesis in Table 2.7. We regress the relative share of black players employed on a corporate control dummy, which is equal to one if the club is listed on the stock market. We firmly reject the hypothesis that clubs under corporate control hire more black players. Before the Bosman ruling, they hire less black players, while after Bosman, there is no significant difference in hiring between clubs under corporate control and those that are not.

In Table 2.8, we present a test for the effect of corporate control on club's performance and wage discrimination. We use specification 2.1 and introduce the corporate control dummy as an

market control, emphasizing that there is “no way to consider public trade of shares as a tool for improving listed clubs' governance and profitability.”

¹Note that Chelsea's ascension to the “Big Four” followed its removal from the stock market, after its shares were bought by a billionaire.

Table 2.7: Share of Black Players and Corporate Control

Dependent Variable:	Relative share of black players employed	
Sample:	<i>Pre-Bosman</i> (1981-1994)	<i>Post-Bosman</i> (1995-2008)
	(1)	(2)
Corporate control dummy	-0.061 ^a (.03)	0.003 (0.02)
Relative wage bill	0.032 (.03)	0.000 (0.03)
Observations	299	270
Club-fixed effect	Yes	Yes

Notes: the dependent variable is relative share of black players employed. Robust standard errors are in parentheses, clustered by club, with ^a, ^b, and ^c denoting significance at the 1%, 5%, and 10% percent level, respectively.

additional regressor. In columns 1 and 3, we find no effect of being a club floated on the stock exchange on club's performance, before or after Bosman. Finally, we present a direct test of whether the rise of corporate control had an effect on wage discrimination. In columns 2 and 4 of Table 2.8, we run the market test discrimination specification 2.1, with the addition of the corporate control dummy, and an interaction term between this dummy and the relative share of black players. We expect a negative estimate of the interaction term whether corporate control reduces the gap between white and black player wages. We find by contrast that the coefficient for the interaction term is positive and significant before Bosman, and positive and insignificant afterwards. The non negative estimate does not come as a surprise given that Arrow (1973)'s prediction relies on an increase in the demand for Black players, which is not what we find (see Table 2.7).

A cautious interpretation of these results is that corporate control of soccer clubs does not reduce discrimination, or at least not enough to be detected by differences between clubs on the stock market and under private ownership. Hence we believe that the mechanism causing the

Table 2.8: Discrimination Market-test: Effect of Corporate Control

Dependent Variable: Sample:	League performance			
	Pre-Bosman (1981-1993)	Post-Bosman (1996-2008)		
	(1)	(2)	(3)	(4)
Relative wage bill	0.399 ^a (0.12)	0.412 ^a (0.12)	0.167 ^b (0.08)	0.167 ^b (0.08)
Relative number of players used	-0.030 ^a (0.01)	-0.031 ^a (0.01)	-0.027 ^a (0.00)	-0.027 ^a (0.00)
Relative share of English black players employed	0.497 ^b (0.25)	0.438 ^c (0.26)	0.018 (0.20)	-0.005 (0.20)
Corporate control dummy	0.036 (0.11)	0.053 (0.11)	0.003 (0.04)	0.003 (0.04)
Corporate control * Relative share of English black players employed		0.988 ^a (0.40)		0.062 (0.27)
Observations	259	259	251	251
Club fixed effect	Yes	Yes	Yes	Yes

Notes: the dependent variable is computed in relative terms as $\left(\frac{Ranking_{it} - min_t}{max_t - min_t}\right)$, where $Ranking_{it}$ is the final ranking of team i at the end of season t ; and min and max are the lowest and highest possible rankings each season. Robust standard errors are in parentheses, clustered by club, with ^a, ^b, and ^c denoting significance at the 1%, 5%, and 10% percent level, respectively.

decrease in wage discrimination that we observe is the Bosman ruling, and not the rise of corporate control of soccer clubs.¹

2.4.7 HETEROGENEITY IN DISCRIMINATION

The results above show that discrimination decreases following the Bosman Ruling. We argue that clubs took advantage of the institutional constraints on player mobility to discriminate against their black English players. Four other tests could make us more confident that our story is correct:

- Clubs that are fighting for relegation or for the championship title need to be competitive, and so should be more likely to hire black players before Bosman.

¹Note that the corporate control hypothesis would not explain the differential rates in mobility, nor the wage penalty still faced by black non-EU players. Moreover, recall that black non-EU players still face discrimination in other European leagues (see Table A.1 in the appendix), where clubs were less likely to be floated on the stock exchange.

- Clubs that have a large share of black players should see a decrease in league performance after Bosman
- Clubs that have a large share of black players should see an increase in their relative wage bills after Bosman.
- Clubs that have a large share of black players should see some black players leave after Bosman

We define clubs that are fighting for relegation or the championship based on their relative wage bills. Clubs with an absolute value of their relative wage bill higher than 0.35 in a given season (around a quarter of clubs) are defined as "competitive". We also look at a subsample of clubs who seem to base their strategy on recruiting black players before the Bosman Ruling. These clubs are in the top 20% of clubs in terms of hiring black English players in any given year in the 5 years before the Bosman Ruling. We test these ideas in Table 2.9 below. Our power to estimate these effects is very low, especially since we are looking at narrow time-windows. Therefore, we only test for these effects in OLS. However, the point estimates all seem to concord with the hypotheses above.

Looking at column 1, competitive clubs hire more black players before Bosman and less afterwards though our results are insignificant. In column 2, 3 and 4, we can see how clubs that hire many black players behave before and after Bosman. As shown in the main tables of the article, they do better than other clubs before Bosman, but that advantage disappears after the Ruling is passed. They decrease their share of black players significantly post Bosman, and their relative wage bill is lower before Bosman but increases slightly after Bosman. Although we cannot be

conclusive since only our estimates on the change in the share of black English players are significant, it seems that both adjustments mechanism to the Ruling are at work: clubs adjust to the Ruling both by renegotiating wages and by letting their black players go to other clubs.

Table 2.9: Discrimination Market-test: Heterogeneity of clubs

Dependent Variable:	Share of black English players	League performance <i>Pre and Post-Bosman</i>	Share of black English Players	Relative Wage bill
Sample:				
Estimator:	OLS (1)	OLS (2)	OLS (3)	OLS (4)
Relative wage bill	-0.032 (0.03)	0.506*** (0.07)	0.031 (0.03)	
Competitive clubs	0.021 (0.04)			
Competitive clubs*Bosman Ruling	-0.052 (0.04)			
Bosman ruling	0.021 (0.03)	0.044 (0.04)	0.026 (0.02)	-0.045 (0.06)
Relative wage bill*Bosman ruling		-0.009 (0.09)	-0.052 (0.04)	
Clubs with High share of black players pre-Bosman		0.115* (0.06)	0.174*** (0.03)	-0.167 (0.10)
Clubs with High share of black players*Bosman Ruling		-0.138 (0.09)	-0.084** (0.04)	0.067 (0.14)
Number of players used black players *Bosman Ruling		-0.028*** (0.01)	-0.004* (0.00)	-0.016* (0.01)
Observations	225	225	225	225
Club fixed effect	Yes	Yes	Yes	Yes

Notes: the dependent variable is computed in relative terms as $\left(\frac{Ranking_{it} - min_t}{max_t - min_t}\right)$, where $Ranking_{it}$ is the final ranking of team i at the end of season t ; and min and max are the lowest and highest possible rankings each season. Robust standard errors are in parentheses, clustered by club, with ^a, ^b, and ^c denoting significance at the 1%, 5%, and 10% percent level, respectively. Clubs with an absolute value of their relative wage bill higher than 0.35 in a given season are defined as "competitive". Clubs with a high share of black players pre-Bosman are Aston Villa, Sheffield Wednesday, Chelsea, Coventry, QPR, Wimbledon and Crystal Palace. Seasons 1994-1995 and 1995-1996 are dropped in all estimations.

2.5 CONCLUSION

In this paper, we find evidence that wage discrimination has been eliminated following a decrease in labor market frictions. As shown in our model, an increase in job-to-job mobility can erode the monopsony power of firms, leading to an eradication of the racial wage gap. Our model appears to fit the empirical facts quite well.

A heartening interpretation of our results is that creating the right labor market conditions can cause wage differentials between white and black employees to disappear even if racist attitudes remain. Otherwise, firms can act on their prejudice, as seems to be the case for black non-EU players.



Labor Mobility and Racial Discrimination

A.1 DATA APPENDIX

Our dataset contains all professional soccer players, regardless of nationality, from the first English league from 1981 to 2008 and comes from Kleven *et al.* (2013). Our dataset is similar to Szymanski (2000) but differs in scope and time. Szymanski uses a panel of 39 clubs from four divisions in the English soccer league over the 1978-1993 period, i.e., pre-Bosman. In terms of scope, we have information on all players from the first division in England. Despite our focus on elite players, our number of clubs in the pre-Bosman period is fairly similar to Szymanski: 41 teams from 1981 to 1993 against 39. In terms of time, we cover the pre- and post-Bosman eras, from 1981 to 2008. To this dataset we add the following variables.

WAGE BILLS. We use wage bills from the Companies House website, a British government agency that collects annual reports from registered companies. Wage data are provided for almost all the English clubs in our sample. We are missing some data from clubs who went bankrupt during the season, such as Crystal Palace in 1998 or Leicester City in 2001, and some data from a few clubs that did not report wage bills in their financial accounts.

These wage data are considered as reliable because they are obtained from audited annual accounts. There are some issues however. First, the reports are not homogeneous over the 30-year period. Some clubs changed the ending date on their company accounts and reported annual results over thirteen months or more, in which case the data were adjusted on a pro rata annual basis. Then, we do not know what proportion of the pay is incentive related (e.g., bonuses for performing well in a cup competition) and what proportion is fixed. Finally, the wage bill is given for all staff, including salaries for scouts, statisticians, physiotherapists, and coaches. However, there are two reasons why it is unlikely to be a significant problem for our analysis. One is that the pay for most of the non-player employees is relatively small compared with the total wages of players; and the other reason is that the share of the pay for non-player employees likely accounts for a similar share of the wage bills in all clubs.

RACIAL INFORMATION. The race information was coded from an examination of players' photographs into categories of either black or not black (which we refer to as white).¹ This method might sound arbitrary because we code players as "black" if they appear to be "black". However, this method is actually a good way to model the potential for discrimination because discriminators prejudice an individual based on appearances (Palacios-Huerta, 2014).² These pictures were

¹Our full coding protocol is available upon request.

²For an explanation of why this appearance-based method is appropriate, Palacios-Huerta (2014) considers the case of the legendary Manchester United player Ryan Giggs. He appears to be Caucasian, and it was unlikely that he

obtained primarily from the reputable website www.transfermarkt.de, and when pictures from that site were not available, we conducted Internet searches. We obtained pictures for nearly all the players in our sample. The players whose photos were missing were primarily youth team players who had had little game time and could thus be discarded from our analysis.

CONTROL VARIABLES. In addition to information on race and wage bill, we have precise data on nationality, age, the number of matches played, the number of goals scored, national team selections (and their level - youth, A, ...), and whether a player participated in the World Cup. We use part of this information to create an objective, albeit imperfect, measure of player's sporting quality (see Appendix A.1.1). Moreover, we added information on team's ranking and attendance from the European soccer statistics website.

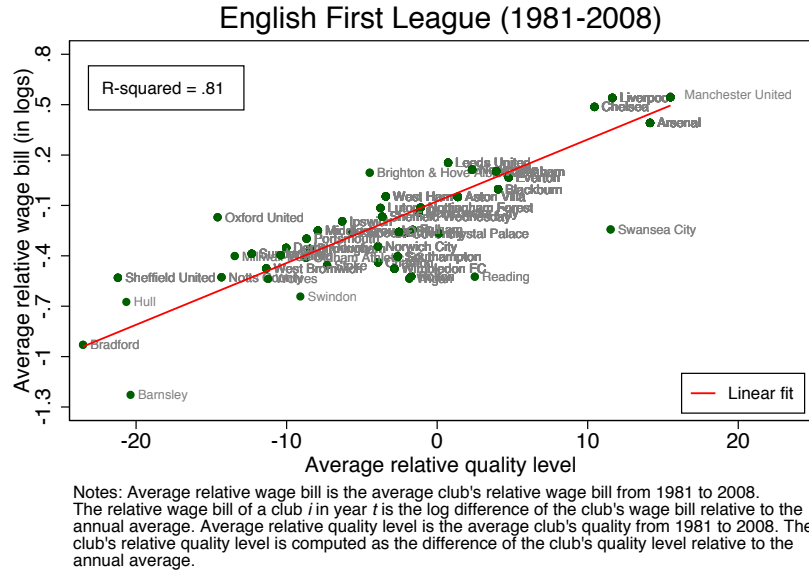
A.1.1 SPORTING TALENT AND WAGE BILL

In order to check whether sporting talent is well approximated by wage bills, we created a team specific talent index that sums the talent of each team's players. This index is computed for each club based on the experience of their players (i.e., their total number of matches played in the Premier League), the age of their players, and whether they are selected by their national team, and at what level (youth team, B team, A team). This variable is a crude measure for the team's sporting talent, but gives a rather sensible proxy for the quality of the team. As an example, the best performers, identified in Figure 2.3, which are Arsenal, Chelsea, Liverpool and Manchester United, are also the best teams in terms of quality with the highest wage bills. A regression of the average team's quality on average wage bill gives a R-squared of 0.81. The wage bill does

faced discrimination as a professional player during his career because discriminators prejudice an individual based on appearance. However, after he became famous, he publicly revealed that he had been victim of racism as a child because of his father's skin color. This revelation came as a surprise to his fans.

approximate quite precisely for the team's sporting talent.

Figure A.1: Average Wage Bill and Team's Quality



A.2 A SEARCH MODEL OF DISCRIMINATION IN SOCCER

This simple model presents three features. First, it is specific to the economics of soccer, relying on the four stylized facts that emerge from Section 2.1. Second, it links race-based wage discrimination to job-to-job mobility constraints. Finally, it offers a micro-foundation for our discrimination market-based test.

We first present player and club behavior, as well as the individual wage setting (A.2.1) before introducing the role of taste discrimination (A.2.2).

A.2.1 CLUBS, PLAYERS AND INDIVIDUAL WAGE SETTING

CLUBS. Clubs are heterogeneous in their wage bill (fact #2 presented in Section 2.1) but homogeneous in size because the number of players per team is fairly rigid. Team success is measured by league position (fact #1) and depends on the quantity of sporting talent (fact #3). Clubs maximize team success simply by minimizing the total cost of talent given total resources.¹ We assume, for simplicity, that total resources, R , are exogenously given when the hiring decisions are made. We consider the production function to be linear in talent, so we abstract from analyzing diversity, talent distribution across players and spillover effects within clubs to focus on the role of prejudice. Following our fact #4, we assume that some clubs hold a ‘taste’ for racial discrimination.² For the sake of simplicity we are agnostic in the model about whether the club’s prejudice comes from fellow players, owners, managers or supporters. However, in our context, it seems more likely that discrimination is driven by employers rather than by consumers.³

PLAYERS. The mass L of players is divided into two types according to their appearance, A and B . Type B could be discriminated against by clubs. All players are heterogeneous in talent independent of their type, and each type has the same distribution of talent. Talent is thus not differentiated by race or color. The decision problem faced by a player is the following. He max-

¹Clubs may also care about profits and not only about success. Considering that teams are profit maximizers is “an assumption with which sports economists have been quite comfortable in the United States, but which often seems less appropriate in the case of European soccer” (Szymanski, 2003: 1170). Garcia-del-Barrio & Szymanski (2009) confirm empirically that soccer club choices in England and Spain “are more closely approximated by win maximization than by profit maximization in both the short term and the long term”.

²Altonji & Blank (1999) reviews the pioneering works introducing search into taste-based theories of discrimination. Note that statistical discrimination, based on stereotypes and made possible by imperfect information, is less of an issue in our setting since clubs easily observe the performance of soccer players.

³Previous research shows that discrimination in soccer is unlikely to be consumer-driven. Preston & Szymanski (2000) find no evidence that having more black players decreases attendance or club revenue, and no correlation between measures of prejudice at the local level and the number of black players hired.

imizes utility over an infinite horizon in continuous time. He receives, at random time intervals, information about new or alternative jobs. This information is encapsulated in the parameter of the Poisson arrival process, λ , which denotes the arrival rate of job offers. This parameter reflects the general state of the labor market, including contracting rules, institutional constraints and barriers to mobility. Job-player matches are destroyed at an exogenous positive rate, δ , when players retire. Players are assumed to be risk neutral, with discount rate r , and must respond to offers as soon as they arrive.

INDIVIDUAL WAGE SETTING. In general there will be many causes explaining individual wage setting, such as the role of appearance that will be considered below. We abstract from this complexity and assume that a club offers to a player two types of contracts: perfect and imperfect. In both cases, the wage (ω) players receive is function of their talent, t . The contract is considered to be “perfect” when the player receives a wage that perfectly rewards his talent ($\omega = t$), and “imperfect” when the player earns only a fraction of what his talent is worth ($\omega = kt$, where $0 < k \leq 1$). In any case, the player accepts the job offer if it pays a higher wage while employed or if the instantaneous utility of being unemployed is lower than that of being employed. Although this is obviously a simplified wage setting, we view it as a useful benchmark that leads to a simple job offer trade-off and will help us to shed some light on the role of prejudice.

The club faces a trade-off when offering an imperfect contract. It diverts a share $(1 - k)t$ of the player’s talent but takes the risk of seeing the player poached by a rival since the player continues to search on the job. A player is poached with probability $\lambda\gamma$, where γ is the probability of receiving a perfect contract. Once the player is poached, the club incurs a turnover cost $c(t)$, which accounts for the leaving, replacement and transitions costs, such as the administrative costs involved with processing a separation, any payroll costs relating to filling the vacant position or time taken to

select and recruit a replacement, among others. We assume turnover costs to be club specific, since clubs may differ in their capacity to recruit new players. For instance, London-based teams may find it easier to attract new players.

The club's value of offering an imperfect contract (J_I) is based on the trade-off between the gain from diverting part of the talent $(1 - k)t$ and the loss that occurs when the player retires (δ) or is poached ($\lambda\gamma$):

$$\begin{aligned} rJ_I &= (1 - k)t - \delta J_I - \lambda\gamma(J_I + c(t)), \\ \Leftrightarrow J_I &= \frac{(1 - k)t - \lambda\gamma c(t)}{r + \delta + \lambda\gamma}. \end{aligned} \quad (\text{A.1})$$

The club's wage setting for a talent t depends on the comparison of J_I versus J_P , the club's value of the perfect job offer, which is null because $\omega = t$:

$$rJ_P = t - \omega - \delta(J_P) \Leftrightarrow J_P = 0. \quad (\text{A.2})$$

The club offers a perfect contract if:

$$\begin{aligned} J_P > J_I &\Leftrightarrow 0 > \frac{(1 - k)t - \lambda\gamma c(t)}{r + \delta + \lambda\gamma}, \\ \Leftrightarrow c(t) &> \frac{(1 - k)t}{\lambda\gamma}. \end{aligned} \quad (\text{A.3})$$

Equation (A.3) states that the club offers a perfect contract to the player if the turnover cost is higher than the rent extracted from the player weighted by the probability of being poached by a rival. To get a simple closed-form solution for the probability γ of offering a perfect contract, we assume that c follows a Pareto distribution with a lower turnover cost bound \tilde{c} and shape

parameter $\alpha \geq 0$. This assumption implies a distribution of turnover cost draws given by

$$G(c) = \left(\frac{c}{\tilde{c}}\right)^{-\alpha}, \quad c \in [\tilde{c}, \infty],$$

where the shape parameter α indexes the dispersion of turnover cost draws. The Pareto parametrization of c is intuitive because most turnover costs are low, but as α decreases, the relative number of high turnover costs increases, and the cost distribution becomes more concentrated at these higher cost levels. Since teams with low turnover costs have their players poached more often, they will offer most of the jobs. Based on the inequality (A.3) and assuming that c is distributed Pareto and linear in talent, such that $c(t) = ct$, yields a closed-form solution for γ :

$$\gamma(t) = \left(\frac{\tilde{c}\lambda\gamma}{1-k}\right)^\alpha = \left(\frac{\tilde{c}\lambda}{1-k}\right)^{\frac{\alpha}{1-\alpha}}. \quad (\text{A.4})$$

The probability of offering a perfect contract (γ) depends primarily on two important variables, λ and k (see Eq. A.4). First, everything else being equal, the arrival rate of job offers, λ , determines the strength of the club's monopsony power. The lower the value of λ , the lower the job-to-job transition, and thus the higher the club probability of offering an imperfect contract to their employees. Second, as k decreases, clubs divert a larger share of the monetary value of a player's talent; therefore offering an imperfect contract to a player is more attractive.

A.2.2 CLUB'S PREJUDICE AND MONOPSONY

Now that we have worked out the endogenous probability of a player receiving a "perfect" job offer, we introduce discrimination into our model. Similarly to most search models with discrimination, such as Bowlus & Eckstein (2002), Bowlus *et al.* (2001) and Black (1995), we introduce

discrimination by supposing that there is hiring discrimination. We are agnostic about whether discrimination comes from all or a subset of clubs, who would be less willing to hire a B-type player. This discrimination is represented in our model by a lower arrival rate of job offers for B-type players, $\lambda\theta$, with $0 \leq \theta \leq 1$. This implies that they are less likely to find a club that wants to offer them a perfect contract. Clubs know that B-types are less likely to be poached, and endogenously offer them lower wages. $\theta < 1$ leads to differences in γ , the probability of receiving a perfect contract. Note that for a given θ , from the Bosman ruling and the associated increase in λ , we expect more B-type players to be poached, and discrimination to decrease.

The intuitive reason is that monopsony power is reduced when job-to-job transition is less constrained, i.e., when the value of job offer arrivals (λ) is high. Clubs should thus offer *perfect* contracts to retain their players and avoid costly turnovers. As a consequence, for sufficiently high λ , preferred and disliked players receive the same talent-adjusted wage. The next proposition sums up our discussion:

Proposition A.2.1 The higher the rate of job offer arrivals (λ), the higher the probability of obtaining a perfect contract (γ). Thus, for a sufficiently high λ , race-based wage differentials disappear despite club prejudice.

If we interpret labor mobility in our model as the transition from imperfect to perfect contracts, then we can also derive a second testable proposition:

Proposition A.2.2 A decrease in job-to-job mobility constraints causes a racially differentiated change in job turnover.

A.2.3 WAGE DISCRIMINATION MARKET TEST

From the assumption on clubs' prejudice and the probability of obtaining a perfect contract γ (Eq. A.4), we can ground Szymanski's discrimination market-based test. The test is simple: discrimination can be said to exist if, for a given wage bill, clubs fielding an above-average proportion of B players systematically outperform clubs with a below-average proportion of B players. This test requires specifying the share of B players in a club, which we derive below based on a very stylized recruitment process.

This stylized hiring process is subject to a zero profit condition such that the club's resources (R) are equal to the club's wage bill (Ω) plus the sum of realized turnover costs (C). We assume each club knows its individual turnover costs c , and fields a fixed number of players, n . Sporting success depends on the quantity of sporting talent, T , hired by the club:

$$T \equiv (n - 1)t + \tilde{t}.$$

Clubs thus hire $n - 1$ players with the same individual talent t , and one marginal player with a potentially different talent \tilde{t} . The share ϕ of imperfect contracts offered to the $n - 1$ players depends on labor market conditions (λ), player type (A or B), and the associated turnover costs. It appears readily that the higher the share of imperfect contracts offered to the $n - 1$ players, the higher the rents extracted from these players, the more talented the marginal player, and the higher the team's sporting talent, T . We assume for simplification that the marginal player cannot

be underpaid and his talent \tilde{t} is such that the zero profit condition holds:

$$\begin{aligned} R = \Omega + C &= (n-1)[\phi kt + (1-\phi)t] + \tilde{t} + (n-1)\phi\lambda(\theta\mu\gamma_B + (1-\mu)\gamma_A)ct \\ &= (n-1)t[\phi(k + c\lambda(\theta\mu\gamma_B + (1-\mu)\gamma_A)) + 1 - \phi] + \tilde{t}, \end{aligned} \quad (\text{A.5})$$

where ϕ is the share of players that are not paid their marginal product ($\omega < t$) and μ the share of B-type players in the team.

Armed with this framework, we illustrate how the share ϕ depends on the value of turnover costs c by focusing on two interesting cases. We view them as useful benchmarks that can shed light on the role of labor mobility on wage discrimination.¹

Case 1 If turnover costs are high, such that $c \geq \frac{(1-k)}{\lambda\theta\gamma_B}$, all players are paid the monetary value of their talent t to avoid being poached, so that $\phi = 0$. No player is underpaid and clubs do not earn any rent.

Case 2 If turnover costs are such that $\frac{(1-k)}{\lambda\gamma_A} \leq c \leq \frac{(1-k)}{\lambda\theta\gamma_B}$, clubs underpay all B-type players, while paying each A-type player its marginal product. Thus, the share of imperfect contracts ϕ is equal to μ . The talent of the marginal player is now

$$\tilde{t} = R - \mu t \times (n-1)(k + \lambda\theta\gamma_B c_B) - (1-\mu)t \times (n-1). \quad (\text{A.6})$$

These two cases depend crucially on labor market conditions, which affect the turnover costs. Our perspective is that the Bosman ruling modifies the general state of the labor market. It lowers the constraints on labor mobility and thus increases the arrival rate of job offers, λ .² Indeed, the

¹Note that a third case is such that if $c \leq \frac{(1-k)}{\lambda\gamma_A}$ all players are underpaid so that $\phi = 1$, and clubs get rents from all players. The marginal player's talent is $\tilde{t} = R - (n-1)t[k - 1 + \lambda c(\theta\mu\gamma_B + (1-\mu)\gamma_A)]$.

²We are not the first to consider that changes in λ may reflect changes in the state of the labor market. As stated in Manning (2003) "in the model of Burdett & Mortensen (1998), the extent of club market power is determined by the rate at which job opportunities arrive relative to the job destruction rate: the lower the arrival rate of job offers, the more market power clubs will have."

end of transfer fees for out-of-contract players and of quotas within the EU dramatically reduced mobility constraints. EU players, independent of their race, now receive a higher number of job offers, from a larger number of firms, compared to the pre-Bosman ruling period. It is thus reasonable to assume that λ is higher after the ruling than before.

In Case 1, differences in sporting talent T across clubs only depend on differences in wage bills and not on the share of black players. This case corresponds to a post-Bosman ruling type situation with high turnover costs due to high λ s. This situation discourages clubs from indulging in discriminatory tendencies and team success does not depend on μ , the club's share of black players. Case 2 can be interpreted as the pre-Bosman ruling situation with different types of contracts due to a taste for discrimination. We view the Bosman Ruling as a move from case 2 to case 1. By combining Eqs. (A.6), (A.5), and $R = \Omega + (n - 1)t\mu\lambda\theta\gamma_{BC}$, we get an interesting testable proposition:

$$\begin{cases} T = \Omega & \text{Case 1} \\ T = \Omega + \mu t(N - 1) * (1 - k) & \text{Case 2} \end{cases}$$

It follows that

Proposition A.2.3 When λ is high (case 1), the team's sporting talent and success only depend on the club's wage bill Ω . When λ is low (case 2), the team's sporting talent and success depends on both the club's wage bill Ω and the share of B-type players μ .

In Case 2, the reason why the share of black players matters is that by diverting part of the wage devoted to B-type players the club earns some rents that are used to increase its overall talent T . This is the essence of the discrimination market test.

A.3 FURTHER EVIDENCE OF DISCRIMINATION

A.3.1 EVIDENCE ON AGE PROFILES OF PLAYERS

The age profile of black players before and after the Bosman ruling also presents some evidence of discrimination. Figures (A.2) and (A.3) depict the age density of black players in “discriminating” and “non-discriminating” teams before and after the Bosman ruling, respectively. We define, although imperfectly, “discriminating” teams as the ones whose proportions of black players in the squad is lower than 25% of the other squads, and “non-discriminating” teams as those where the proportions of black players is higher than 75% of the other squads. Of course, this way of selecting discriminating and non-discriminating teams is far from perfect, but it is consistent with our empirical strategy and model. In the model, firms that discriminate are more likely to have their black players poached by rival firms. We should therefore expect these firms to have less older black players.

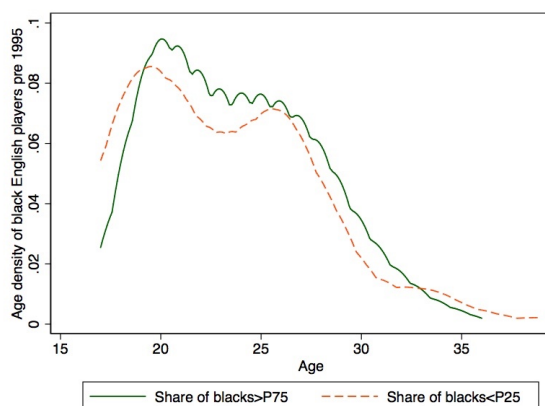


Figure A.2: Pre-Bosman

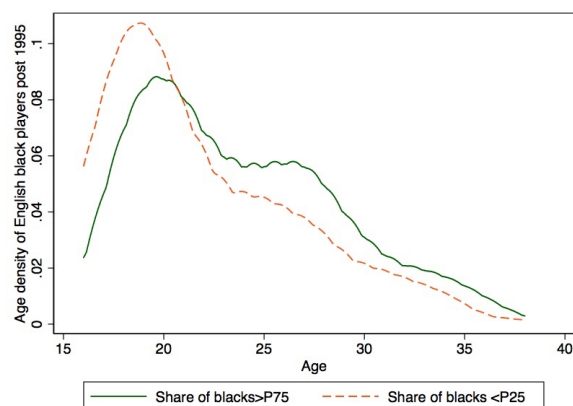


Figure A.3: Post-Bosman

In Figure A.2, we observe that the age densities of black players in pre-Bosman are quite similar in discriminating and non-discriminating clubs, but there is a huge change post-Bosman (Fig-

ure A.3): the age density of black players is more right-skewed in discriminating clubs. These differences could be explained by job-to-job mobility constraints. When players are young, they tend to play for their local clubs or for any club that wants them, whether these discriminate or not. Then, when labor mobility is low, as is the case pre-Bosman, black players employed in discriminating clubs can not move. This explains the similar age profiles depicted in Figure A.2 between discriminating and non-discriminating clubs. By contrast, when job-to-job mobility is high, as was the case post-Bosman, black players in discriminating clubs are mostly youngsters since older players leave as soon as they can. This explains why the age density of black players is more right-skewed in discriminating clubs in the post-Bosman era (Figure A.3).

A.3.2 EVIDENCE ON OTHER EUROPEAN COUNTRIES

Club wage bill data outside of England is much harder to find. Although we have access to data for Belgium and France, in none of these cases can we go back beyond 2000, much less 1995. Table A.1 reports results using the market test for discrimination in France (2002-2009) and Belgium (2005-2008).¹ We exploit here an interesting feature of the data, which is the distinction between two types of non-EU citizens: Players from countries who have signed the Cotonou agreement, mostly from Africa, the Caribbeans and Oceania, and those who have not signed the agreement, essentially from South Americans countries. Players from “Cotonou countries” can play freely in France and Belgium since 2003, whereas the others are still subject to quotas. We use the same market test as the one described in section 2.4.3. We find some evidence that black players from non-Cotonou (and non-EU) countries face wage discrimination in Belgium and France. These players do not benefit from the Bosman ruling and free mobility of labor.

¹Wage bill data are missing for small Belgian clubs.

Table A.1: Discrimination Market-test: Club's League Success in France and Belgium

Dependent Variable: Country: Time Period: Estimator:	Club's League Success			
	France		Belgium	
	2002-09		2005-08	
	OLS	Within	OLS	Within
	(1)	(2)	(3)	(4)
Relative wage bill	0.371 ^a (0.04)	0.369 ^a (0.04)	0.295 ^a (0.07)	0.335 ^a (0.05)
Relative number of players used	-0.029 ^a (0.01)	-0.031 ^a (0.01)	-0.001 (0.01)	-0.002 (0.01)
Relative share of Native black players employed	-0.101 (0.21)	-0.087 (0.20)	-0.538 (0.35)	-0.634 ^c (0.35)
Relative share of EU black players employed	-0.741 (1.12)	-0.377 (1.11)	0.84 (0.19)	0.227 (0.18)
Relative share of non-EU white players employed	-0.451 (0.35)	-0.335 (0.30)	0.146 (0.30)	0.283 (0.32)
Relative share of black non-EU players employed, of which:				
Cotonou players	0.195 (0.24)	0.229 (0.25)	-0.280 (0.24)	-0.099 (0.31)
non-Cotonou players	0.691 ^b (0.33)	0.750 ^c (0.40)	2.074 ^a (0.42)	2.207 ^a (0.51)
Observations	137	137	51	51

Notes: the dependent variable is the goal difference in the match. Robust standard errors in parentheses with ^a, ^b, and ^c denoting significance at the 1%, 5%, and 10% percent level, respectively. Standard errors are clustered at the club level (see text for details). Among black non-EU players only non-Cotonou ones are constrained post-2003.

A.3.3 ADDITIONAL RESULTS ON LONG-TIME WINDOWS

We present additional results using long-time windows with match-level data and the IV strategy. Our results can be interpreted as follows. Before Bosman (columns 1 and 2), a team that switched all its players for black players could expect to win the match by one goal. In other words, because of wage discrimination, the club with only black players can hire more talented players. A similar effect could be achieved by doubling the wage budget. By contrast, the non-significant estimates on the share of black players, reported in columns 3 and 4, suggest that wage discrimination has disappeared post-Bosman.

In columns 2 and 4, we use an instrumental variable approach to account for possible mea-

Table A.2: Discrimination Market-test: Match Success and Long Span - Within and Within-IV

Dependent Variable:	Match Success			
	<i>Pre-Bosman</i> (1981-1993)		<i>Post-Bosman</i> (1996-2008)	
Sample:	Within	Within-IV	Within	Within-IV
Estimator:	(1)	(2)	(3)	(4)
Difference in log wage bill	0.698 ^a (0.18)	1.597 ^a (0.43)	0.503 ^a (0.12)	0.343 (0.27)
Difference in number of players used	-0.047 ^a (0.01)	-0.053 ^a (0.01)	-0.029 ^a (0.01)	-0.030 ^a (0.01)
Difference in share of black players employed	0.855 ^a (0.33)	1.162 ^a (0.43)	-0.117 (0.31)	0.065 (0.40)
Observations	4494	3073	4402	3264
Hansen p-value		0.79		0.20
Club-pair fixed effect	Yes	Yes	Yes	Yes

Notes: the dependent variable is the goal difference in the match. Robust standard errors in parentheses with ^a denoting significance at the 1% percent level. Standard errors are two-way clustered at the visiting and receiving club level (see text for details).

surement errors and reverse causality. The difference in wage bills is instrumented with the corresponding difference in the instrument variables used in subsection (2.4.1). The result of the first-stage estimates are reported in Table (A.4) in Appendix A.3.4. The instruments pass the standard validity assessments (see the bottom of Table A.4). The F-test of joint significance of the excluded exogenous variables is rejected at the 1% level and is above the recommended threshold of 10. The test of overidentifying restrictions for the excluded instruments is also passed and the Angrist-Pischke first-stage chi-squared statistics reject the null of underidentification (Angrist & Pischke, 2009). Moreover, it is again reassuring that the standard errors on the second-stage estimates (col. 2 and 4) are not much larger than those in the within model (col. 1 and 3, respectively)

Table A.3: First Stages: League Success and Long Span

Dependent Variable: Sample:	Relative Log Wage Bill	
	Pre-Bosman (1981-1993)	Post-Bosman (1996-2008)
Relative share of black English players	0.009 (0.140)	-0.003 (0.140)
Relative number of players used	0.007 ^c (0.004)	-0.002 (0.003)
Lagged relative log attendances	0.248 ^a (0.008)	0.522 ^a (0.098)
Lagged relative cup performance	0.011 ^b (0.005)	0.004 (0.004)
Relative record transfer fee	0.125 ^a (0.036)	0.023 ^a (0.004)
Observations	259	268
Club fixed effects	Yes	Yes
F test of excluded instruments	14.83 ^a	28.69 ^a
Angrist-Pischke underidentification $\chi^2(3)$	46.41 ^a	83.99 ^a
Test of overidentifying restrictions	6.50	3.85
$\chi^2(2)$ p-value	0.72	0.15

Notes: robust standard errors in parentheses, clustered at the club level, with ^a, ^b and ^c denoting significance at the 1%, 5% and 10% level respectively.

A.3.4 FIRST STAGE RESULTS

A.4 TESTING FOR STRUCTURAL BREAKS

In the main body of the article, we use a split sample approach to investigate the effect of the Bosman ruling. Another approach is to test for a structural break in 1995, by using an interaction term between the Bosman ruling and the share of black English players. We complement this approach by interacting the wage bill and the Bosman ruling, in order to test for structural breaks in the relationship between performance and wages. The advantage of the structural break approach is to test directly for the difference between the two periods, and to keep the coefficients for the non-interacted variables (the fixed effects and the variable for total players used) constant throughout the entire period. Moreover, the testable predictions for the model imply a decrease in discrimination following liberalization, not necessarily its disappearance.

The results depicted in Tables A.5 and A.6, and are unsurprisingly similar to what was pre-

Table A.4: First-Stages: Match Success and Long span

Dependent Variable:	Difference in Log Wage Bill	
Sample:	Pre-Bosman (1981-1993)	Post-Bosman (1996-2008)
Difference in share of black English players	0.035 (0.081)	0.116 (0.123)
Difference in number of players used	0.009 ^a (0.002)	-0.001 (0.002)
Difference in lagged stadium attendance	0.011 ^a (0.003)	0.009 ^a (0.003)
Difference in lagged cup performance	0.009 ^a (0.003)	0.004 ^b (0.002)
Relative record transfer fee	0.088 ^a (0.032)	0.271 ^a (0.044)
Observations	3129	3264
Pair of clubs fixed effects	Yes	Yes
F test of excluded instruments	11.74 ^a	17.09 ^a
Angrist-Pischke underidentification $\chi^2(3)$	36.49 ^a	53.46 ^a
Test of overidentifying restrictions	0.461	3.034
$\chi^2(2)$ p-value	0.79	0.22

Notes: robust standard errors in parentheses are two-way clustered at the visiting and the receiving club level, with ^a and ^b denoting significance at the 1% and 5% level respectively.

sented in the main body of the text. Reassuringly, there is no differential effect of wages after Bosman Table A.5. Indeed, the share of black English players has a positive and significant effect on team performance before Bosman, while the post-Bosman interaction term with the share of English black players is negative and significant. The results are similar for the short time span reported in Table A.6. The coefficient for the share of black players is positive and significant before Bosman, though the post-Bosman interaction term in this case is only significant at the 10% level.

Table A.5: Structural Break: Long Span

Dependent Variable:	League Success			
Sample:	<i>Pre</i> and <i>Post</i> -Bosman (1981-2008)			
Estimator:	OLS	Within	Within-IV	AB
	(1)	(2)	(3)	(4)
Relative wage bill	0.509 ^a (0.05)	0.363 ^a (0.09)	0.409 ^b (0.20)	0.621 ^a (0.11)
Relative wage bill*Bosman ruling	-0.034 (0.07)	0.009 (0.08)	-0.019 (0.13)	-0.170 (0.13)
Relative share of English black players employed	0.575 ^a (0.18)	0.460 ^b (0.19)	0.465 ^b (0.19)	1.301 ^b (0.56)
Share of English black players employed*Bosman ruling	-0.722 ^a (0.25)	-0.584 ^b (0.28)	-0.605 ^b (0.28)	-1.379 ^b (0.62)
Number of players used	-0.026 ^a (0.00)	-0.026 ^a (0.00)	-0.026 ^a (0.00)	-0.022 ^c (0.01)
Bosman ruling	-0.008 (0.02)	-0.015 (0.03)	-0.013 (0.03)	-0.020 (0.08)
Observations	520	517	515	519
Club fixed effect	Yes	Yes	Yes	Yes

Notes: the dependent variable is computed in relative terms as $\left(\frac{Ranking_{it} - \min_t}{\max_t - \min_t}\right)$, where $Ranking_{it}$ is the final ranking of team i at the end of season t ; and min and max are the lowest and highest possible rankings each season. Robust standard errors are in parentheses, clustered by club, with ^a, ^b denoting significance at the 1% and 5% level respectively. Seasons 1994/1995 and 1995/1996 are dropped as in the main body of the text.

Table A.6: Structural Break: Short Span

Dependent Variable:	Match Success	
Sample:	<i>Pre</i> and <i>Post</i> -Bosman (1989-2000)	<i>Pre</i> and <i>Post</i> -Bosman (1990-1999)
Estimator:	Within (1)	Within (2)
Relative wage bill	0.415 ^b (0.17)	0.528 ^a (0.19)
Relative wage bill*Bosman ruling	0.216 ^c (0.12)	0.129 (0.14)
Share of English black players employed*Bosman ruling	-0.904 ^c (0.51)	-1.049 ^c (0.57)
Share of black players employed	0.745 ^b (0.31)	0.919 ^a (0.32)
Number of players used	-0.046 ^a (0.01)	-0.034 ^a (0.01)
Bosman	0.038 (0.07)	-0.019 (0.10)
Observations	3624	3497
Club-pair fixed effect	Yes	Yes

Notes: the Match Success dependent variable is the goal difference in the match. Robust standard errors in parentheses with ^a, ^b, and ^c denoting significance at the 1%, 5%, and 10% percent level, respectively. Standard errors are two-way clustered at the match and the receiving club level (see text for details). Seasons 1994/1995 and 1995/1996 are dropped in the first column only.

3

Gender Discrimination in Recruitment Committees

Even though women make up the majority of Ph.D candidates in many academic disciplines in Europe, they are underrepresented in faculty positions and the gender gap increases as the air gets more rarefied. In the European Union, women made up between 40-60% of PhD graduates, 40% of APs, and only 21% of full professors, according to the European Commission's 2015 SHE

figures.^{1 2}

Many different explanations for vertical segregation in academia have been put forward: for instance, less investment in human capital, perhaps due to the unequal distribution of household labour, psychological differences, such as attitudes towards risk and competition,³ and also discrimination. Teasing out which of these potential causes is the most salient is important since misidentifying the true causes of gender disparities can lead to detrimental policies that may actually worsen gender gaps.⁴

If we believe that discrimination is holding back the careers of women, and that discrimination is caused by men with own-gender preferences, then one way to improve the prospects of women could be to increase the number of women among the academics making hiring and promotion decisions. A new reform by the French government took this line of reasoning and imposed gender quotas in academic recruitment committees.⁵ The aim of this article is to directly test the effect of the reform and *in fine* whether own-gender preferences exist.

In what follows, I use administrative data provided by French universities to investigate whether having more female members in academic recruitment committees leads to better outcomes for women candidates. Surprisingly, the raw correlations directly contradict the presence of own-gender preferences: the higher the share of women evaluators in a committee, the lower women

¹Chenu & Martin (2016) study the career trajectories of sociology professors in France, our country of interest, and show that only 11% of freshly-minted women sociology assistant professors from 1996-1999 had become full professors by 2012, compared to 35% of men. In 2011, women made up 50% of sociology assistant professors, but only 27% of full professors.

²Geuna & Shibayama (2015) show that women are more likely to leave academia at all stages in their career than men. However most positions in French academia are not tenure-track. Researchers can stay at the assistant professor level without fear of losing their position.

³See Bertrand (2011) for a survey on the effect of psychological differences on labour market outcomes.

⁴For instance, Antecol et al. (2016), show that a well-meaning policy aimed at inciting women to take up parental leave, a gender-neutral tenure clock stop, reduced the rate of women receiving tenure by 22% and increased the rate for men by 19%.

⁵The committee must be made up of at least 40% of each gender. (Link to the decree).

are ranked, even when controlling for publications and academic connections.

I then directly evaluate the impact of the reform, by exploiting the fact that disciplines are differentially affected by the quota; the quota is more likely to bind in fields where the share of women jurors was lower than 40% before the reform. I compare the ranks of women before and after the reform, using the disciplines that were already respecting the quota as a control group. Women receive worse ranks after the reform, and the effect is significant and very large. However, since we do not observe the votes of individual jury members, we do not know whether it is women who have opposite-gender preferences, or men who decide to vote against women whenever more women are their co-jurors, and we have to be cautious in interpreting these results. In fact, one element leans towards this second interpretation: the negative effect of the reform in disciplines where women are underrepresented is driven entirely by committees helmed by men.

These results cast doubt on whether gender quotas in recruiting are effective in solving vertical segregation. Though this reform improves the representativity of women,¹ the relatively small number of women eligible to participate in committees will mean that the administrative work of each will substantially increase, potentially harming their ability to publish. As argued by Vernos (2013), when discussing gender quotas in ERC evaluation panels, "quotas might make matters worse by overworking already-stretched female scientists". If gender quotas are not efficient in countering own-gender preferences, then the 2015 reform may end up doing more than good.

Although the struggles of women in academia and the barriers they face at many stages of their career are well documented, there have been few studies looking at the effect of the feminisation of evaluators on the prospects of women. Some articles try to study this effect indirectly. For instance, a recent study in France by Breda & Hillion (2016) showed that women were favoured

¹See Azmat (2014) for positive and negative aspects of increasing gender representativity in decision-making teams.

in oral exams for posts as secondary-school teachers in disciplines that had few women instructors. The articles that have tried to study a direct effect have found conflicting results. Some find that having women evaluators has a positive effect on the outcomes for women, on field data (Lincoln et al. (2012), Boring (2017), Zeltzer, 2015)¹ or audit studies (Edo et al., 2015)². A few find no effect (Abrevaya & Hamermesh (2012), Feld et al. (2016), Williams & Ceci, 2015).³ Ellemers *et al.* (2004), and Broder (1993), find opposite-gender preferences, in evaluation of doctoral students' work ethics and grant applications (using evaluations from men and women on the same paper) respectively. Whether own- or opposite-gender preferences exist in practice remains an open question.

Similarly, studies evaluating the effect of the gender of candidates and recruiters on the ranking of applicants find contradicting results. For instance Bagues & Esteve-Volart (2010), using data from law recruitments in Spain, find that having more women in committees does not increase the probability of women being hired, whereas Paola & Scoppa (2015), looking at recruitments of Economics and Biology professors, find that having at least one woman in the committee increases the probability of women being promoted. Bagues et al. (2017) use data on 100 000 applications to the post of professor in Italy and Spain. They find that having more women in the recruiting committee harms the chances of women being promoted, except for full professorships

¹The first article finds that having a male committee chair for research prizes raises the probability of men of receiving a prize. Boring (2017) shows that male students give higher evaluation grades to their male professors. Zeltzer finds that male generalists refer patients to male specialists relatively more often than do female doctors.

²They find that women recruiters are more likely to call back women.

³Abrevaya & Hamermesh (2012) look at whether female referees accept papers written by women more often, using article and reviewer fixed effects, and find no effect. Feld et al. (2016) compare students that are graded anonymously with students that are not, which enables them to distinguish between endophily and exophobia. They find evidence of endophily if instructors and students have the same nationality, but no effect for gender. Williams & Ceci (2015) run lab experiments, and find that women and men in biology, engineering and psychology (not in economics however) prefer to hire women over men with identical profiles. These preferences do not change depending on whether the evaluator is male or female.

in Spain.

My work builds on Bagues *et al* (2017) in several ways. In their data, candidates do not know their jurors, and are unlikely to work with them after the committee has convened. In the committees I study, half of the jurors are professors in the university and will have the highest-ranked candidate as a colleague, which could raise the stakes of hiring, when compared to the decision of promoting an assistant professor to full professor in another university. Moreover, I use ranked data instead of data on promotion decisions, which allows for identification of effects over the whole range of candidates, instead of the marginal candidates who are the only ones harmed by discrimination in the context of promotions. A candidate who is always promoted or always rejected can also be affected by discrimination, but this effect would not be identified in promotion decisions, and relying on estimates from these studies could lead us to underestimate the size of gender preferences. Studying ranks on the other hand, allows us to estimate discrimination parameters no matter the ability or likelihood of being hired of the candidate.

One problem in identifying preferences is possible endogeneity in the assignment of candidates to jury, such as candidates not applying depending on the gender composition of the committee. Bagues *et al* (2017), Bagues & Esteve-Volart (2010), and Paola & Scoppa (2015) all use random assignment of candidates to committees to take care of this concern. I rely instead on variation across time and fields from a policy reform. The difference between the two approaches is relevant if we believe that random assignment in itself may increase or decrease the probability of observing biases,¹ especially since in most labour market contexts, evaluations do not rely on random assignments of jurors to candidates.

More generally, this article also contributes to the literature that considers the effect of gender

¹See Li (2017), for evidence of a bias/efficiency tradeoff when considering assignment of candidates to jurors in scientific funding.

quotas on women. Economists have mostly focused on analysing the effect of quotas in boardrooms or in the political arena. For instance, Ahern & Dittmar (2012) find that the 2003 reform in Norway, which imposed gender quotas in the boardroom of companies, had a negative impact on firm value. Bertrand et al. (2014) analyse the effect of the same reform on the labour market outcomes of women. They find that the reform has a positive effect on the wages of women in the boardroom, and that these women are more qualified than before, but that these gains do not trickle-down the wage ladder.¹ My article shows that quotas in academic hiring committees do not have a positive effect on the hiring of women, similar to the results of Bertrand *et al* (2014) in another labour market context.

In section 3.1, I discuss how own- or opposite-gender preferences could intervene in the ranking of committees, and how we can estimate them. In section 3.2, I explain how I construct the data set, and present some descriptive statistics. In section 3.3, I present results on the hiring and ranks of candidates. In section 3.4, I look at how the quota affected hiring, and discuss these results in section 3.5. I conclude in section 3.6.

3.1 MODEL

In the administrative data used in the paper, I observe a single ranking from each committee over all candidates. I consider that jury j has a latent utility function U_{ij} when hiring candidate i which is defined as follows:

¹From articles that look at quotas in politics, Casas-Arce & Saiz (2015) find that following a reform that imposed quotas on electoral lists in Spain, parties that previously had few amount of women candidates had a very large (6 percentage points) increase in vote shares. Beaman *et al* (2009) and (2012) show that in villages which had been randomly selected to have mandatory leadership positions for women in village councils, women were more likely to win local elections in the years that followed, and that voter attitudes towards women in leadership positions improved, as well as the aspirations and educational achievements of women in these villages.

$$U_{ij} = q_{ij} + \mu g_i g_j$$

where q_{ij} is how qualified the candidate is for a particular post, as evaluated by the committee j , and μ is a parameter that evaluates how committees rank different genders, when g_i is the gender of the candidate, and g_j the gender composition of the committee.¹ In this article I consider two different methods to recover the parameter of interest, μ , a random utility model and a rank-ordered logit. Although the rank-ordered logit directly takes ranks into account, there is no clear method for performing IV, which is why both methods are considered in this article.

3.1.1 RANK-ORDERED LOGIT

One way to estimate this model could be to regress the rank of the candidate using OLS on the independent variables. However, ranking is simply a preference ordering, and the distance between the 1st and 2nd rank might not be the same as between 2nd and 3rd. The standard method to recover parameters from ranking data in economics is the rank-ordered logit, which doesn't rely on the distance between ranks.

This method can be viewed as a multiplicative form of the traditional multinomial logit, and was proposed by Beggs et al. (1981), later refined in Hausman & Ruud (1987). The intuition is that respondents rank items based on repeated multinomial logit decisions, first choosing item i from M choices, then item i' from $M - 1$ choices and so forth, with the choice set decreasing in size until only one alternative is left. Combes et al. (2008), use this method to analyse whether connections matter for academic promotions, using ranking data from the French *agrégation* in economics, a context very similar to ours.

¹Since we have ranking data, I do not consider intransitive preferences, and do not study heterogeneity of utilities among committee members.

Going back to our proposition for the form of U_{ij} , consider the case where q_{ij} has the following form:

$$q_{ij} = x_{ij}\beta + \eta_j + \epsilon_{ij}$$

where x_{ij} is a vector with the characteristics of the candidates, η_j is a jury fixed effect and ϵ_{ij} is an error term. Consider two variables, q_{ij}^* , which is equal to the rank k given by jury j to alternative i , and r_{ij} , which is equal to the alternative i given the rank k by jury j . If ϵ_{ij} is distributed Type I extreme value, so that $Pr(\epsilon_{ij} < u) = e^{-e^{-u}}$, then the likelihood of observing a particular ranking in the j^{th} concours is:

$$Pr(q_{r_{1j}j} > q_{r_{2j}j} > \dots > q_{r_{Mj}j})$$

$$l_j(\beta) = \prod_{i:q_{ij}^*=1}^{M-1} \frac{\exp(x_{ij}\beta + \mu g_i g_j)}{\sum_{i':r_{i'j} \geq r_{ij}}^M \exp(x_{i'j}\beta + \mu g_{i'} g_j)}$$

Notice here that the jury fixed effects, η_j cannot be recovered, since they do not enter the likelihood function. We can only recover parameters that vary for candidates within jury. The log-likelihood function of J independent committees is then:

$$\sum_{j=1}^J \sum_{i:q_{ij}^*=1}^{M-1} x_{ij}\beta + \mu g_i g_j - \sum_{j=1}^J \sum_{i:q_{ij}^*=1}^{M-1} \left(\log \sum_{i':r_{i'j} \geq r_{ij}}^M \exp(x_{i'j}\beta + \mu g_{i'} g_j) \right)$$

The model can be estimated through maximum likelihood. Since the log likelihood is globally concave, β and μ will be the unique maximisers. We can recover the parameters μ by considering (for instance) the effect on the rank of women (relative to men) of having a female president (or a committee made up of many female members) or the effect for both genders of having a president

of the same gender.

3.1.2 RANDOM UTILITY MODEL

The first models using a latent utility framework date back to Thurstone (1927). In those early models, ranked data was analysed by pairwise difference. A part of the estimation will be devoted to this approach, since it allows for a more flexible approach to instrumental variable regressions.

We have a latent utility function where the candidate i has a better ranking than candidate i' if

$$U_{ij} \geq U_{i'j}$$

This implies

$$q_i \geq q_{i'} - \mu g_i g_j + \mu g_{i'} g_j$$

Assume that q_i , the candidate's intrinsic quality as evaluated by jury members depends on two factors: an observed term, x_i , and an unobserved term, $\epsilon_{ij} \sim N(0, \sigma)$ which represents other variables that the jury may take into account. If we represent the probability of candidate i receiving a better ranking than candidate i' by a binary variable Y , then:

$$Pr(Y = 1|x) = Pr(U(i) \geq U(i'))$$

$$Pr(x_i \beta + \mu g_i g_j + \epsilon_{ij} \geq x_{i'} \beta + \mu g_{i'} g_j + \epsilon_{i'j})$$

$$Pr(\beta(x_i - x_{i'}) + \mu(g_i - g_{i'})g_j \geq \epsilon_{i'j} - \epsilon_{ij})$$

$$= \Phi(X\beta + \mu Gg_j)$$

where $X = x_i - x_{i'}$ and $G = g_i - g_{i'}$. We can then use a probit regression to recover the

parameters. Of course, when $g_i = g_{i'}$, $G = 0$, so we can only estimate μ on pairs of candidates of different gender. Only dyads with candidates of different gender can give us information on own-gender preferences.¹

3.2 DATA

3.2.1 FRENCH HIRING COMMITTEES

French academic hiring committees are created *ad hoc* for each position that has to be filled. These committees have a jury president, who has broad powers over the nomination of committee members. The president also has a deciding vote in case of a tie between candidates. Once a committee has been created, candidates can apply *via* a web platform called GALAXIE,² and post their CVs. In some cases (but not all), they are aware of the jury composition at the time of their application.³

The committee then decides which candidates can be auditioned. Once these candidates are auditioned, the committee then makes a decision and ranks the candidates it deems worthy (if any) of being hired by the committee. If the first-ranked candidate refuses an offer, the position is then offered to the second-ranked candidate, then to the third-ranked candidate if the second also refuses, and so forth. All ranked candidates can potentially receive an offer from the university, if the candidates ranked above them refuse the offer. Therefore, if the committee decides that these candidates do not meet the requirements needed to work at the university some of the auditioned candidates are not ranked. The candidates are aware of the ranks that they receive, which helps

¹Simulations, presented in Section B.2 show that these estimates are unbiased for the data generating process presented above, and that the standard errors from OLS (if the parameter values are not too large) or probit regressions have the right power once we cluster at the committee level, to take correlation of errors (since we may use the same individual from the same committee multiple times in our sample) into account.

²Link to GALAXIE and available positions.

³We could be worried by auto-censure mechanisms, although if these mechanisms exist, they would reinforce our results. Paola & Scoppa (2015) find that women tend to drop out when applying for positions more often than men, but this effect is independent of the composition of the committee.

them to inform their decision on which offers they should accept or reject. These ranks may also give a signal to candidates on whether it is relevant to pursue an academic career. This is important in our context since we know that women are more likely to drop out of academia than men.

During the analysis, the ranking variable includes the K -ranked candidates, and the candidates that audition but are not ranked, who are all given the rank $K+1$. Since I only use order and not distance in my regressions, this normalisation is harmless. Candidates that are not auditioned are discarded for the regression analysis, but are included when computing the gender ratio of applicants.

3.2.2 DESCRIPTIVE STATISTICS

The dataset was compiled using administrative data on 196 hiring committees from 2 different French universities.¹ For each committee, I have access to the names and ages of the jurors and the names and ranks of the candidates. I do not know the behaviour of individual jury members, so I can only infer the behaviour of women in committees from two variables: the gender of the jury president, and the percentage of women in the hiring committee.

I present some statistics on these variables in Table 3.1. As we can see, the number of women presidents and the share of women in committees varies strongly between fields. In some fields, women presidents are few and far between: Only 1 out of 22 recruitments in Physics for instance, or 0 in economics. These are also fields where few women apply for jobs, which means that it could be difficult to properly estimate the impact of this variable. For instance, if we look at Table 3.2, we can see that male presidents hire women 34% of the time, compared to 38% of the time for female presidents, although Pearson's χ^2 test rejects that the groups are different. However, these raw figures do not take into account field fixed effects.

¹I have data from 2013 to 2016 for one university, and 2010 to 2017 for the other.

Table 3.1: Some descriptive statistics

	# recruitments	# of women recruitments	#Women presidents	% of women in the committee
Biology	38	20	16	49
Economics	11	2	0	36
History	15	5	7	39
Languages	32	14	15	52
Law	12	4	1	38
Maths	29	6	4	35
Physics	22	6	1	41
Political Science	16	5	4.5	33
Sociology	19	8	4	34
Total	196	70	52.5	41

Table 3.2: Recruitments and gender

President's gender:	Man recruited	Woman recruited	Total
Man	94	49.5	143.5
Woman	32	20.5	52.5
Total	126	70	196

Results include one co-presidency with two presidents of each gender. Pearson's chi-squared statistic is 0.294, with p-value=0.588.

Table 3.3: Descriptive statistics on hires and applicants

	% of women applicants	% of women recruited
Biology	52	53
Economics	21	15
History	34	33
Languages	54	44
Law	38	34
Maths	20	21
Physics	27	27
Political Science	30	31
Sociology	43	42
Total	38	36

Table 3.3 presents the percentage of women who apply for a position, by field. This variable is important, since the (average) percentage of women hired tracks the (average) percentage of women who apply. Since women are more likely to be presidents in fields (Biology and Languages) where women are more likely to apply, the raw correlation may not be informative. The percentage of women who apply to a contest is controlled for in the estimations.

As described in section 3.1, we will use data on pairs of female-male candidates to estimate a random-utility model. Since this data relies both on ranking, the gender composition of each contest and the number of contests, the composition in fields changes slightly compared to the data presented in Table 3.1.¹ To look at dyads makes sense, since we cannot know if there is a bias towards women in ranking if no women apply. Nearly half of the dyads come from Biology and Maths.

¹For instance, there are more dyads in sociology than physics (58 to 45) but less contests (19 to 22).

Table 3.4: Descriptive statistics on dyads

	# dyads FM	Dyads "won" by women
Law	4	1
Economics	27	9
History	41	16
Languages	110	39
Maths	114	53
Physics	45	27
Biology	100	42
Political Science	18	8
Sociology	58	21
Total	517	216

Dyads represent pairs of female-male candidates ranked by the same jury. The second column shows in how many dyads the woman received a better rank, by discipline.

ADDITIONAL DIMENSIONS

In the model, juries rely on the observed quality of the candidate x_i to make their decisions. Although I do not observe the CVs of candidates, I can control for two of the most important components of candidate quality, publications and connections. To control for publications, I collect the h-index and citations of candidates using the Scholar H-index calculator (2010), which creates h-indices and citation counts from Google Scholar. These variables will be used as measures, albeit imperfect,¹ of candidate quality.

Academic connections are another important potential confounder (see Combes et al. (2008)). In order to get more information on potential connections between jurors and candidates, I scraped the French website, theses.fr, which gathers information on which institution French doctors received their PhD from, as well as the name of their PhD supervisor. The data is harder

¹As pointed out by Bagues & Zinovyeva (2015), uninformed jurors are likely to only consider citation or h-index measures of candidates to determine whether the candidates are qualified or not, whereas more informed jurors will rely on information that can not be observed.

Table 3.5: Descriptive statistics on academic connections

	# Total Candidates	Candidates w/ supervisor in jury	Candidates w/ PhD from the institution	Candidates from Paris
Law	21	0	0	3
Economics	33	1	1	13
History	60	1	11	33
Languages	120	13	21	89
Maths	147	2	8	50
Physics	77	5	17	37
Biology	129	14	35	87
Political Science	41	2	7	18
Sociology	79	4	12	31
Total	707	42	112	351

Data on candidates' PhDs is scraped from online sources.

to collect for foreign PhD students, but I gathered information from other online sources on where they received their PhD.¹

The probability of having one's PhD supervisor in the committee is small (6%) but not negligible, rising to almost 11% in Biology. We can also look at endogamy, how often candidates with a PhD from the same institute apply for a position there. Once again, this is highly dependent on fields, with candidates in Biology (around 27%) very likely to have received a PhD from the same institution where they are applying for a position. We can also note that almost half of candidates received their PhD in Paris.

THE REFORM

In what follows, I first present results that look at the correlation between having more female evaluators and the rank of female candidates. Although one may be worried about endogeneity of assignment to evaluators, for instance if women specifically choose to apply only if there are many women in the committee, I control for the share of women applicants, which should eliminate

¹Most juries are made up almost exclusively of French professors, so the probability of students who have received a PhD abroad of being in a committee with their PhD supervisor is very small.

Table 3.6: Effect of the reform on the mean share of women jurors and number of female presidents

Discipline	Mean share of women jurors		Share of female presidents	
	Pre-reform	Post-reform	Pre-reform	Post-reform
Law	0.36	0.43	0.13	0
Economics	0.26	0.48	0	0
History	0.37	0.47	0.42	0.66
Languages	0.54	0.49	0.60	0.35
Maths	0.31	0.37	0.18	0.11
Physics	0.32	0.48	0	0.08
Biology	0.49	0.48	0.35	0.5
Political Science	0.25	0.5	0.27	0.2
Sociology	0.3	0.50	0.27	0

this concern. Nevertheless, I use the 2015 reform as an instrument, which should take care of other endogeneity concerns.

Since the 1st January 2015, each committee must be made up of at least 40%¹ of each gender. This reform affected fields in different ways. I present the average female/male juror ratio in Table 3.6 for each discipline before and after the reform. We can see a clear increase in the proportion of women sitting on committees due to the reform, except for the disciplines where the quotas were already respected, Biology and Languages. I use fields where the average proportion of women in committees pre-reform was less than 0.35 as an instrument for the increase in the feminisation of committees.² There does not seem to be a systematic effect on the gender of the jury president however.

¹For maths, this ratio was dropped exceptionally to 14% and to 33% for Political Science until 2016 (Link to the decree). In other fields, this ratio is binding but once the committee has been approved, if a committee member drops out, the replacements do not need to respect the gender quota. The data includes the final composition of the committee.

²The results are robust to using other cutoffs.

Table 3.7: Effect of the reform on the mean h-index of jurors

Discipline	Mean h-index of jurors	
	Pre-reform	Post-reform
Economics	15.06	10.96
History	8.86	8.33
Languages	4.03	4.67
Law	8.39	8.21
Maths	10.27	11.60
Physics	13.50	11.44
Biology	12.49	10.64
Political Science	18.36	12.51
Sociology	14.14	13.09

COMMITTEE MEMBERS

The reform also affects jury composition in other dimensions. For instance, due to the additional constraints necessary to fill the quota, the quality of evaluators may decrease, which may in turn positively or negatively affect the likelihood that a woman is highly ranked. In order to test this effect, I collected data on the h-index of jury members, and present in Table 3.7 how the reform affected the mean h-index of jurors.

3.3 ESTIMATION

I estimate the two models from section 3.1 on my dataset. I look at results without taking into account possible endogeneity of assignment of candidates to committees, then present results using the reform.

3.3.1 RESULTS ON THE FIRST-RANKED CANDIDATE

First I look at the effect of two gender variables, the gender of the committee president and the proportion of females in the committee on the probability that the first-ranked candidate is a woman, using a probit regression. Since in our dataset, most candidates who are ranked first accept the offer, this tests directly for whether more women are hired when there are more women in the committee. The results are in Table 3.8. The first column has no controls, and there is a significant, positive effect of having more women in the committee in this case. However, once we control for the percentage of women candidates, this effect vanishes, and the coefficient becomes negative and insignificant. I add further controls for the mean age of committee members, discipline fixed effects and a dummy for whether the post is for an assistant or associate professorship.¹ None of these further controls are statistically significant.

Although we can likely rule out a significant positive effect of the feminisation of committees on the probability of women being hired, this doesn't preclude finding an effect on the *ranking* of women. Indeed, even if gender had no effect on the committee's decision, we could still find negative effects if better male candidates apply to committees which are mostly composed of women. In the next section we look directly at the rankings of candidates, which enables us to control more directly for the suitability of each applicant for the post.

3.3.2 RESULTS ON RANKS

RANK-ORDERED LOGIT

In this section I analyze ranks using the rank-ordered logit. With this method I can only control for variables that affect ranking within committees. For all non-individual variables, I consider

¹Maitre de Conférences or Professeur des Universités in French.

Table 3.8: Correlation of the gender of jurors on the probability of women being first-ranked

Dependent variable:	Gender(W) of the first ranked candidate			
Estimator:	Probit	Probit	Probit	Probit
Marginal effects	(1)	(2)	(3)	(4)
Gender (W) of president	0.006 (0.08)	-0.036 (0.07)	-0.038 (0.08)	-0.037 (0.08)
% of women in the committee	0.607** (0.26)	-0.042 (0.25)	0.001 (0.35)	0.056 (0.35)
% of women candidates		0.878*** (0.10)	0.768*** (0.26)	0.787*** (0.26)
Mean age of the committee members				0.015 (0.01)
Professorship				-0.003 (0.10)
Discipline fixed effects	No	No	Yes	Yes
University FE	No	No	Yes	Yes
# observations	196	196	162	162

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. The gender variable is coded as 1 for a woman and 0 for a man. Contests with only members of a specific gender are dropped in columns three and four.

their effect on the rank of women relative to men. Similar to the results above, we cannot estimate these parameters when only men or only women applied. Therefore, we have less observations than in the descriptive statistics presented in section 3.2.

A first interesting result is to look at the effect of being a woman on the rank given by the committee. This effect is negative, though insignificant once we control for the h-index of candidates. Next, consider the interaction between how women are ranked, and the gender composition of the committee. There is a significant and large negative effect of the share of women in the committees on the ranking of women in columns 2 and 3, where the only control is the h-index of candidates. However, once more control variables are added, this effect becomes insignificant, although the point estimate remains high. One way to interpret the estimate is to compare it to the coefficient that controls for publications one row below. The effect of having a jury made

Table 3.9: Correlation of the gender of jurors and the rank of candidates

Dependent variable:	Rank			
Estimator:	Logit	Logit	Logit	Logit
	(1)	(2)	(3)	(4)
Woman in a jury with a female president		0.131 (0.25)	0.162 (0.25)	0.203 (0.25)
More women in a jury for women		-0.592** (0.27)	-0.491* (0.27)	-0.834 (0.95)
H-index of the candidate	0.04*** (0.01)		0.042*** (0.01)	0.043*** (0.01)
Candidate is a woman	-0.15 (0.10)			0.439 (0.62)
Effect of having more women candidates				-0.519 (0.68)
PhD supervisor in the committee				0.150 (0.21)
PhD from a Parisian university				-0.100 (0.14)
PhD from the same institute				0.208 (0.17)
PhD from a foreign university				-0.031 (0.18)
Fixed Effects	No	No	No	No
# observations	648	648	648	648

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. The gender variable is coded as 1 for a woman and 0 for a man. The h-index variable can be interpreted as the effect of increasing the h-index for all candidates. Other variables must be interpreted as the differential effect of the variable for women with respect to men.

up of 10% more women in this case is comparable to having a h-index that is one point lower for women. The control variables are mostly insignificant (bar the h-index), but they have the expected sign: for instance, having a supervisor in the committee, or having a PhD from the institute where you are applying increases your ranking by the committee.

RANDOM UTILITY MODEL

Next I use the method with dyads to estimate whether own-gender preferences exist. I use a probit regression, and cluster at the jury level to take correlation between errors for dyads from the same jury into account. In the first column I only consider the two possible gendered variables that could affect the probability that a woman is ranked higher than a man: the number of women in the committee, and the gender of the president. I add discipline fixed effects in the second column. The percentage of women is significant and negative, though only at the ten percent level in the second column. I add more control variables in the next two columns, like the h-index difference, data on where the candidate received his thesis, and the mean-age of the committee members. Older committee members seem to favour women, and candidates with more publications are more likely to have a better rank, but the other variables are not significant. The point estimates indicate the percentage of women in the committee still has a negative effect, but is non-statistically significant once we add all the controls.

In Table 3.II, the sample is split into two different categories: STEM (Biology, Physics and Maths) and non-STEM, since a lot of the debate around the under-representation of women is about STEM fields (e.g. Nimmesgern, 2016). Buser et al. (2014) find that men are more likely to major in STEM fields than women with equivalent grades. There is also some evidence that women in France receive better evaluations in fields where they are under-represented.¹ I find that

¹e.g. Breda & Ly (2015) and Breda & Hillion (2016).

Table 3.10: Correlation of the gender of jurors and the probability that a woman is better ranked within a dyad

Dependent variable:	Probability that W is more highly ranked			
Marginal effects	(1)	(2)	(3)	(4)
Gender (W) of president	0.021 (0.07)	0.035 (0.07)	0.016 (0.07)	0.036 (0.07)
% of women in the committee	-0.539** (0.27)	-0.486* (0.29)	-0.477* (0.27)	-0.375 (0.27)
% of women candidates			-0.204 (0.24)	-0.256 (0.23)
H-index difference			0.017*** (0.00)	0.018*** (0.00)
Professorship			0.100 (0.06)	-0.059 (0.08)
Candidate (W) has a Phd from the same institute				0.014 (0.05)
Candidate (W) has a Phd from Paris				-0.033 (0.05)
Candidate (W) has a foreign PhD				-0.037 (0.07)
Candidate(W)'s PhD advisor is in the committee				0.034 (0.08)
Mean age of the committee members				0.027*** (0.01)
Discipline fixed effects	No	Yes	Yes	Yes
Establishment fixed effects	No	No	Yes	Yes
# observations	517	517	517	517

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Dyads represent pairs of female-male candidates ranked by the same jury. The gender variable is coded as 1 for a woman and 0 for a man. Standard errors are clustered at the jury level

the point estimate of having more women in the jury is negative in both STEM and non-STEM, but is stronger in STEM fields and insignificant once we control for discipline fixed effects.

3.3.3 INVESTIGATING ENDOGENEITY CONCERNS

One important issue is that the selection of candidates both into the candidate and the audition pool is possibly endogenous. For instance, candidates may put less effort in their application depending on the composition of the jury, or even decide not to apply at all. I address this concern by considering whether the composition of the jury affects variables other than the ranking within

Table 3.11: Correlation of the gender of jurors and the probability that a woman is better ranked within a dyad. Split by subject

Dependent variable:	Probability that W is more highly ranked			
Marginal effects	(1)	(2)	(3)	(4)
Gender (W) of president	0.050 (0.11)	-0.001 (0.09)	0.094 (0.09)	-0.026 (0.09)
% of women in the committee	-0.951** (0.47)	-0.083 (0.30)	-0.615 (0.45)	-0.170 (0.31)
% of women candidates	0.098 (0.29)	-0.317 (0.22)	0.315 (0.50)	-0.355 (0.25)
H-index difference			0.023*** (0.01)	0.013** (0.01)
Candidate (W) has a Phd from the same institute			0.051 (0.07)	-0.014 (0.07)
Candidate (W) has a Phd from Paris			-0.057 (0.06)	0.003 (0.06)
Mean age of the committee members			0.020** (0.01)	0.022** (0.01)
Discipline	STEM	Social Science	STEM	Social Science
Discipline fixed effects	No	No	Yes	Yes
Establishment fixed effects	No	No	Yes	Yes
# observations	259	258	259	258

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Dyads represent pairs of female-male candidates ranked by the same jury. The gender variable is coded as 1 for a woman and 0 for a man. Standard errors are clustered at the jury level

Table 3.12: Regression of the gender composition of applicant pools on the gender composition of the recruitment committee

Dependent variable:	Gender composition of candidates			
Marginal effects	(1)	(2)	(3)	(4)
% of women in the committee	0.922*** (0.16)	0.715*** (0.17)	0.730*** (0.18)	0.468** (0.18)
Gender (W) of president	0.108** (0.05)	0.119** (0.05)	0.048 (0.05)	0.045 (0.05)
Discipline fixed effects	No	No	Yes	Yes
# observations	196	162	196	162

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. The gender variable is coded as 1 for a woman and 0 for a man. Standard errors are clustered at the jury level. Columns 2 and 4 drop contests with only male or female applicants.

field. In the following tables, I present results both with contested and uncontested (only candidates of a single gender applied) committees. In Table 3.12, I regress the gender ratio of candidates on the gender ratio of jury members, while controlling for discipline fixed-effects.

The applicant pool depends significantly on the gender composition of the committee, either through sub-discipline effects (i.e. sub-disciplines within each field have different gender compositions) or through endogenous selection of candidates (candidates observe the gender composition of the committee and then decide whether to apply or not). This implies that the results above could be endogenous, since candidate pools may have selection bias on unobservable candidate quality. In the following section, I try to solve this endogeneity concern by using variation from a policy reform.

3.4 EFFECT OF THE QUOTA

In this section, I look at the effect of the gender quota reform of 2015, which is used as an exogenous shock to analyze the effect of the share of women in the committee on the ranking of women.¹ I use a diff-in-diff, and an instrumental variable regression using a linear probability

¹Results on the hiring of women, i.e. on the first ranked candidates, are not significant, and using a Wald test, I cannot reject the null hypothesis of no endogeneity. Results are presented in Appendix B.3.

Table 3.13: Effect of the reform on the candidate pool

Dependent variable:	Gender composition of candidates	
Marginal effects	(1)	(2)
Post-Reform	0.07 (0.06)	-0.00 (0.07)
Group	Treatment Group	Control Group
# observations	99	97

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. The gender variable is coded as 1 for a woman and 0 for a man.

model¹ for the random utility model.

I use the reform to construct a control and a treatment group: the group variable is the set of disciplines with a ratio of women in the jury below 35% before the reform in 2015.² To check whether this approach is valid, I test whether the gender composition of the applicants changes significantly before and after the reform in both groups. As can be seen below in Table 3.13, there are no significant differences in the gender composition of applicants after the reform in both the control and treatment groups.

The results using the rank-ordered logit are presented in Table 3.14. As we can see, there is a strong negative effect on the rank of women of being in a field that had a strong increase in the number of women jurors due to the reform. This effect is robust to the inclusion of other covariates, such as the h-index of candidates or the effect of having a PhD from the same institute. Table B.11 in the Appendix includes even more covariates and interaction terms, and the results there are also significant. The mean h-index of jurors does not seem to have any sizeable effect on the ranks of women.

The method of dyads allows us to estimate the model through IV, instrumenting for the potentially endogenous variable with the effect of being in our group post reform. The first-stage

¹The coefficients from probit regressions are biased whenever the errors are heteroscedastic, which could be the case here since the instrument is binary. I present results using probit regressions in Appendix B.7.

²Economics, Maths, Political Science, Physics, Sociology

Table 3.14: Effect of the quota on the rank of female candidates

Dependent variable:	Rank		
Estimator:	Logit	Logit	Logit
	(1)	(2)	(3)
Post Reform*Group	-1.173*** (0.44)	-1.358*** (0.47)	-1.326*** (0.48)
Woman in a jury with a female president	0.006 (0.25)	0.071 (0.25)	0.103 (0.26)
Post-reform	0.615** (0.31)	0.697** (0.31)	0.677** (0.33)
Group	0.565** (0.26)	0.297 (0.26)	0.265 (0.28)
Candidate is a woman	-0.482*** (0.17)	0.353 (0.44)	0.302 (0.49)
H-index of the candidate		0.041*** (0.01)	0.042*** (0.01)
Effect of having more women candidates on women's rank		-1.559* (0.83)	-1.584* (0.84)
PhD supervisor in the committee			0.099 (0.22)
PhD from the same institute			0.197 (0.16)
PhD from a Parisian university			-0.095 (0.12)
# observations	648	648	648

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. The gender variable is coded as 1 for a woman and 0 for a man. The h-index variable can be interpreted as the effect of increasing the h-index for all candidates. Other variables must be interpreted as the differential effect of the variable for women with respect to men.

Table 3.15: IV estimate of the effect of the increase in women jurors on the probability that a woman is better ranked within a dyad

Dependent variable:	Rank		
Estimator:	IV	IV	IV
Marginal Effects	(1)	(2)	(3)
% of women in the committee	-2.496** (0.99)	-2.207** (0.99)	-2.422*** (0.91)
Gender (W) of president	0.059 (0.08)	0.062 (0.08)	0.065 (0.08)
Group	-0.158 (0.11)	-0.123 (0.11)	-0.200* (0.10)
Reform	0.143* (0.08)	0.131* (0.08)	0.140* (0.08)
H-index difference		0.009* (0.00)	0.010** (0.00)
Candidate (W) has a Phd from the same institute		0.023 (0.05)	0.034 (0.05)
Candidate(W)'s PhD advisor is in the committee		-0.006 (0.09)	0.001 (0.09)
Mean age of the committee members		0.007 (0.01)	0.008 (0.01)
% of women candidates			-0.264 (0.25)
# observations	517	517	517

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Dyads represent pairs of female-male candidates ranked by the same jury. The gender variable is coded as 1 for a woman and 0 for a man. Errors are clustered at the contest level.

estimates are presented in the appendix.¹ Being in the group affected by the reform has a highly significant and positive effect on the share of women in committees, as expected. The instrumental variable results are presented in Table 3.15.² Once again, the share of women in the committee has a significant negative effect.

The results can be considered as a Local Average Treatment Effect (LATE) i.e. the effect of increasing the share of women jurors in disciplines that have few women. These effects may be stronger or weaker than the ones we would find when analysing all fields. However, the measure

¹The F-test statistic for the first-stage is above 10 in all three cases, as advised by Staiger & Stock (1997), and the instrument is significantly correlated with the endogenous variable.

²Of course, one could simply look at the probability of a woman being better ranked than a man after the reform in the disciplines affected, using a difference-in-difference approach rather than instrumental variables. Table B.5 in the Appendix provides results using this methodology.

that we are considering is the relevant one, since the policy proposals that are usually debated aim at increasing the proportion of women evaluators in fields where they are not represented, not in fields where parity is already achieved. Whether increasing the proportion of women evaluators also increases the rank of women in fields where parity is already respected is a moot point. In fact, the global effect of the reform, presented in Appendix B.6, is close to 0.

3.4.1 TOP RANKS ONLY

One concern in the approach above is that some committees may not care about how they rank lower quality candidates, since they are very unlikely to be hired. To investigate this concern, I present results in Table 3.16 using the same methodology as that used for Table, but dropping all ranks below 3. In this specification, the reform still has a significant effect once the h-index of candidates is controlled for. In this specification, the gender of the jury president now has a negative effect on the ranking of women.

3.5 DISCUSSION

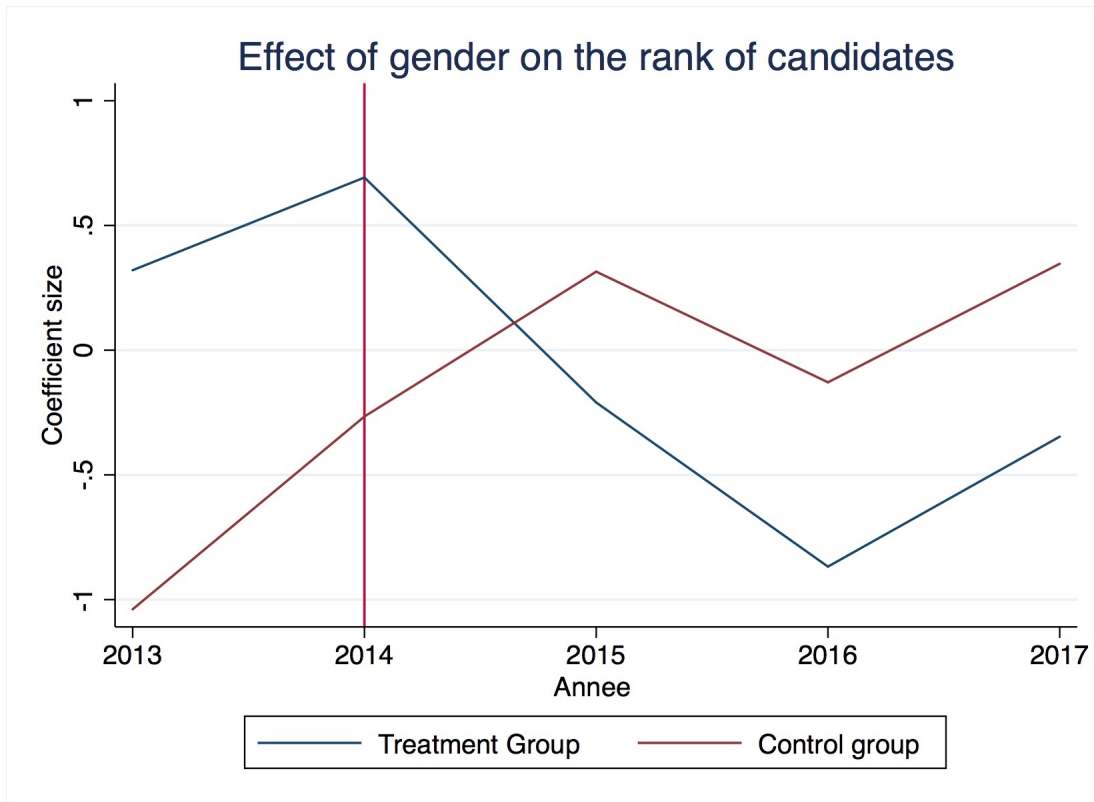
I provide a graphical interpretation of the results with the instrument in Figure 3.1. In this Figure, I plot the coefficients from a rank-ordered logit of being a woman, in the two years before and after the reform. The treatment group consists of recruitments in disciplines that are affected by the reform, while the control group is made up of disciplines that were already respecting gender parity in committees, and should therefore not be affected by the reform, as explained above. In disciplines affected by the reform, the rank of women decreases, whereas women are ranked slightly higher in the control group. This does not seem to be driven by outliers. In Figure 3.2, I look at the effect of the reform on the ranks of women by discipline, with disciplines sorted

Table 3.16: Rank-Ordered Logit using the quota: top 3 ranks only

Dependent variable:	Rank		
Estimator:	Logit	Logit	Logit
	(1)	(2)	(3)
Post-Reform*Group	-0.870 (0.58)	-1.044* (0.56)	-1.144** (0.56)
Woman in a jury with a female president	-0.683** (0.31)	-0.670** (0.30)	-0.670** (0.31)
Group	-0.075 (0.44)	0.005 (0.42)	0.058 (0.43)
Reform	0.506* (0.28)	0.280 (0.31)	0.209 (0.32)
Candidate is a woman	-0.135 (0.21)	0.546 (0.52)	0.459 (0.56)
H-index of the candidate		0.025* (0.01)	0.025* (0.01)
Effect of having more women candidates		-1.279 (0.94)	-1.398 (0.95)
PhD supervisor in the committee			0.546* (0.28)
PhD from the same institute			-0.046 (0.20)
PhD from a Parisian university			-0.036 (0.15)
Mean H-index of jurors			0.015 (0.03)
Discipline fixed effects	No	No	No
# observations	470	470	470

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Only the first three ranked candidates are included
The gender variable is coded as 1 for a woman and 0 for a man. The h-index variable can be interpreted as the effect of increasing the h-index for all candidates. Other variables must be interpreted as the differential effect of the variable for women with respect to men.

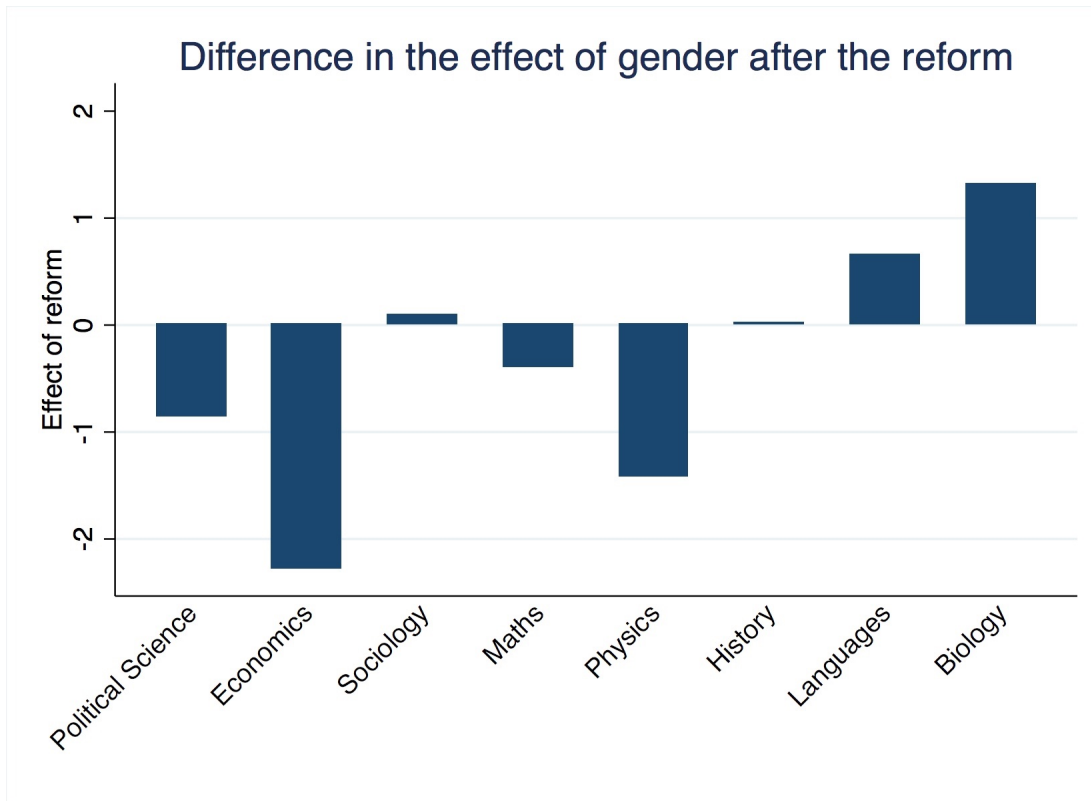
Figure 3.1: Rank-ordered logit on gender of candidates by year



from left to right according to their pre-reform average share of female jurors. The figure bears out what we see in the regression: disciplines that had few women jurors, and thus were more affected by the quota seem to be less favourable to women after the reform, while disciplines that were already respecting the quota (Biology and Languages), seem to have been positively affected by the reform.

One issue in interpreting these results is that we cannot identify what is driving this effect. One possibility is that women in positions of authority have opposite-gender preferences, a phenomenon that has already been analysed in the sociology literature under the expression "Queen

Figure 3.2: Effect of the reform on the ranks of women by discipline. Disciplines most affected by the reform are ordered from left to right.



Bee” syndrome.¹ An alternative to the Queen Bee syndrome is that women are penalised when they decide to promote women, as discussed in Johnson & David (2016). The effect that we see is then due to women internalising the retribution they could face from male colleagues if they promote women, rather than opposite-gender preferences. Akerlof & Kranton (2000) postulate instead that men discriminate against women when their identities are threatened. It is possible that this is the case in our setting, since having more women in what were traditionally masculine settings could be seen as a threat to the masculinity of the jurors. The interpretation of our effects in this case would not be of opposite-gender bias in recruiting, but of own-gender bias from men

¹see e.g. Staines et al. (1974).

Table 3.17: Effect of the reform by gender of the jury president

Dependent variable:	Rank	
Estimator:	Logit	Logit
	(1)	(2)
Post Reform*Group	-1.886*** (0.51)	0.987 (0.77)
Post-reform	0.867** (0.34)	0.552 (0.59)
Group	0.280 (0.30)	0.402 (0.57)
Candidate is a woman	0.861* (0.50)	-1.733 (1.13)
H-index of the candidate	0.051*** (0.01)	0.018 (0.02)
Effect of having more women candidates	-2.641*** (0.92)	2.477 (1.93)
# observations	476	169

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. The gender variable is coded as 1 for a woman and 0 for a man. Column 1 is for committees with a male president, and column 2 for committees with a female president. The h-index variable can be interpreted as the effect of increasing the h-index for all candidates. Other variables must be interpreted as the differential effect of the variable for women with respect to men.

that appears only when the gender identities of jurors are threatened. Our results are consistent with all three hypotheses.

However, an element that leans towards the third hypothesis is presented in Table 3.17. In this Table, I split the sample into two groups: those that have a female or a male president, and look at the impact of the reform across these two groups. The negative effect of the reform is driven entirely by committees that are helmed by men. This is consistent with the idea that men retaliate against women following the reform. Irrespective of the channel that causes this effect, the results from the above section strongly suggests that gender quotas and other coercitive measures may have unintended effects. In this specific case, increasing the proportion of female jurors has led to a negative effect on the ranking of women by committees.

3.6 CONCLUSION

In this article, I have showed that imposing gender quotas in recruitment committees does not help the recruitment of women, and may actually harm their careers. It is unclear whether this effect is driven by women being less likely to favour women in hiring, or through men voting against women as a reaction to the quota, though there is some evidence in favour of the second channel. The results do not preclude there being discrimination against women in hiring,¹ but they do suggest that there may not be an easy solution to the problem of under-representation of women in academia.

¹However, the coefficient for the ranking of women in Table 3.9 with the most controls (column (4)) is positive and insignificant.

B

Gender discrimination in recruitment committees

B.1 SEMIPARAMETRIC ESTIMATION

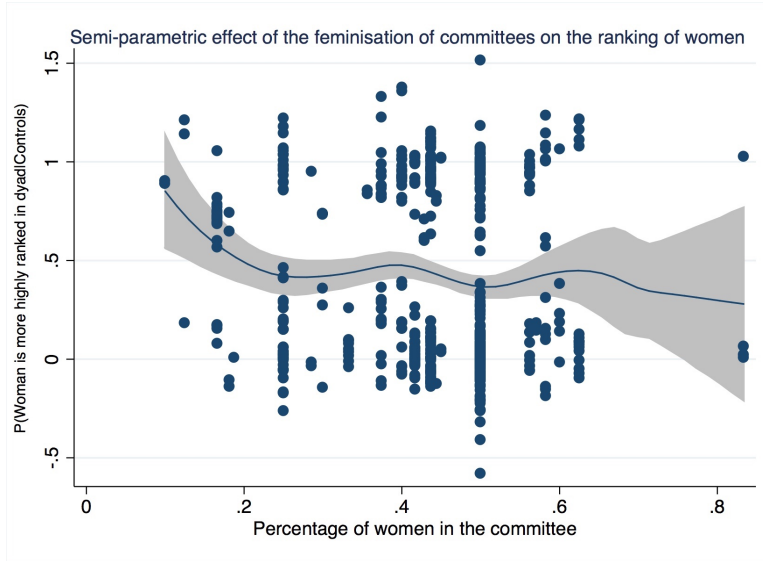
One hypothesis from Paola & Scoppa (2015) was that the effect of the feminisation of committees on the likelihood of being promoted was non-linear. In their paper, they found that there was a positive effect on the probability of a woman being promoted when committees moved from having no women to having 1 woman, and no further effect thereafter. Although we do not observe committees with no women in our dataset, we can check whether assuming a linear effect of feminisation of the committee is correct.

Consider the following partially linear model,

$$W_{ij} = X\beta + m(Gg_j) + \epsilon_{ij}$$

where W_{ij} is a binary variable, equal to 1 if a women is better ranked than a man for a specific dyad, X , β and G are defined in section 3.1 and m is a function that we wish to estimate. Using Robinson (1988)'s double residual estimator, we can first estimate β consistently, then recover $m(Gg_j)$ non-parametrically. As can be seen below, from a semi-parametric regression that controls for the h-index of candidates, the effect of the feminisation of committees on the ranking of women does not seem to be non-linear.

Figure B.1: Semi-parametric estimation of the relationship of interest



Using Hardle & Mammen (1993)'s test, that relies on square differences between a linear estimator and the estimated non-parametric function, we cannot reject the null hypothesis that the linear form and the non-parametric estimates are not different.

B.2 DIFFERENCES BETWEEN THE RANDOM UTILITY MODEL AND THE RANK-ORDERED LOGIT

Are the two methods presented in the article equivalent? Consider the following data generating process. I generate a jury utility variable U_{ij} such that:

$$U_{ij} = \beta q_{ij} + \mu g_i g_j + \epsilon_{ij}$$

where g_i and g_j are binomial variables, q_{ij} is normally distributed with mean 5 and standard deviation of 1. Using the values from the article, I use 0.05 for β and -0.6 for μ .

From this utility variable I then simulate artificial committees in the following way: from a 1000 observations of U_{ij} I create 100 committees with 10 candidates in each committee. 60 committees have a female jury president. Each committee then creates a ranking r_{ij} of each candidate based on the value of U_{ij} .

This simulation allows me to answer two different questions: do both methods have the right power, and can both methods accurately recover the μ/β ratio?

Let us consider the first question. The right standard errors would give us a 5% probability of having a t-statistic above the absolute value of 1.96, if the true effect was in fact 0. I simulate the DGP described above 500 times, with a true effect of gender bias, $\mu = 0$. I present results both with ϵ_{ij} either normally distributed or type I extreme value in Table B.1. In all cases, errors are clustered at the contest level, since candidates evaluated by the same jury are compared multiple times, generating correlation between errors.

The methods that use the random utility approach are very similar in terms of power and estimates. They seem to over-reject when errors are distributed Extreme Value, and under-reject

Table B.1: Power of estimation methods

Estimator	RO Logit	OLS	Probit	Logit
<i>Normal errors</i>				
Mean	-0.03	-0.00	-0.00	0.00
SD	0.11	0.04	0.04	0.04
# of t-statistics $\leq 1.96 $ (%)	947 (95%)	941 (94%)	941 (94%)	941 (94%)
<i>Type I Extreme Value</i>				
Mean	0.00	0.00	0.00	0.00
SD	0.11	0.04	0.04	0.04
# of t-statistics $\geq 1.96 $ (%)	943 (94%)	960 (96%)	960 (96%)	959 (96%)

Results from 1000 simulations. True effect is 0. For the right amount of power, tests should reject the null hypothesis only 5% of the time. Results from the probit and logit estimations are marginal effects.

when errors are distributed normally. However, the power estimates are within 1% of the valid rejection rates. The rank-ordered logit has a wider standard deviation of estimates, but is closer to the valid rejection rates.

In Table B.2, we can see the mean values of the estimators for the parameters presented above. I present the median value for the μ/β ratio, since the mean is affected by outliers, i.e. when β is estimated to be very close to 0. The mean point estimates are similar for the rank-ordered logit, probit and OLS. The results are similar no matter the method considered.

This is the case when the values of the parameters are small. However, with larger values of β and μ , the LPM suffers from bias, as can be seen in Table B.3. This well known result stems from the predicted values of Y , $X\beta + \mu G_{ij}$, being outside the range of possible values that can be taken on by a binary variable, i.e. 0-1.

B.3 PROBIT AND HETEROSCEDASTICITY

In this section I present results using an instrument variable probit. Although the endogenous variable, the share of women in the jury, is continuous, the instrument considered is binary which

Table B.2: Estimates from simulations

Estimator	RO Logit	OLS	Probit	Logit
<i>Gender bias</i>				
Mean	-0.61	-0.15	-0.38	-0.60
SD	0.11	0.04	0.10	0.17
<i>Effect of quality</i>				
Mean	0.05	0.01	0.03	0.05
SD	0.04	0.01	0.03	0.04
<i>Ratio</i>				
Median	-0.086	-0.086	-0.85	-0.085
SD	0.07	0.10	0.10	0.10

Means and standard deviation of estimators of μ and β from 1000 simulations. Error term is distributed type I EV. Results are similar with normally distributed errors. Results from the probit and logit estimations are full effects and not marginal. The marginal effects are similar in size to the OLS results.

Table B.3: Estimates from simulations: large parameters

Estimator	RO Logit	OLS	Probit	Logit
<i>Gender bias</i>				
Mean	-6.01	-0.40	-3.33	-6.04
SD	0.30	0.02	0.20	0.36
<i>Effect of quality</i>				
Mean	3.00	0.22	1.66	3.01
SD	0.13	0.01	0.08	0.14
<i>Ratio</i>				
Median	-0.50	-0.55	-0.50	-0.50
SD	0.02	0.02	0.02	0.02

Means and standard deviation of estimators of μ and β from 1000 simulations. Error term is distributed type I EV. Results are similar with normally distributed errors. Results from the probit and logit estimations are full effects and not marginal. The marginal effects are similar in size to the OLS results.

may introduce heteroscedasticity. Under heteroscedasticity, the results from probit regressions are biased, which is why a LPM is used in the main text. The marginal effects are significant, but slightly smaller in magnitude compared to those in the main text.

Table B.4: Probability that W is more highly ranked

Dependent variable:	Rank		
Estimator:	Probit	Probit	Probit
Marginal Effects	(1)	(2)	(3)
% of women in the committee	-1.948*** (0.44)	-1.939*** (0.44)	-1.892*** (0.43)
Gender (W) of president	0.048 (0.06)	0.050 (0.06)	0.050 (0.06)
Reform	0.114** (0.06)	0.115** (0.06)	0.111** (0.06)
Group	0.120* (0.07)	0.115* (0.07)	0.164** (0.07)
Candidate (W) has a Phd from the same institute		0.028 (0.04)	0.026 (0.04)
Candidate(W)'s PhD advisor is in the committee		-0.014 (0.07)	-0.004 (0.07)
H-index difference			0.009* (0.00)
% of women candidates			-0.182 (0.20)
# observations	517	517	517

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Dyads represent pairs of female-male candidates ranked by the same jury. The gender variable is coded as 1 for a woman and 0 for a man. Errors are clustered at the contest level.

B.4 DIFFERENCE IN DIFFERENCE

I present results in Tables B.5 and B.6 that are similar to the graphical interpretation in the text. Instead of instrumenting the share of women in the jury, I rely on a difference in difference analysis, comparing variation in the probability of women being hired before and after the reform in treatment and control group. Being in the treatment group after the reform significantly decreases the probability of a woman being hired, even when controlling for H-index differences and connections. The point estimates found using probit and OLS are extremely similar.

Table B.5: Probit results, Difference-in-Difference

Dependent variable:	Rank		
Estimator:	Probit	Probit	Probit
Marginal effects	(1)	(2)	(3)
Post Reform*Group	-0.352*** (0.11)	-0.312*** (0.12)	-0.374*** (0.11)
Gender (W) of president	0.008 (0.08)	0.009 (0.08)	0.005 (0.08)
Reform	0.212*** (0.08)	0.191** (0.08)	0.220*** (0.07)
Group	0.247*** (0.09)	0.208** (0.09)	0.116 (0.09)
H-index difference		0.012** (0.00)	0.013*** (0.00)
Candidate (W) has a Phd from the same institute		-0.016 (0.05)	0.001 (0.05)
Candidate(W)'s PhD advisor is in the committee		-0.008 (0.08)	0.003 (0.08)
% of women candidates			-0.535*** (0.19)
# observations	517	517	517

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Dyads represent pairs of female-male candidates ranked by the same jury. The gender variable is coded as 1 for a woman and 0 for a man. Errors are clustered at the contest level.

B.5 FIRST STAGE RESULTS

B.6 FULL EFFECT OF THE REFORM

In this section I look at the direct effect of the reform without considering heterogeneity in treatment effects across disciplines, using only a dummy variable for the effect of the reform. With this methodology, the net effect of the reform seems to be close to 0.

B.7 ADDITIONAL TABLES

In Table B.9, I present results with additional variables, such as the effect of the mean age and mean h-index of the jurors. I also add interaction terms between these variables and the reform as well as discipline fixed effects. The age of the jurors has a significant effect: older juries seem to

Table B.6: OLS results, Difference-in-Difference

Dependent variable:	Rank		
Estimator:	OLS	OLS	OLS
	(1)	(2)	(3)
Post Reform*Group	-0.356*** (0.12)	-0.320** (0.12)	-0.382*** (0.12)
Reform	0.211*** (0.08)	0.190** (0.08)	0.219*** (0.07)
Group	0.245*** (0.09)	0.208** (0.09)	0.117 (0.09)
H-index difference		0.010** (0.00)	0.011*** (0.00)
Candidate (W) has a Phd from the same institute		-0.018 (0.05)	-0.003 (0.05)
Candidate(W)'s PhD advisor is in the committee		-0.009 (0.08)	0.004 (0.08)
% of women candidates			-0.529*** (0.19)
# observations	517	517	517

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Dyads represent pairs of female-male candidates ranked by the same jury. The gender variable is coded as 1 for a woman and 0 for a man. Errors are clustered at the contest level.

prefer female candidates, however the other variables do not have a significant effect. The effect of gender is negative, but this seems to be driven by the interaction with the effect of the age of jurors.

In Table B.10, I use the reform to directly estimate the effect of the reform on hiring, using the same approach as in the main part of the text, but on the gender of first-ranked candidates rather than on dyads. However, in this specification, I do not control for the connections or the h-index of candidates. The effect on hiring is positive but insignificant. The test statistics that look at the correlation between the residuals in the IV probit and the reduced equation for the endogenous variable are insignificant, indicating that endogeneity may not be a problem in our case.

In Table B.11, I use the same specification as in Table 3.14, but include other measures of publications that could be more relevant than the h-index. I include variables with age- and author-

Table B.7: First stage: IV

Dependent variable:	Rank		
Estimator:	IV	IV	IV
	(1)	(2)	(3)
Post Reform*Group	0.141*** (0.04)	0.141*** (0.04)	0.160*** (0.04)
Gender (W) of president	0.022 (0.02)	0.023 (0.02)	0.025 (0.02)
Post-Reform	-0.028 (0.02)	-0.027 (0.02)	-0.035 (0.02)
Pre-reform*Group	-0.162*** (0.03)	-0.158*** (0.03)	-0.135***
H-index difference		-0.001 (0.00)	-0.001 (0.00)
Candidate (W) has a Phd from the same institute		0.016 (0.01)	0.013 (0.01)
Candidate(W)'s PhD advisor is in the committee		-0.002 (0.02)	-0.005 (0.02)
% of women candidates			0.133** (0.06)
F-Statistic	15.40	15.21	19.47
# observations	517	517	517

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Dyads represent pairs of female-male candidates ranked by the same jury.
The gender variable is coded as 1 for a woman and 0 for a man.

discounted h-indices and citations. The variable that seems most significant in this case the age-discounted h-index. Including these other citation measures makes no difference to the effect of the reform.

Table B.8: Rank-Ordered Logit using the quota

Dependent variable:	Rank		
Estimator:	Logit	Logit	Logit
	(1)	(2)	(3)
Post-reform	0.016 (0.22)	0.001 (0.22)	-0.014 (0.22)
Candidate is a woman	-0.203 (0.13)	0.076 (0.33)	0.140 (0.47)
H-index of the candidate		0.042*** (0.01)	0.043*** (0.01)
Effect of having more women candidates		-0.529 (0.67)	-0.588 (0.69)
PhD supervisor in the committee			0.142 (0.21)
PhD from the same institute			0.202 (0.16)
PhD from a Parisian university			-0.095 (0.02)
# observations	648	648	648

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. The gender variable is coded as 1 for a woman and 0 for a man. The h-index variable can be interpreted as the effect of increasing the h-index for all candidates. Other variables must be interpreted as the differential effect of the variable for women with respect to men.

Table B.9: Rank-Ordered Logit using the quota

Dependent variable:	Rank		
Estimator:	Logit	Logit	Logit
	(1)	(2)	(3)
Post-Reform*Group	-1.126** (0.44)	-1.393*** (0.47)	-1.662*** (0.51)
Woman in a jury with a female president	0.012 (0.24)	0.084 (0.25)	0.198 (0.26)
Reform	0.600** (0.31)	0.766** (0.30)	0.728** (0.33)
Group	0.599** (0.26)	0.316 (0.26)	0.109 (0.36)
Candidate is a woman	-2.242** (0.98)	-1.819* (0.99)	-4.131*** (1.28)
Mean age of jurors	0.035* (0.02)	0.047** (0.02)	0.087*** (0.03)
H-index of the candidate		0.032** (0.01)	0.128 (0.11)
Effect of having more women candidates		-1.977** (0.86)	-0.799 (1.02)
Reform*H-index		0.042 (0.03)	0.041 (0.03)
PhD supervisor in the committee			0.112 (0.23)
PhD from the same institute			0.172 (0.16)
PhD from a Parisian university			-0.061 (0.13)
Mean h-index of jurors			-0.014 (0.04)
Age*H-index			-0.002 (0.00)
Discipline fixed effects	No	No	Yes
# observations	648	648	648

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. The gender variable is coded as 1 for a woman and 0 for a man. The h-index variable can be interpreted as the effect of increasing the h-index for all candidates. Other variables must be interpreted as the differential effect of the variable for women with respect to men.

Table B.10: IV Probit on gender of first-ranked candidate

Dependent variable:	Gender of first-ranked candidate		
Estimator:	Logit	Logit	Logit
	(1)	(2)	(3)
Gender of winner			
% of women in the committee	0.984 (2.87)	1.361 (2.97)	1.162 (3.25)
Group	-0.357 (0.36)	0.220 (0.33)	0.377 (0.33)
Reform	0.022 (0.30)	-0.063 (0.31)	-0.087 (0.31)
Gender (W) of president		-0.135 (0.24)	-0.155 (0.24)
% of women candidates		2.848*** (0.75)	2.960*** (0.76)
Mean age of the committee members			0.042* (0.02)
Establishment fixed effects			0.350 (0.35)
$\tan(\rho)$	0.020 (0.32)	-0.171 (0.33)	-0.153 (0.34)
$\ln(\sigma)$	-2.215*** (0.06)	-2.266*** (0.06)	-2.303*** (0.06)
# observations	196	196	196

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. The gender variable is coded as 1 for a woman and 0 for a man. This regression considers first-ranked candidates only.

Table B.11: Rank-Ordered Logit using the quota: Other publication measures

Dependent variable:	Rank		
Estimator:	Logit	Logit	Logit
	(1)	(2)	(3)
Post-Reform*Group	-1.097** (0.45)	-1.403*** (0.48)	-1.355*** (0.49)
Woman in a jury with a female president	-0.011 (0.26)	0.010 (0.25)	0.051 (0.26)
Group	0.508* (0.26)	0.292 (0.26)	0.215 (0.28)
Reform	0.586* (0.31)	0.732** (0.31)	0.724** (0.32)
H-index of the candidate	0.004 (0.03)	0.009 (0.04)	0.012 (0.04)
Age-discounted H-index	0.138** (0.06)	0.138** (0.06)	0.132** (0.06)
Coauthor-discounted H-index	0.023 (0.05)	0.022 (0.05)	0.022 (0.05)
Citations	0.000 (0.00)	0.000 (0.00)	0.000 (0.00)
Age-discounted Citations	-0.001* (0.00)	-0.001* (0.00)	-0.002** (0.00)
Coauthor-discounted Citations	-0.000 (0.00)	-0.000 (0.00)	-0.000 (0.00)
Candidate is a woman	-0.398** (0.17)	0.448 (0.46)	0.314 (0.52)
Effect of having more women candidates		-1.718** (0.87)	-1.757** (0.87)
PhD supervisor in the committee			0.085 (0.23)
PhD from the same institute			0.232 (0.17)
PhD from a Parisian university			-0.099 (0.13)
Mean h-index of jurors			0.016 (0.02)
# observations	644	644	644

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. The gender variable is coded as 1 for a woman and 0 for a man. The h-index variable and other citation measures can be interpreted as the effect of increasing the h-index for all candidates. Other variables must be interpreted as the differential effect of the variable for women with respect to men.

4

Homophily and Migration

4.1 INTRODUCTION

Even though workers are mobile, economic outcomes differ wildly between locations in the US.¹ One reason for this may be differences in the ability of workers to exploit spatial arbitrage opportunities, possibly due to high mobility costs. Differentials in mobility costs have often been used to explain the differences in mobility between high and low-skilled workers (see e.g. Topel, 1986 or Bound & Holzer, 2000) or between black and white workers, as in Raphael and Riker (1999).² An alternative explanation³ is that individuals do not only value economic outcomes when con-

¹See, for instance, Lkhagvasuren (2012) 3% of US workers move across states each year.

²According to Raphael and Riker, black workers are 7% less likely to move following a shock.

³See e.g. Notowidigdo (2011) for further evidence that mobility costs may not be very high.

sidering which city they want to live in, but other factors which may be difficult to quantify, and could be race-specific, such as amenities.

In this article, I consider whether inter-city migration, following labor market shocks, is affected by own-race preferences. This is a complicated question to answer since location choice is itself determined by other factors (wages, rents) that are endogenous, since they are determined jointly by supply and demand. In other words, lower rents attract more workers, but more housing demand increases rents. To solve this problem, I rely on a spatial general equilibrium model with 11 equations, that includes a labor demand and labor supply curve, equilibrium on the housing market and transfers. Individuals are heterogeneous in race (black and white) and skill (split by education level). In the location choice equation, individuals choose where to live depending on local wages, rents, amenities, and the local share of co-ethnic residents. The equations are then instrumented by Bartik shocks, including interactions with lagged shares of hispanic and asian residents, to instrument for all the endogenous variables in the model. This allows me to recover my parameters of interest such as the elasticity of substitution between black and white workers, or the inverse elasticity of housing demand. I then simulate labour market shocks to see how own-group preferences affect the location choices of workers, and how this affects the racial wage gap.

I find strong effects of racial preferences on mobility for low-skilled and high-skilled workers: for given wages and housing costs, workers prefer to live in cities that have a higher share of own-group residents, even when taking network externalities into account.¹ These preferences are stronger for low-skilled workers. I do not consider the channels through which these own-group preferences operate. The results that I find could also be consistent with race-specific amenities

¹Network amenities are captured through birth-state effects.

that are not captured by the amenity index, race-specific human capital or networks that are not capture by the state of birth effects.

The effects that I find are important since black workers are disproportionately located in MSAs that have low wages, whatever their skill level, as can be seen in Figure C.1. This figure plots the residual wages of black workers¹ in a given city against the share of black workers in the city. There is a negative and significant correlation between these two variables, for both high- and low-skilled workers, indicating that black workers are not sorting into the cities that are offering them the best wages.²

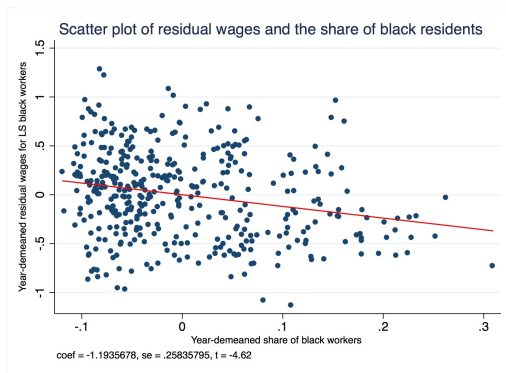


Figure 4.1: Low-skilled workers

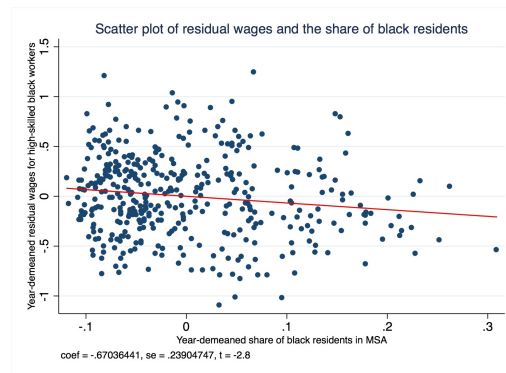


Figure 4.2: High-skilled workers

Past articles have often considered whether location choices are affected by discrimination and own-group preferences, but usually within city. Theoretically, the literature starts with Schelling (1969) and (1971), who shows that even small preferences for living with people of one's own group can lead to total segregation in equilibrium. Empirically, Cutler *et al* (1999) examine the rise and fall of ghettos in the US, and try to disentangle the root causes of within-city segregation. They

¹Residual wages in this case are the city fixed effects from a Mincerian wage regression that includes education, gender, age, age squared, occupation, industry and marital status dummies, on individuals of a given self-reported racial status.

²Similar graphs for white workers are included in the Appendix, but they do not show the same patterns

make a distinction between 3 different hypotheses: the "port of entry" hypothesis, where blacks prefer to live together, especially if they have just moved to a new location; collective action racism, where whites actively segregate through redlining and violence; and decentralised racism, where whites are able to self-segregate by paying a premium on the neighbourhoods where they choose to live. They find that segregation rose in the US until the 1970s, principally caused by collective action racism, before declining thereafter, with the remaining segregation caused by decentralised racism i.e. whites have higher willingnesses to pay for living in neighbourhoods with a majority of their own ethnicity. Card et al. (2008) consider whether within-city location choices of whites exhibit tipping points depending on the population share of minorities in that neighbourhood. Using census-tract level data, they find that whenever blacks make up a significant share of census tract population (depending on the city, between 5 and 10%), a large share of the white population decides to leave. Rathelot & Safi (2014) find that minorities in France have lower mobility rates if they live in neighbourhoods with high minority residency. Bayer et al. (2004) also look at how own-group preferences can shape neighborhood formation, using a multinomial logit model similar to the one proposed in this article. They find evidence of sorting across neighbourhoods, and high marginal willingnesses to pay for areas composed with a high percentage of one's own ethnicity.

Very little work has been done, however, on segregation at the city / Metropolitan Statistical Area (henceforth MSA) level, despite the high level of variation in black and white shares of population across MSAs (e.g. New Orleans and Memphis vs. Portland and Albuquerque). Since the choice of residence within MSAs can be highly constrained, through redlining and other more subtle forms of discrimination, explaining residential choice at the neighbourhood-level through preferences can be challenging. In contrast, we can expect redlining to be less of an issue when

considering moving to another city, since individuals can move to any MSA they wish, even if they are not able to move to a specific neighbourhood.

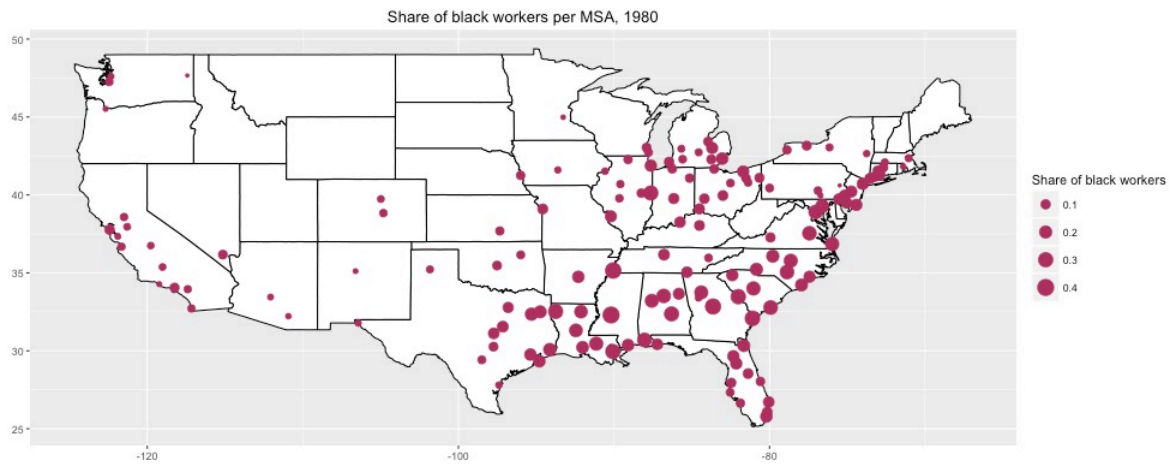


Figure 4.3: Black workers/Total population per MSA in 1980

Even though segregation at the neighbourhood level is higher,¹ segregation across cities is large and persistent. As shown in Tables C.1 and C.2, the percentage of black residents in MSAs varies in the sample between 40 and 1%. There seems to be a persistence in this ordering, since 9 out of the 10 cities with the largest share of black residents in 1980 can also be found in the same top 10 ranking in 2000. I represent this graphically in Figures 4.3 and 4.4, which show the share of black residents for the MSAs in my sample. It is easily seen graphically that the relative order of the share of black residents in cities is extremely stable. Of course, if there are homophilic preferences, then

¹According to Bayer et al. (2004), the average black resident in Detroit lived in a neighbourhood that was 80% black, but the average white resident in a neighbourhood that was 90% white.



Figure 4.4: Black workers/Total population per MSA in 2000

this could be an equilibrium outcome, and the stability of this ranking is not surprising.

Why do some MSAs remain segregated, or accrue a larger than expected share of black residents? An explanation can be seen on Figures C.3 and C.4 below, where I plot the average change in population of an MSA against the change in the share of black residents in the same MSA. These two variables are negatively correlated, which can be explained by own-group preferences. If homophily is important, then instead of leaving when their MSA is hit by negative shocks, some workers prefer to stay because the share of their group within the MSA increases, which mitigates the negative shock to wages. This relationship holds for both high-skilled and low-skilled workers, which makes it unlikely that this is entirely driven by housing or credit constraints. The aim of this article will be to see whether these effects exist and if so to try and quantify

them in a general equilibrium model.

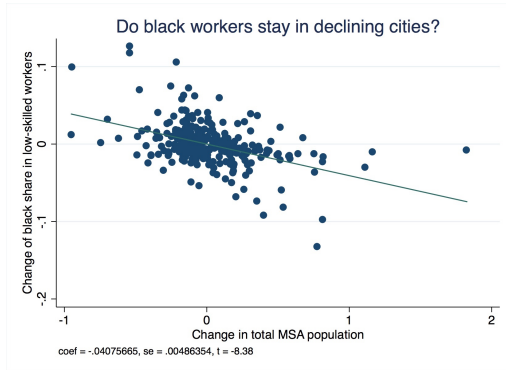


Figure 4.5: Low-skilled workers

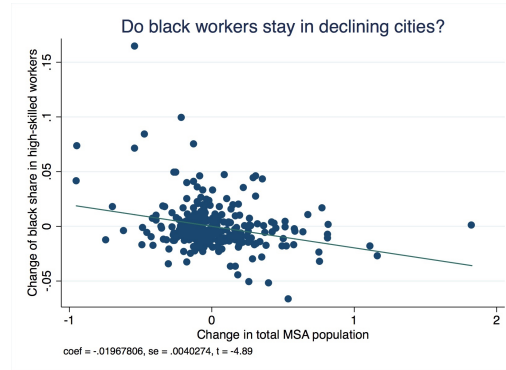


Figure 4.6: High-skilled workers

I first present the general equilibrium model in section 4.2. I then explain the data in section 4.3, before showing the results in section 4.5

4.2 MODEL

I present a spatial general equilibrium model *à la* Rosen (1974)-Roback (1982) with homophily, based on previous work by Notowidigdo (2011) and Diamond (2013).

I consider 4 different groups (high and low-skilled, white and black) which interact in different ways depending on which market we are considering. For instance, some workers of different types may be complements in production, but compete for the same houses. The overall architecture of the model is based on a labour supply and demand curve, endogenous transfer and amenity provision, and housing prices determined by equilibrium on the housing market. When considering an individual's labour supply decision, I allow workers to have homophilic preferences depending on their group and skill type, which will affect where workers of different groups decide to locate.

The model will then be estimated through simultaneous equations GMM in section 4.5.

4.2.1 LABOUR DEMAND

Assume a large number of cities i , which produce a homogenous tradable good using two inputs, high-skill labour H_{it} and low-skill labour L_{it} ¹ with a CES production function, as in Notowidigdo (2011). Estimation of the elasticities from nested CES functions can be challenging. Card & Lemieux (2001) estimate a similar model with skill and age cohorts by using a two-step procedure, while Notowidigdo (2011) reduces the number of parameters to estimate by calibrating the skill-group elasticities. I assume perfect substitution between black and white workers within skill level, an assumption which holds in the data.²

$$Y_{it} = (\theta_{it}^H H_{it}^\rho + \theta_{it}^L L_{it}^\rho)^{\frac{1}{\rho}} \quad (4.1)$$

where ρ is equal to $1 - \frac{1}{\sigma_{HL}}$, with σ_{HL} the elasticity of substitution between different skills, and the θ s are city-level technological efficiency parameters. Since blacks and whites are perfect substitutes, we can rewrite H_{it} and L_{it} as: $H_{it} = (\eta_t H_{it}^W + H_{it}^B)$ and $L_{it} = (\phi_t L_{it}^W + L_{it}^B)$, where ϕ_t and η_t reflect productivity differences.³

Wages are set on the demand curve and so are equal to marginal product. Let P_t be equal to the price of the tradable good. This gives us four different equations for the different groups. To prevent the exposition of the model from being too cumbersome, I only present the results for one group, since the wages for the other groups can be derived analogously. We can write the

¹For simplicity, there is no capital in the model. The empirical section uses decadal changes, and since capital is more elastic than labour, it is likely that most of the incidence from negative supply shocks falls on workers.

²Boustan (2009) estimates the elasticity of substitution between black Southern migrants and Northern whites, and finds an extremely high elasticity of substitution of 23.8, in the period from 1940 to 1970. I try to estimate a similar elasticity of substitution for our time period in section C.3 of the appendix, and cannot reject the null hypothesis that blacks and whites are perfect substitutes in production within skill-categories.

³These differentials may not reflect actual productivity differences, and could be caused by discrimination.

wages of a specific group (here high-skilled whites) as:

$$W_{it}^{HW} = \frac{\partial P_t Y_{it}}{\partial H_{it}^W} = P_t \theta_{it}^H H_{it}^{\rho-1} Y_{it}^{1-\rho} \eta_t \quad (4.2)$$

We can then rewrite the log labour demand curves as:

$$w_{it}^{HW} = \ln\left(\frac{W_{it}^{HW}}{P_t}\right) = \ln(\theta_{it}^H) + (\rho - 1)\ln(H_{it}) + (1 - \rho)\ln(Y_{it}) + \ln(\eta_t) \quad (4.3)$$

One potential problem in estimation is that the city-level productivity parameters are likely to be correlated with the labor supply parameters, because of spillovers. One way to improve precision is to take first-differences to get rid of time-invariant productivity differences, but the time-varying components could depend on both the high and low-skilled employment ratios, and on the number of workers in the city, i.e. $\theta_{it} = F(H_{it}, L_{it})$. In some of the specifications, I follow Diamond (2013) and approximate the labour demand curves above with a first-difference, log-linear aggregate labor demand, which has the advantage of both taking spillovers into account and of being more flexible:

$$\Delta w_{it}^{HW} = \gamma_{HHW} \Delta \ln(H_{it}) + \gamma_{HLW} \Delta \ln(L_{it}) + \Delta \ln(\eta_t) + \Delta \epsilon_{it}^{\theta HW} \quad (4.4)$$

$$\Delta w_{it}^{HB} = \gamma_{HHB} \Delta \ln(H_{it}) + \gamma_{HLB} \Delta \ln(L_{it}) + \Delta \epsilon_{it}^{\theta HB}$$

$$\Delta w_{it}^{LW} = \gamma_{LHW} \Delta \ln(H_{it}) + \gamma_{LLW} \Delta \ln(L_{it}) + \Delta \ln(\phi_t) + \Delta \epsilon_{it}^{\theta LW}$$

$$\Delta w_{it}^{LB} = \gamma_{LHB} \Delta \ln(H_{it}) + \gamma_{LLB} \Delta \ln(L_{it}) + \Delta \epsilon_{it}^{\theta LB}$$

where $\Delta \epsilon_{it}^{\theta}$ represent changes in technology that are not captured by changes in population.

4.2.2 TRANSFER PAYMENTS

An important factor which can explain differences in mobility between skill-groups are transfer payments, which can help equalise situations across different cities, and are principally paid out to low-skilled workers.

However, the provision of benefits may also depend on the race of the potential recipients. The opinion of Americans on welfare and crime spending depends heavily on their attitudes towards race.¹ I allow the welfare spending in cities to also depend on the cities' racial composition, to test this hypothesis. The elasticity of welfare payments to economic downturns is allowed to vary depending on which group the incidence of the shock falls. The transfers to unskilled individuals are assumed to have a constant elasticity with respect to changes in their wage. Hence:

$$T_{it} = W_{it}^{WL\psi^W} W_{it}^{BL\psi^B} \exp(\epsilon_{it}^T)$$

Taking log differences, we get the following equation:

$$\Delta \ln(T_{it}) = \psi^W \Delta w_W^L + \psi^B \Delta w_B^L + \Delta \epsilon_{it}^T \quad (4.5)$$

Transfer payments are countercyclical, in order to mitigate the effect of economic downturns. We should therefore expect transfer payments to negatively covary with changes in wages of low-skilled workers, i.e. $\psi < 0$.

¹Alesina et al. (1999) show that public good provision depends on a city's ethnic fractionalisation. See also Gilens (1996) or Wilson & Nielsen (2011).

4.2.3 HOUSING SUPPLY

Next I consider the housing market. Following Davis & Heathcote (2007) and Davis & Palumbo (2008), I consider the main inputs of housing production to be construction materials and land. As in Saiz (2010), land developers are price takers, and housing demand HD is a function of population so that:

$$P_{it}^h = f(CC_{it}, P_L(HD))$$

where P_{it}^h is the price of housing, f is an unknown function, CC_{it} are construction costs, and P_L is the price of land. Land is assumed to be fixed, so that as population increases, there is less land available for development, and land prices increase. There is no uncertainty, so that prices are simply the discounted values of rents: $P_{it}^h = \frac{R_{it}}{i}$, where i is the interest rate.

The elasticity of land prices (and in turn house prices) to changes in housing demand is not homogenous across cities, and depends on the land available for development, and regulations. I allow for heterogeneity across cities in housing demand elasticities by using the land availability index from Saiz (2010), and the Wharton regulation index from Gyourko et al. (2008). Both of these indices try to capture how elastic land prices are to changes in demand. The land availability index tries to capture how geographical constraints affect the land available for development. The WRI does the same but for regulation and planning constraints. We can approximate the unknown function $f()$ with a log-linear housing supply. Taking log-differences, we can rewrite the equilibrium on the housing market as:

$$\Delta \ln(R_{it}) = \mu_i \Delta \ln(HD_{it}) + \Delta \epsilon_{it}^{CC} \quad (4.6)$$

Where μ , as in Saiz, is $\mu_0 + \mu^{land} x_i^{land} + \mu^{reg} x_i^{reg}$, to take into account heterogeneity in the elas-

ticity of housing supply, and HD is housing demand, which is the sum of housing expenditure by each group: $HD = \sum_z \zeta^z Z_{it}^z W_{it}^z$, where Z_{it} is the number of individuals of type z and ζ^z the share on income spent on housing by each group. The error term, ϵ_{it}^{CC} represents changes in unobserved construction costs. To identify μ , we need to find variation that is independent from changes in construction costs, so housing demand will have to be instrumented.

4.2.4 LABOUR SUPPLY

One potential problem is that the price of housing could be correlated with unobserved changes in the error terms. For instance rents could be higher because of changes in amenities, which biases our estimates. For this reason, I adopt a version of the BLP (Berry et al., 1995) framework that accounts for correlation between amenities and house prices. This approach was first used for neighbourhood choice in Bayer et al. (2004) and Bayer et al. (2007), and then to location choice between MSAs in Diamond (2013) and Piyapromdee (2014).

Individual workers j maximise the following Cobb-Douglas utility function:

$$U_{ijt}^z = \ln(G_{jt}^{1-\zeta^z}) + \ln(Q_{jt}^{\zeta^z}) + u_j(D_{it})$$

$$\text{subject to } G_{jt} * P_t + Q_{jt} * R_{it} \leq W_{it}^z$$

Individuals spend a fixed share ζ of income W on the housing good Q_{jt} , and their remaining income on a nationally traded good G_{jt} with price P_t . Only variables that are city-specific matter for city choice, so the wage, rents, and the variable D , which consists of mobility costs, amenities and potentially homophilic preferences.

Estimates by Davis & Ortalo-Magne (2011) show that housing seems to be a normal good with

an elasticity close to 1, and so a Cobb-Douglas utility function fits the data along a large number of dimensions. I nevertheless allow the share ζ to vary across skill-groups¹.

The indirect utility that a worker j gains from living in a city i , v_{jit} , is:

$$v_{jit} = w_{it}^z - \zeta^z r_{it} + u_j(D_{it})$$

where $w_{it}^z = \ln(W_{it}^z/P_t)$ and $r_{it} = \ln(R_{it}/P_t)$. For low-skilled workers, $W_{it}^z = s_W W_{it}^z + s_L T_{it}$, where s_W is the share of income from wages, and s_L is the share of income from transfers, assumed to be 0 for high-skilled workers. I now define $u_j(D_{it})$ as follows:

$$u_j(D_{it}) = a_{it}\beta_t^{z,a} + \beta_t^{z,st} x_j^{z,st} + \beta_t^{z,div} x_j^{z,div} + \omega^z \left(\frac{H_{it}^z + L_{it}^z}{H_{it} + L_{it}} \right) + \epsilon_{it}^{z,City} + \epsilon_{ijt}^{z,A}$$

I consider that individuals gain utility from living in the same state ($\beta^{st} x^{st}$) or division ($\beta^{div} x^{div}$) that they were born in, from the city's amenities ($a_{it}\beta_t^{z,a}$), and from a homophily parameter that depends on the racial composition of the city ($\omega^z (\frac{H_{it}^z + L_{it}^z}{H_{it} + L_{it}})$). Using birth states and divisions is a useful way to integrate mobility costs into our framework: the more attractive cities will be those that attract the most workers born in other states and census divisions. Notice that some of these parameters are common for all workers of type z living in a particular city i . We can combine those parameters into a single parameter δ as follows:

$$\delta_{it}^z = \lambda^z (w_{it}^z - \zeta^z r_{it}) + \omega^z \left(\frac{H_{it}^z + L_{it}^z}{H_{it} + L_{it}} \right) + \beta^{z,a} a_{it} + \epsilon_{it}^{z,City} \quad (4.7)$$

¹see e.g. Guerrieri et al. (2013) and Michaels et al. (2012)

We can then rewrite the indirect utility function as simply:¹

$$v_{jit} = \delta_{it}^z + \beta_t^{z,st} * x_j^{z,st} + \beta_t^{z,div} x_j^{z,div} + \epsilon_{ijt}^A$$

What is the probability that an individual chooses city i? We need to find the probability that $v_{jit} \geq v_{jkt} \forall k$. If the ϵ_{ijt}^A are distributed type I extreme value, then we get the familiar form below for this probability (McFadden (1973)).

$$Pr(v_{jit} \geq v_{jkt} \forall k) = \frac{\exp(\delta_{it}^z + \beta_t^{z,st} x_j^{z,st} + \beta_t^{z,div} x_j^{z,div})}{1 + \sum_{i=1}^n \exp(\delta_{it}^z + \beta_t^{z,st} x_j^{z,st} + \beta_t^{z,div} x_j^{z,div})} \quad (4.8)$$

In the first step of the estimation, I treat the δ^z s as parameters to estimate, along with the other parameters, through maximum likelihood. I then take first differences to get rid of fixed effects in amenity provision, and estimate the equation below, using δ^z as a dependent variable.

$$\Delta \delta_{it}^z = \lambda^z (\Delta w_{it}^z - \zeta^z \Delta r_{it}^z) + \omega^z \Delta \left(\frac{H_{it}^z + L_{it}^z}{H_{it} + L_{it}} \right) + \beta^{z,a} \Delta a_{it} + \Delta \epsilon_{it}^{z,City} \quad (4.9)$$

Here, our main parameter of interest is ω ,² which tells us how mean utility changes in response to changes in the composition of one's own group in the population.

4.2.5 ENDOGENOUS AMENITY SUPPLY

Diamond shows that amenities react endogenously to changes in the skill ratio of the city, which could be a problem when estimating the labour supply equations. Another problem, specific to

¹For identification of the parameters, I normalise the equation above by dividing by σ_j^z , the parameter that governs how much dispersion there is in taste for cities, hence the appearance of λ^z in the equation.

²There is no heterogeneity in tastes above and beyond the idiosyncratic error term, for ease of computation.

our set-up, is that the changes in amenities may be correlated with changes in ethnic composition.¹ I therefore add endogenous amenity supply to the model as follows, which will allow us to check whether amenities vary at the MSA level with the racial composition of the cities.²

$$\Delta a_{it} = \beta^{HL} \Delta \ln\left(\frac{H_{it}}{L_{it} + H_{it}}\right) + \beta^{BW} \Delta \ln\left(\frac{B_{it}}{W_{it} + B_{it}}\right) + \Delta \epsilon_{it}^{EA} \quad (4.10)$$

where $B_{it} = H_{it}^B + L_{it}^B$ and $W_{it} = H_{it}^W + L_{it}^W$

4.2.6 EQUILIBRIUM

The equilibrium on this market following an exogenous shock to local labour demand is as follows:³

1. Labour supply is equal to labour demand for all four groups.

$$\Delta w_{it}^z = \gamma_{HWZ} \Delta \ln H_{it} + \gamma_{HLZ} \Delta \ln L_{it} + \Delta \eta_t^z + \Delta \epsilon_{it}^{wZ}$$

$$Z_{it} = \sum_{j \in z} \frac{\exp(\delta_{it}^z + \beta_t^{z,st} x_j^{z,st} + \beta_t^{z,div} x_j^{z,div})}{1 + \sum_{i=1}^n \exp(\delta_{it}^z + \beta_t^{z,st} x_j^{z,st} + \beta_t^{z,div} x_j^{z,div})}$$

2. Amenity demand is equal to endogenous amenity supply

$$\Delta \delta_{it}^Z = \lambda_{it}^z (s_W \Delta w_{it}^z + s_T * \Delta t_{it} - \zeta^z \Delta r_{it}) + \omega^z \Delta \left(\frac{H_{it}^z + L_{it}^z}{H_{it} + L_{it}} \right) + \beta_t^{z,a} \Delta a_{it} + \Delta \epsilon_{it}^{z,City}$$

¹Bayer et al. (2004) use boundary changes as an instrument when estimating a similar equation at the neighbourhood level.

²A stronger, alternative assumption, would be to check whether ethnic *diversity* affects amenity provision at this level. There is a large literature on the negative relationship between ethnic diversity and public good provision, (see e.g. Habyarimana et al. (2007) or Algan et al. (2016) for examples). However testing this hypothesis would require a fractionalisation index with many different groups - something that is beyond the scope of this paper for now

³See Bayer & Timmins (2005) for proof of uniqueness in this class of models. The potential problem comes from the spillover effects, that can cause multiple equilibria. As an example, two cities with identical characteristics can self-sort into all-black or all-white cities depending on how agents coordinate. Uniqueness relies on preferences for self-segregation being weak enough relative to preferences for characteristics of the city.

$$\Delta a_{it} = \beta^{HL} \Delta \ln\left(\frac{H_{it}}{L_{it}}\right) + \beta^{BW} \Delta \ln\left(\frac{B_{it}}{W_{it}}\right) + \Delta \epsilon_{it}^{EA}$$

3. Housing prices adjust to changes in population, so that the change in housing supply is equal to changes in housing demand:

$$\Delta \ln(R_{it}) = \mu_i \Delta \ln(HD_{it}) + \Delta \epsilon_{it}^{CC}$$

4. Transfers respond to shocks on low-skill wages:

$$\Delta \ln(T_{it}) = \psi^W \Delta w_W^L + \psi^B \Delta w_B^L + \Delta \epsilon_{it}^T$$

This gives us 11 equations to estimate. I present both reduced form evidence, and structural estimates, using data from the US census.

4.3 DATA

For all the aggregate-level variables I use data from the 5% sample of the 1980, 1990 and 2000 individual-level American census (IPUMS).¹ All differences in the paper are decadal changes. My main sample used to compute wages, population and rents, includes all native-born households between the ages of 25 and 55, currently not in group quarters, working at least 35 hours a week and 48 weeks a year and living in a metropolitan area. I use the sample weights to obtain the appropriate aggregate variables. I also exclude workers outside of the labour force, and those earning business or farm income. Wages are computed as the log of yearly wage income divided by weeks worked. I use data on monthly gross rents from the census to compute the housing costs of renters, and I use the discount rate from Peiser and Smith (1985) to impute rents to home-owners. Wages, transfer and rents are all deflated using the CPI-U.

¹See Flood et al. (2015).

The skill groups are composed of high-skill workers, who have some amount of college education, and low-skill workers who have a high-school degree or less.

In computing the shares of each racial group, I use the self-reported racial groups in the sample. These shares are over the population of working class adults (immigrants and natives) in the city, so that the shares of whites and blacks do not add up to one. For ease of computation, I use a smaller sample when estimating the conditional logit consisting only of household heads. Other members of the household are assumed to move with the household head. Individuals living in MSAs that are not constantly identified as such in 1980, 1990 and 2000, or individuals that do not live inside a MSA, are considered as living in a state-specific "rural" area for the purposes of the conditional logit estimation.

4.3.1 AMENITIES

The changes in amenities are calculated by changes in the amenities used in Diamond (2013), which encompasses data on school quality, crime, environment, transport and retail amenities. These data are amalgamated into a single index using principal component analysis. The only difference between this index and the one in Diamond is that the periods considered are 1980-1990 and 1990-2000 rather than 1980-1990 and 1980-2000. However, PCA does not seem to work as well when using decadal changes. The factor loadings for the transportation and retail index are not the same as in Diamond. Having more movie theaters for instance has a negative factor loading in the retail index. More worryingly, the factor loadings for the single amenity index are not consistent. Retail and environment, for instance, have a negative factor loading, along with crime. This might explain some of the strange results we get in the full regression, where amenities do not seem to have any effect on city choice.

Table 4.1: Amenity index

Variable	Factor Loading	Unexplained variance
<i>School index</i>		
Government K–12 spending per student	0.7071	0.4459
Student–teacher ratio	-0.7071	0.4459
<i>Retail</i>		
Apparel stores per 1,000 residents	0.6891	0.1702
Eating and drinking places per 1,000 residents	0.6722	0.2104
Movie theaters per 1,000 residents	-0.2707	0.8719
<i>Transportation</i>		
Public buses per capita	0.6487	0.2402
Public transit index	0.6109	0.326
Average daily traffic—interstates	-0.3276	0.8062
Average daily traffic—major roads	-0.3141	0.8218
<i>Crime</i>		
Property crimes per 1,000 residents	0.7071	0.482
Violent crimes per 1,000 residents	0.7071	0.482
<i>Environment index</i>		
Government spending on parks per capita	0.7071	0.4851
EPA air quality index	-0.7071	0.4851
<i>Jobs</i>		
Patents per capita	0.7071	0.4769
Employment rate	0.7071	0.4769
<i>Overall amenity index</i>		
Retail index	-0.6242	0.3907
Transportation index	0.6109	0.4164
Crime index	-0.1029	0.9835
Environment index	-0.0496	0.9962
School index	0.1618	0.959
Job index	0.4449	0.6905

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Standard errors are clustered at the MSA level. Dependent variables are percentage changes.

4.4 IDENTIFICATION

In order to identify the parameters in the model, I use an instrument *a la* Bartik (1991), which has been used extensively in the literature. Bartik's shift-share instrument interacts local employment shares with employment shocks at the national level, which should be exogenous to local employment conditions. These instruments can be considered as proxies for technological shocks. For instance, Blanchard & Katz (1992) use this instrument to estimate the response of migration to local labour demand shocks, Piyapromdee (2014) to estimate the impact of immigrants on local wages and rents, Diamond (2013) to find the impact of college workers on changes in amenities and Serrato & Wingender (2011) to estimate the impact of local government spending. Similar to Diamond's approach, I use a Bartik shock to wages in order to instrument all the endogenous variables in the model. The Bartik shock at the level of the local labor market can be computed as follows:

$$\hat{B}_{it}^z = \sum_{r=0}^N (w_{rt}^z - w_{rt-1}^z) \frac{Z_{irt-1}}{Z_{it-1}}$$

The growth in log wage at the industry level for workers of group z , excluding city i , is interacted with the 10 year lagged share of workers working in the industry at the city level, and then summed over all industries.¹ I will use Bartik shocks to instrument wages, population and house prices. I also follow Diamond (2013) in using the interaction between the Bartik shocks and the variables that determine the elasticity of land to population. When estimating the labour demand equations, for example, if housing supply is particularly inelastic then the influx of population responding to a technological change will be small since rents will rise disproportionately.

¹In the literature on Bartik instruments, a reference base line e.g. 1980 is sometimes used instead of simple differences, in order to account for serial correlation of shocks. This is the case in Diamond (2013) for instance. However, this approach then conflates short-term and long-term elasticities, which is why I prefer to eschew this approach and use decadal differences instead.

I only consider Bartik instruments at the skill level, and do not use Bartik instruments at the skill-race level. However, I do use an additional instrument to take care of the endogeneity of the share of one's own group in the city, even though through differences in skill composition and propensity to migrate, Bartik shocks turn out to be good instruments for the share of black or white workers in a city. I interact the 10 year lag share of residents who are Asian or Hispanic immigrants with the two Bartik shocks. This instrument makes sense since we know that immigrants are likely to live in enclaves (Card, 2009), that they are more responsive to labour shocks (Cadena & Kovak, 2016), and that Asian and Hispanics immigrants were the fastest growing immigrant group over the considered time period.

Each equation from the model above is estimated with 8 instruments based on Bartik-shift shares, and interactions with housing supply parameters.¹

4.5 ESTIMATION

4.5.1 REDUCED FORM

First I investigate whether the instrument that uses the 10-year lagged share of immigrants is a good instrument for the change in the share of one's group. I regress the current share of white and black residents on the Bartik shocks for high-skilled and low-skilled workers, as well as the Bartik-migrant shock, in Table 4.2. The share of white and black residents is impacted negatively and significantly by these instruments, as expected, though the R^2 for changes in the share of black workers is low. In Appendix C.4, I present at the reduced form effect of the instruments on the other variables used in the estimation.

¹I use Bartik instruments for low-skilled and high-skilled workers, interacted with the regulation and land-use indices, and the additional instrument for the shares of black and white workers.

Table 4.2: Are the Bartik shocks good instruments?

Dependent variable:	Δ Share of black workers	Δ Share of white workers
Estimator	OLS	OLS
Bartik shock for HS workers \times l1 o. share of immigrants	-0.318*** (0.12)	-0.491*** (0.06)
Bartik shock for LS workers \times l1 o. share of immigrants	0.068 (0.14)	0.425*** (0.07)
Bartik shock for HS workers	-3.757*** (1.89)	-5.189*** (0.71)
Bartik shock for LS workers	0.512 (2.34)	4.665*** (0.99)
Year fixed effect	Yes	Yes
Observations	280	280
R^2	0.03	0.39

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Standard errors are clustered at the MSA level. Dependent variables are percentage changes.

4.5.2 GMM

The full model is evaluated through simultaneous equation GMM. Since the model is overidentified (there are more moments than there are parameters), simultaneous equation GMM is more efficient than equation-by-equation GMM. I take η and ϕ , the wage premia for whites in a given skill category, from the data. In some estimations, I also calibrate ζ , the share of income spent on housing by each group, by finding the average amount of income spent by households on housing (including imputed rents for homeowners) in the census data. I find that $\zeta^{HW} = 0.37$, $\zeta^{LW} = 0.40$, $\zeta^{HB} = 0.35$, $\zeta^{LB} = 0.45$.

All the parameters from the model are estimated jointly in a two-step GMM, with standard errors clustered at the MSA level. In all the following tables, column 1 presents the model where ζ is estimated and I do not control for changes in amenities or homophilic preferences, column 2 adds the effect of changes in one's own group. In both these columns labour demand is estimated using the specification with the CES production function. I use the calibrated ζ s specified above

in column 3 and in column 4 the ζ s are calibrated and I control for changes in amenities. In these last two columns, I use the log-linear labour demand curve. Column 4 is added for the sake of comparison with Diamond, but in this estimation the amenity index has a significant negative impact on the desirability of cities. I consider model 3 the main specification.

Let us first look at the labour demand parameters, presented in Table 4.3. For this table, we are interested in the effect of labour supply shocks on wages. In the first two columns, I estimate ρ , where the elasticity of substitution between low and high-skilled workers is $\frac{1}{1-\rho}$. My estimates for the elasticity are high when compared to those of previous articles which estimate the elasticity at the national level. For instance, Katz & Murphy (1992) find an elasticity of 1.4. However, when estimated using variation at the local level, elasticities of substitution tend to be higher. Card (2009), finds values that range between 1.5 and 2.5, which is close to my estimates. The elasticity estimated in column 2 for instance is of 1.81.

In the last two columns, we must be aware that there are two effects at play. Changes in the supplies of high-skilled and low-skilled workers affect wages because they can be more or less easily substituted, but they also affect the overall productivity of the city through spillovers as the city gets larger or through changes in the high-skilled/low-skilled ratio. Hence, it is not obvious *a priori* which signs these parameters should have.

None of the parameters are statistically significant in my estimation. Looking at the point estimates, the spillover effect seems to dominate the substitution effect, at least for white workers. Having more high skilled workers raises the wages of both high-skilled and low-skilled workers, while having more low-skilled workers decreases them. This is consistent with the skill-composition of the city having an effect on the production technology at the city level, as in some of the immigration literature. For instance, Peri et al. (2015) find negative wage and employment

Table 4.3: Labour demand parameters

Model	(1)	(2)	(3)	(4)
ρ	0.601*** (0.08)	0.449*** (0.05)		
Change in supply of HS on HS black wages			0.550 (0.37)	0.370 (0.27)
Change in supply of LS on HS black wages			-0.478 (0.49)	-0.224 (0.38)
Change in supply of HS on LS black wages			0.492 (0.42)	0.424 (0.42)
Change in supply of LS on LS black wages			-0.792 (0.69)	-0.715 (0.69)
Change in supply of HS on HS white wages			0.304 (0.19)	0.088 (0.26)
Change in supply of LS on HS white wages			0.053 (0.31)	0.429 (0.44)
Change in supply of HS on LS white wages			1.300 (0.81)	1.260 (0.85)
Change in supply of LS on LS white wages			-1.760 (1.22)	-1.704 (1.25)
Observations	280	280	280	276

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Standard errors are clustered at the MSA level. All estimates are from decadal changes on the logged variables. Each equation includes a technology (Bartik) shock, a time fixed effect, and a wage premia depending on the group. Instruments include Bartik shocks for high and low-skilled workers, and the interaction of these shocks with housing elasticity parameters. The last column is missing four observation because of missing data for amenities.

Table 4.4: Transfers and Housing Supply

Model	(1)	(2)	(3)	(4)
Elasticity of transfers to low-skilled white wages	-0.771 (1.33)	-0.503 (1.21)	-1.247 (1.40)	-0.880 (1.68)
Elasticity of transfers to low-skilled black wages	-0.081 (1.34)	0.906 (1.25)	1.145 (1.62)	1.094 (1.38)
μ_0	2.278*** (0.37)	0.285*** (0.10)	0.287 (0.20)	0.250** (0.12)
$\mu^{landavailability}$	0.458** (0.21)	1.199*** (0.23)	1.226*** (0.26)	1.112*** (0.20)
$\mu^{regulation}$	-0.060*** (0.01)	-0.017** (0.01)	-0.012 (0.01)	-0.013** (0.01)
Observations	280	280	280	276

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Standard errors are clustered at the MSA level. μ_0 gives the base elasticity of housing prices with respect to housing supply. Each equation includes a time fixed effect. Instruments include Bartik shocks for high and low-skilled workers, and the interaction of these shocks with housing elasticity parameters.

effects on high-skilled natives from random unexpected refusals in HiB work permits.

Table 4.4 presents the model parameters pertaining to transfers and housing supply. The transfer payment elasticity for low-skilled white workers that I find is much lower in magnitude to Notowidigdo (2011). The transfer elasticity with respect to black wages is positive though insignificant, indicating that black workers receive lower transfers than whites in response to shocks, which is a finding similar to Alesina et al. (1999)'s.

Saiz (2010) estimates a very similar equation to the one whose results are presented in Table 4.4. In order to get the full inverse housing supply elasticities, the estimated γ s have to be multiplied with the land availability and regulation indices. Doing so leads us to find an average inverse elasticity of 0.55 when using the estimates from model 3, smaller than Saiz's elasticity of 0.6. Finding a lower elasticity makes sense since Saiz looks at changes over 30 years, while this estimation only considers 10 year changes.

Table 4.5 presents the elasticities of labour supply to wages and rents in columns 1 and 2, and estimated disposable income in columns 3 and 4, with ζ calibrated. Columns 2, 3 and 4

Table 4.5: Sensitivity to wages and rents

Model	(1)	(2)	(3)	(4)
Sensitivity of high-skilled blacks to wage changes	-1.419*** (0.43)	2.603*** (0.62)	2.756*** (1.32)	1.649* (0.87)
Sensitivity of high-skilled blacks to rent changes	0.736** (0.34)	-0.165 (0.36)		
Sensitivity of low-skilled blacks to wage changes	-2.742*** (0.87)	3.218*** (0.51)	2.243** (0.94)	3.448*** (0.70)
Sensitivity of low-skilled blacks to rent changes	-0.524 (0.64)	-0.763* (0.42)		
Sensitivity of high-skilled whites to wage changes	-0.278 (0.21)	2.887*** (0.25)	2.759*** (0.46)	3.316*** (0.75)
Sensitivity of high-skilled whites to rent changes	0.786*** (0.24)	-0.258** (0.11)		
Sensitivity of low-skilled whites to wage change	-1.542*** (0.49)	5.327*** (0.75)	3.960*** (1.51)	5.453*** (0.90)
Sensitivity of low-skilled whites to rent changes	0.736** (0.34)	-0.165 (0.36)		
Observations	280	280	280	276

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Standard errors are clustered at the MSA level. Estimates are demand elasticities with respect to the city characteristic considered, for a small city. Each equation includes a time fixed effect. Instruments include Bartik shocks for high and low-skilled workers, and the interaction of these shocks with housing elasticity parameters.

include the effect of own-race preferences. Column 4 includes the amenity index. Without using own-race preferences, the model gives negative estimates for higher wages, and positive estimates for rents. This implies that given the location choices of workers, the only way of rationalising them is to suppose that workers prefer higher rents and lower wages. Hence, without own-race preferences, the model is mis-specified. Adding own-race preferences and calibrating ζ gives us positive estimates for having higher real income. The elasticity to wages depends on race, with black workers more responsive than their white counterparts. We can also notice that low-skilled workers are more sensitive to wages in this estimation, which is different from what is commonly assumed in the literature, although bear in mind that preferences for living in one's own birth state have been controlled for.

Finally, the parameters of interest ω , are presented in Table 4.6. These parameters are statistically significant and positive for black and white workers across all estimations. These pref-

Table 4.6: Homophily and amenities

Model	(1)	(2)	(3)	(4)
ω^{HB}		3.736*** (0.70)	4.534*** (1.15)	2.620*** (0.94)
ω^{LB}		4.516*** (0.75)	4.788*** (0.95)	3.465*** (1.04)
ω^{HW}		2.815*** (0.29)	2.537*** (0.34)	2.609*** (0.75)
ω^{LW}		5.525*** (0.65)	4.261*** (0.92)	5.203*** (0.87)
Change in amenity index for HS blacks				-0.132 (0.17)
Change in amenity index for LS blacks				-0.124 (0.29)
Change in amenity index for HS whites				-0.403** (0.17)
Change in amenity index for LS whites				0.337* (0.20)
Change in college ratio on amenity index				5.822* (3.26)
Change in black/white ratio on amenity index				-3.589 (1.34)
ω included	No	Yes	Yes	Yes
Calibrated housing share	No	No	Yes	Yes
CES labour demand	Yes	Yes	No	No
Amenity index included	No	No	No	Yes
Hansen's J: p-value	0.00	0.03	0.05	0.07
Observations	280	280	280	276

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Standard errors are clustered at the MSA level. Each equation includes a time fixed effect.

Table 4.7: Network amenities

Year	1980		1990		2000	
	Whites	Blacks	Whites	Blacks	Whites	Blacks
<i>Birth state</i>						
LS workers	3.519	3.313	3.589	3.630	3.540	3.69
HS workers	2.768	2.902	2.864	3.109	2.884	3.117
<i>Birth division</i>						
LS workers	1.225	1.041	1.305	0.921	1.268	0.825
HS workers	1.187	0.776	1.216	0.730	1.162	0.778

Parameters estimated through maximum likelihood from a conditional logit model of MSA choice.

erences are sizeable, and comparable in magnitude to those of higher real wages. The effect of higher amenities however is only positive for low-skilled white workers. One surprising result is that these preference are significant and negative for high-skilled white workers, who are typically thought to have the highest preferences for amenities.

The network amenities, preferences for living in your state of birth or Census division of birth, are presented in Table 4.7. From these results, it seems that black workers are slightly more attached to their state of birth than white workers, but slightly less to their region of birth. Counterfactual simulations will help us understand how these parameters contribute to explaining the differences between blacks and whites in reaction to negative shocks, as we saw in Figures C.3 and C.4.

4.5.3 DISCUSSION

As seen above, the coefficient on homophily is positive and significant for low-skilled workers. The racial composition of the city seems to matter more for low-skilled than high-skilled workers, at least for white workers. There are two potential explanations: One is that this reflects differences in preferences between low-skilled and high-skilled workers. Although possible, this seems at odd with some of the literature. For instance, Bayer et al. (2014) find that rising education standards for black males can increase segregation within a city.¹ The other is that I only consider city choice in this article, whereas in reality, individuals choose both a city and a neighbourhood to live in. It's possible that high-skilled (and richer) workers of both groups can choose neighbourhoods that conform to the sociodemographic characteristics that they want, unlike low-skilled

¹Bayer et al. (2014) find that as racial inequality narrows, segregation can increase. When there is a critical mass of highly-educated black workers in the population, new neighbourhoods are created than can leave them with the choice of living in middle-class black or white neighbourhoods, whereas they may previously only have had a choice between middle-class white and lower-class black neighbourhoods.

workers, and that this explains the results we find.

4.6 WHAT IS THE EFFECT OF HOMOPHILY?

We can use the results from the section above to gauge what a world without homophily would look like, and in particular what the effects would be on inequality. Using the results from the model, I can generate counterfactual cities, and see for instance what the effect on the wage gap would be if individuals sorted into cities only because of real wages and idiosyncratic preferences for cities.

4.6.1 RACIAL WAGE GAP

In my first counterfactual, I try to predict which cities individuals would live in if there were no homophily. I then see what part of the regional wage gap, the part of the wage gap that is due to individuals sorting into cities for reasons other than wages, can be explained by homophily. I do not expect this gap to completely disappear, since network amenities will ensure some persistence of the current sorting patterns. Given the patterns that we saw in Figure C.1, which showed that black workers were disproportionately located in cities with low wages, we should expect better sorting to reduce the racial wage gap.

To estimate the effect of these counterfactual cities, I create a new δ for each group, using equation 4.11 and the parameter values from the third column, and set ω^z to 0:

$$\hat{\delta}_{it}^z = \hat{\lambda}^z(w_{it}^z - \zeta^z r_{it}^z) + \hat{\epsilon}_{it} \quad (4.11)$$

where $\hat{\epsilon}_{it}$ are the residuals from the same equation including the homophily parameter, and

are intended to capture city fixed effects.¹ I then compute the probability of individuals in my sample choosing each city using the new δ , by using the conditional logit set out in equation 4.8. Within each skill-race group, the only additional heterogeneity comes from the birth state of individuals. Creating the counterfactual cities entails finding the probability of living in a given MSA for natives of one of the 51 states or federal districts, and multiplying these probabilities by the number of people in our sample born in these states, for each race-skill group. In Figure 4.7, I plot the difference in population in the counterfactual cities versus the actual cities against the share of black and white workers in these cities. Consistent with the evidence from Figure C.1, we can see that there is resorting for both groups of workers, though the R^2 is much higher for black workers.²

From these figures, I can estimate the wage gap due only to location choices by multiplying the average weekly log wage of workers in each city for each group by the proportion of workers. I consider both nominal wages and real wages net of housing costs, and present the results in Tables 4.8 and 4.9. Allowing for better sorting does not necessarily improve outcomes in Table 4.8, but this issue is resolved once we take rents into account in Table 4.9. We can see that better sorting across space would reduce the wage gap for most years (for instance by 0.08 log points in 1980 for low-skilled workers), but that a large share of the wage gap, whether real or nominal remains unexplained. This is because there is more inequality between white and black workers within city, than there are between workers of the same group across space.

¹This is an upper bound on how much homophily matters since I ignore the further effects on wages and rents of the reallocation of workers across cities.

²0.97 for black workers and 0.67 for white workers.

Table 4.8: Racial wage gap

Year	1980		1990		2000	
	HS	LS	HS	LS	HS	LS
<i>Real log wage</i>						
White workers	6.678	6.450	6.669	6.367	6.754	6.375
Black workers	6.381	6.164	6.402	6.164	6.473	6.174
Wage gap	0.297	0.286	0.267	0.203	0.281	0.201
<i>Counterfactual log wage</i>						
White workers	6.691	6.396	6.701	6.325	6.796	6.377
Black workers	6.447	6.276	6.442	6.010	6.444	6.133
Counterfactual wage gap	0.243	0.120	0.259	0.315	0.352	0.245

Table 4.9: Racial real wage gap

Year	1980		1990		2000	
	HS	LS	HS	LS	HS	LS
<i>Real log wage</i>						
White workers	4.096	3.665	4.106	3.626	4.192	3.636
Black workers	3.942	3.031	3.968	3.054	4.044	3.071
Wage gap	0.154	0.633	0.139	0.572	0.148	0.565
<i>Counterfactual log wage</i>						
White workers	4.109	3.688	4.124	3.630	4.220	3.645
Black workers	4.015	3.275	4.077	3.071	4.087	3.173
Counterfactual wage gap	0.084	0.412	0.047	0.559	0.133	0.472

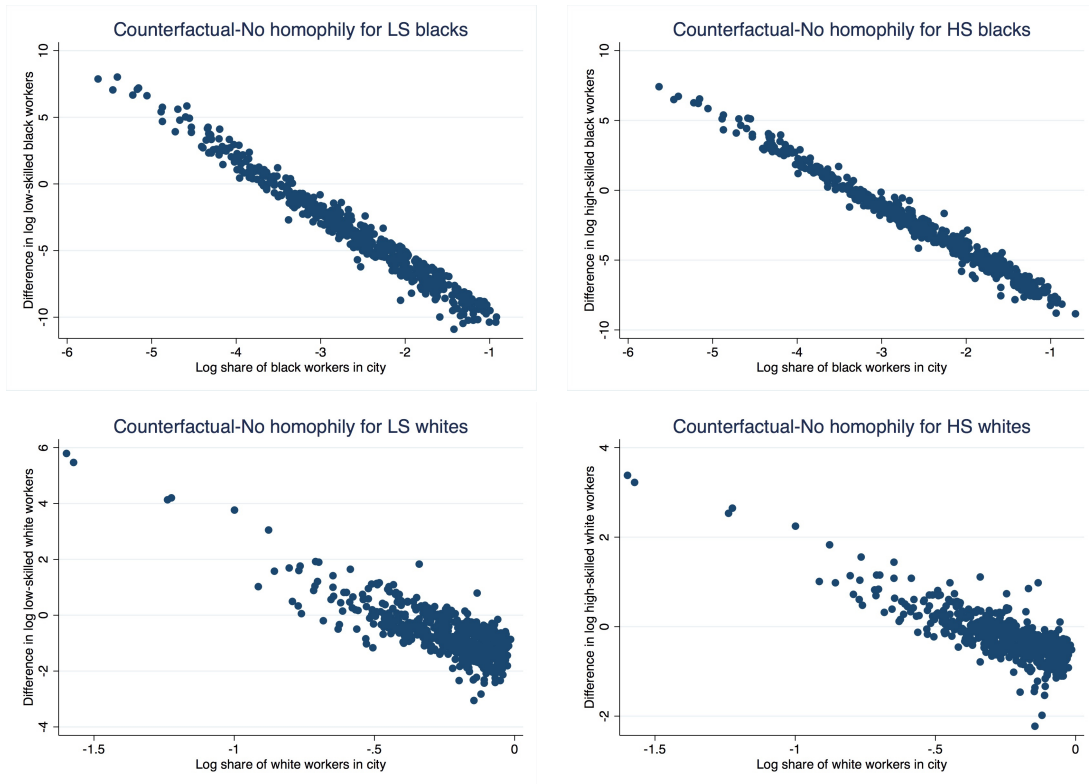


Figure 4.7: Difference between the true and counterfactual populations with no homophily, as a function of the initial shares of black and white workers.

4.6.2 REACTION TO SHOCKS

In Figures C.3 and C.4, we saw that cities tended to have a larger share of black residents as they declined in population. In this section I investigate whether my model can replicate this fact, and how much homophily and other factors contribute to this phenomenon. My simulation is in three steps, as follows:

WAGES ONLY

First I use the Bartik shocks as an exogenous shock on high and low skilled wages and on transfers, use the parameter values from model three to predict new values for δ_{it+1} , and see the new locations that workers choose as a response to the Bartik shock. Formally, these new δ s can be written as shown in Equation 4.12.

$$\hat{\delta}_{it}^z = \lambda_{it}^z (s_W \times (w_{it-1}^z + B_{it}^z) + s_T \times \hat{t}_{it} - \zeta^z r_{it-1}) + \omega^z \left(\frac{H_{it-1}^z + L_{it-1}^z}{Pop_{it-1}} \right) + \hat{\epsilon}_{it-1} \quad (4.12)$$

where $\hat{\epsilon}_{it-1}$ is computed as in section 6.1, and $\hat{t}_{it} = t_{it-1} + (\psi^W + \psi^B)B_{it}^L$. These new estimated δ s are then plugged into the conditional logit model to predict how workers from each group will be reallocated across cities as a function of the shock on wages and transfers only, as in section 6.1. I then multiply the probabilities of living in a particular MSA for each given race-skill-birth state group by the number of workers corresponding to this group at period t .

WAGES AND RENTS

Following the shock to wages, workers reallocate across space to take into account the change in wages due to the Bartik shocks. However, as they change cities, they have an effect on wages and rents. I ignore the effect on wages in this simulation since none of the parameters in the third column of my estimation are significant. I model the effect on rents in a similar way to that of wages. First, I recover a city fixed effect $\hat{\epsilon}_{it}^{CC}$ from the following equation, using the values of μ from column 3.

$$r_{it} = \mu \times \ln(HD_{it}) + \epsilon_{it}^{CC} \quad (4.13)$$

Then, I predict how rents will change following the bartik shock on wages by finding the

new housing demand, and using the equation below to predict rents \hat{r}_{it} from the new housing demand:

$$H\hat{D}_{it} = \sum_z \zeta^z \hat{Z}_{it} (w_{it-1}^z + B_{it}^z) \quad (4.14)$$

$$\hat{r}_{it} = \mu \times \ln(H\hat{D}_{it}) + \hat{\epsilon}_{it}^{CC} \quad (4.15)$$

where \hat{Z}_{it} is the predicted population in cities following the Bartik shock to wages. I then generate new δ s using Equation 4.18 :

$$\hat{\delta}_{it}^z = \lambda_{it}^z (s_W \times (w_{it-1}^z + B_{it}^z) + s_T \times \hat{t}_{it} - \zeta^z \hat{r}_{it}) + \omega^z \ln\left(\frac{H_{it-1}^z + L_{it-1}^z}{Pop_{it-1}}\right) + \hat{\epsilon}_{it-1} \quad (4.16)$$

WAGES, RENTS AND OWN-RACE PREFERENCES

Finally, I use the reaction in wages and rents to the bartik shocks to predict the change in the share of black and white workers living in a particular city. I take the \hat{Z}_{it} predicted with the δ s from the previous section, and use them to predict the share of each group living within a given city. I take immigration in period t as given. The log shares are computed as follows:

$$\frac{\hat{H}_{it}^z + \hat{L}_{it}^z}{imm_{it} + \sum_z \hat{Z}_{it}} \quad (4.17)$$

and I re-estimate a final set of δ s using the following formula:

$$\hat{\delta}_{it}^z = \lambda_{it}^z (s_W \times (w_{it-1}^z + B_{it}^z) + s_T \times \hat{t}_{it} - \zeta^z \hat{r}_{it}) + \omega^z \ln\left(\frac{\hat{H}_{it}^z + \hat{L}_{it}^z}{imm_{it} + \sum_z \hat{Z}_{it}}\right) + \hat{\epsilon}_{it-1} \quad (4.18)$$

Table 4.10: Response to shocks of the share of white workers in a city.

Model	Wages only	Wages and rents	Wages, rents and preferences
Dependent var.	Δ share of white workers	Δ share of white workers	Δ share of white workers
Skill-neutral bartik shock	1.005* (0.54)	0.862* (0.47)	5.274** (2.35)
Year FE	Yes	Yes	Yes
Observations	280	280	280

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Standard errors are clustered at the MSA level. The change in log share is computed as the difference between the estimated log share, and the 10 year lagged real share of white workers in the city.

IMPACT ON SHARE OF BLACK AND WHITE WORKERS IN MSAs

The goal of this exercise is to see how the share of black and white workers in a city respond to shocks. I present regressions of the model-generated change in shares of black and white workers on a skill-neutral bartik shock in Tables 4.10. and 4.11.¹ The simulations are consistent with what we see in Figures C.3 and C.4. If we look at the reaction of workers to shocks on wages alone, then cities sort racially based on whether the city is hit by a positive or a negative wage shock. This can be due to a combination of factors: preferences for wages are higher for low-skilled white workers than for black workers, which will influence sorting following the initial shock. The location of shocks and the distribution of birth-states across groups can also have an influence. The results show that if cities are hit by a negative shock, they tend to become more black. Incorporating the effect on rents has a mitigating influence, since lower rents compensate for the lower wages following a negative Bartik shock, and might induce some wavering workers into staying. However, once the effect on the change in the share of white or black workers is added, then the effect becomes much stronger.

¹To be more precise, the bartik shock is computed using all workers and not low-skill or high-skill workers only.

Table 4.11: Response to shocks of the share of black workers in a city.

Model	Wages only	Wages and rents	Wages, rents and preferences
Dependent var.	Δ share of black workers	Δ share of black workers	Δ share of black workers
Skill-neutral bartik shock	-6.578*** (2.41)	-6.042*** (2.32)	-27.121** (10.53)
Year FE	Yes	Yes	Yes
Observations	280	280	280

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Standard errors are clustered at the MSA level. The change in log share is computed as the difference between the estimated log share, and the 10 year lagged real share of black workers in the city.

4.7 CONCLUSION

In this article, I combined a spatial general equilibrium problem with preferences for living in MSA's with a large share of one's racial group. Despite large differences in the proportion of different racial groups living across MSAs, little work has been done on this subject. I find that these preferences matter for both racial groups studied, no matter the skill level of workers. Using simulations, I show that reducing these preferences could help to reduce the racial wage gap. One way to pursue these ideas further would be to integrate both neighbourhood and MSA choices into a single model.



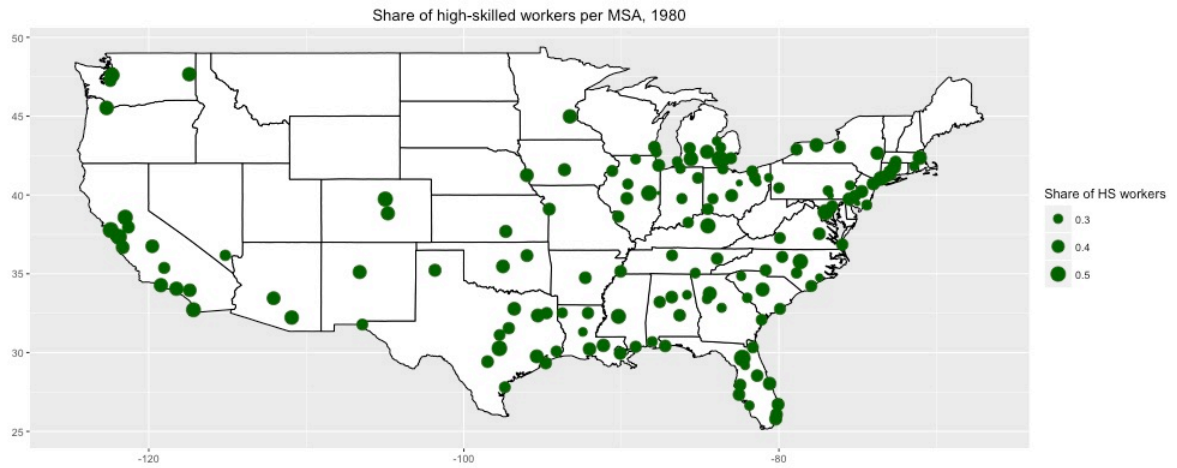
Homophily and Migration

C.1 FIGURES

Here are some additional figures that may be of interest to the reader. The first are maps that show the share of high-skilled workers in a given MSA.

The second are figures that reproduce the graphs given in the main part of the article, but for white workers only. We can see that there is no significant trend for white workers, although for high-skilled workers the relationship is downward sloping.

We can also see that both high-skilled and low-skilled white workers are less likely to stay in declining cities, although this relationship is significant only for high-skilled workers.



C.2 TABLES

C.3 PERFECT SUBSTITUTION BETWEEN WHITE AND BLACK WORKERS

Consider a more complicated version of the labour demand presented in the main body of the article. In this case, we have a two-factor CES nests 2-level, 2 factor nested CES production function, as in Ottaviano & Peri (2012) and more recently Piyapromdee(2014). Estimation of the elasticities from the CES functions can be challenging, since we cannot estimate the CES aggregates until we know all the elasticities. Card & Lemieux (2001) estimate a similar model with skill and age cohorts by using a two-step procedure.

I follow Card and Lemieux's formulation of the CES elasticities, though I allow blacks and whites to have different elasticities of substitution within skill group, so that:

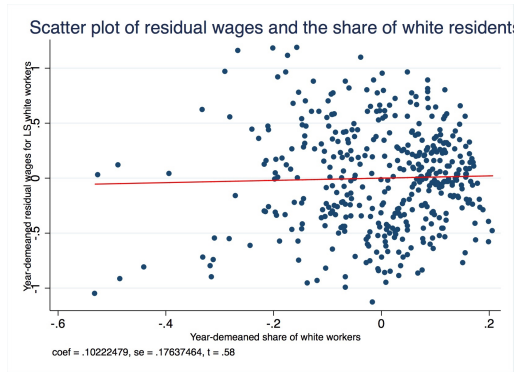
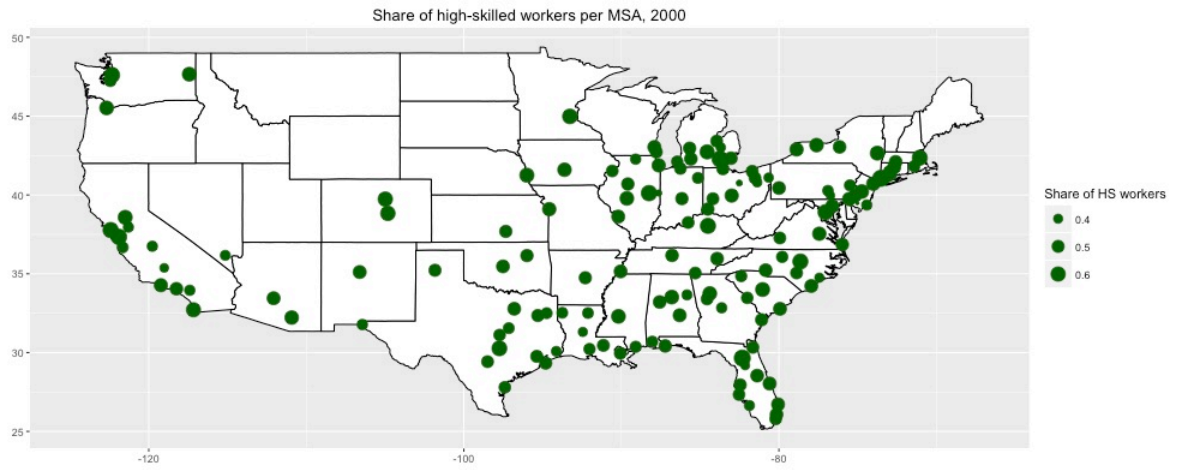


Figure C.1: Low-skilled workers

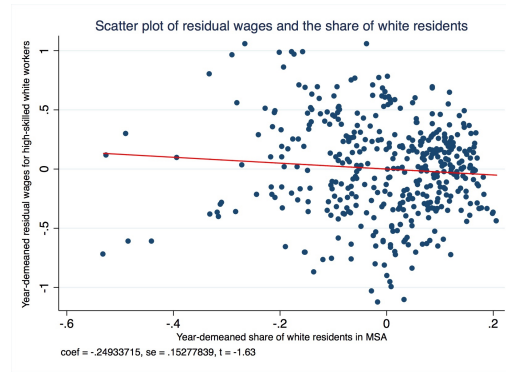


Figure C.2: High-skilled workers

$$Y_{it} = (\theta_{ht}H_{it}^\rho + \theta_{lt}L_{it}^\rho)^{\frac{1}{\rho}}$$



Figure C.3: Low-skilled workers

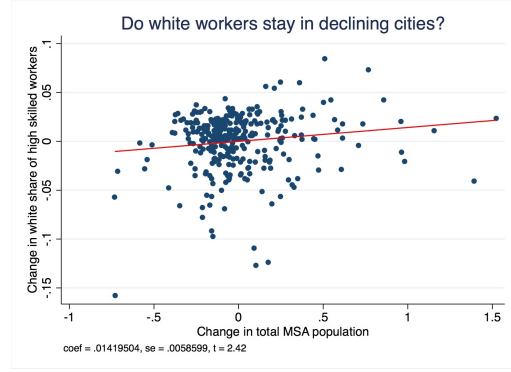


Figure C.4: High-skilled workers

$$H = (\theta_{hbt} H_{it}^{B\xi} + \theta_{hwt} H_{it}^{W\xi})^{\frac{1}{\xi}}, L = (\theta_{lbt} L_{it}^{B\phi} + \theta_{lwt} L_{it}^{W\phi})^{\frac{1}{\phi}}$$

H and L are CES aggregates of different group types, B and W, of high- and low-skilled workers respectively, the θ s are technological efficiency parameters, and ρ , ξ and ϕ are functions (e.g. $\rho = 1 - 1/\sigma_{HL}$) of the elasticities of substitution between factors. Wages are equal to marginal product, so that:

$$\frac{\partial Y_{it}}{\partial H_{it}} = \theta_{ht} H_{it}^{\rho-1} Y_t^{1-\rho}$$

$$\frac{\partial H_{it}}{\partial H_{it}^W} = \theta_{hwt} H_{it}^{W\xi-1} H_{it}^{1-\xi}$$

$$w_{it}^{HW} = \frac{\partial Y_{it}}{\partial H_{it}} \frac{\partial H_{it}}{\partial H_{it}^W} = \theta_{ht} H_{it}^{\rho-\xi} Y_t^{1-\rho} \theta_{hwt} H_{it}^{W\xi-1}$$

Here we are interested in identifying ξ . I adopt a similar approach to Card (2009), and estimate the effect of relative supply on relative wage gaps at the "city" (MSA here) level. These estimates could be biased if high wage gaps attract more high-skilled workers. Card's article solves

Table C.1: % of residents who are black in 1980

Metropolitan Area	% of residents who are black
Jackson	40.1
Memphis	36.8
Macon-Warner Robins	35.0
Augusta-Aiken	30.8
Savannah	31.8
New Orleans	30.3
Shreveport	29.7
Danville	29.7
Montgomery	29.5
Mobile	27.6
...	...
Providence-Fall River-Pawtucket	2.7
Tucson	2.6
Portland	2.3
York	2.2
Ventura-Oxnard-Simi Valley	2.2
Albuquerque	2.1
Minneapolis-St. Paul	1.9
Honolulu	1.8
Spokane	1.4
Allentown-Bethlehem-Easton	1.4

this concern by instrumenting the relative supply of skilled workers with a shift-share instrument based on the skill composition of immigrants by nationality and immigrant enclaves by city. In a similar spirit I instrument for the relative share of black and white workers by skill by using the Bartik instruments presented in the main body of the article. Consider the log wage ratio between high-skilled black and white workers:

$$\ln\left(\frac{w_{it}^{HB}}{w_{it}^{HW}}\right) = \ln\left(\frac{\theta_{hbt}}{\theta_{hwt}}\right) + (\xi - 1)\ln\left(\frac{H_{it}^B}{H_{it}^W}\right)$$

Table C.2: % of residents who are black in 2000

Metropolitan Area	% of residents who are black
Jackson	42.4
Memphis	40.0
Savannah	36.3
New Orleans	35.2
Montgomery	34.6
Shreveport	34.4
Macon-Warner Robins	34.1
Fayetteville	33.5
Danville	32.1
Augusta-Aiken	31.6
...	...
El Paso	3.1
Corpus Christi	3
San Jose	2.9
Tucson	2.8
Albuquerque	2.5
Portland	2.5
Honolulu	2.4
Allentown-Bethlehem-Easton	2.2
Ventura-Oxnard-Simi Valley	2.1
Spokane	1.2

Regressing the log mean wage¹ ratio of high-skilled blacks and whites on the log supply of high-skilled black workers over high-skilled white workers (and analogously for low-skilled workers), allow us to estimate ξ and ϕ . The ratio of productivities can be captured with time fixed effects. Following Card, I do not allow for city-fixed effects, but instead control for the city size and the college ratio for all specifications. I give estimates of ξ and ϕ in Table C.3, instrumenting the log supply with Bartik instruments in columns 2 and 4. Remember that these estimates give us the negative inverse of the elasticity, since $\xi - 1 = -\frac{1}{\sigma_{HWB}}$ so that estimates close to 0

¹Here wages are "residual wages": the city-level residuals from a regression of wages on education, occupation, age and marital status.

Table C.3: Estimation of elasticities of substitution

Dependent variable:	$\ln(w^B/w^H)$ for high-skilled workers		$\ln(w^B/w^H)$ for low-skilled workers	
Estimator	OLS	IV	OLS	IV
$\ln(B/W)$ supply	0.002 (0.00)	0.002 (0.00)	0.002 (0.00)	-0.005 (0.00)
Time fixed effects	Yes	Yes	Yes	Yes
MSA controls	Yes	Yes	Yes	Yes
Observations	432	432	276	276

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Standard errors are clustered at the MSA level. Dependent variable is the log mean wage ratio between blacks and whites in each skill group, where wages are the city-level residuals from a regression of wages on education, occupation, age and marital status. Change in relative black-white supply, measured as the log difference in employed workers of each group, is instrumented by the difference between the Bartik shock at the relevant skill-level. All specifications control for the college employment ratio and the city size.

are consistent with perfect substitution. The point estimates give us an elasticity of substitution close to 500 for both skill levels. The estimates are similar whether we consider the OLS or the IV specifications. The t statistic for the Bartik instruments in the first stage is over 35 for both regressions.

The estimated elasticity is extremely high: for similar point estimates, Card concludes that the elasticity of substitution between high-school dropouts and high-school graduates is essentially infinite. I therefore assume perfect substitution between black and white workers in the section above.

C.4 REDUCED FORM EFFECT OF THE INSTRUMENTS ON THE DEPENDENT VARIABLE

Table C.4: Are the Bartik shocks good instruments?

Dependent variable:	Δw^{HW}	Δw^{LW}	Δw^{HB}	Δw^{LB}	Δr_{it}
Estimator	OLS	OLS	OLS	OLS	OLS
Bartik shock for HS workers	3.428*** (0.67)	1.023 (0.68)	3.330*** (1.22)	1.589 (1.35)	2.153 (1.64)
Bartik shock for LS workers \times	-2.144** (0.86)	0.262 (0.95)	-0.654 (1.46)	2.471 (1.90)	-7.594*** (2.15)
Bartik shock for HS workers \times	-0.059 (0.33)	-0.178 (0.33)	-0.313 (0.67)	-0.333 (0.75)	-0.316 (0.76)
Land unavailability index					
Bartik shock for HS workers \times	0.049 (0.03)	0.043 (0.04)	0.096** (0.04)	-0.009 (0.04)	-0.016 (0.06)
Wharton index					
Bartik shock for LS workers \times	-0.167 (0.43)	-0.190 (0.52)	-0.424 (0.83)	-0.789 (1.02)	0.510 (1.10)
Land unavailability index					
Bartik shock for LS workers \times	-0.124** (0.05)	-0.117* (0.07)	-0.134* (0.07)	-0.018 (0.07)	-0.171 (0.17)
Wharton index					
Bartik shock for LS workers \times	-0.311*** (0.05)	-0.263*** (0.06)	-0.205** (0.10)	-0.152 (0.12)	-0.839*** (0.15)
110. share of immigrants					
Bartik shock for LS workers \times	0.236*** (0.04)	0.170*** (0.04)	0.145** (0.07)	0.173** (0.08)	0.253** (0.10)
110. share of immigrants					
Observations	280	280	280	280	280
R^2	0.64	0.64	0.18	0.15	0.38

Bibliography

Abrevaya, J. & Hamermesh, D. S. (2012). Charity and Favoritism in the Field: Are Female Economists Nicer (To Each Other)? *The Review of Economics and Statistics*, 94(1), 202–207.

Ahern, K. & Dittmar, A. K. (2012). The changing of the boards: The impact on firm valuation of mandated female board representation. *The Quarterly Journal of Economics*, 127(1), 137–197.

Akerlof, G. A. & Kranton, R. E. (2000). Economics and identity. *The Quarterly Journal of Economics*, 115(3), 715–753.

Alesina, A., Baqir, R., & Easterly, W. (1999). Public goods and ethnic divisions. *The Quarterly Journal of Economics*, 114(4), 1243–1284.

Algan, Y., Hémet, C., & Laitin, D. D. (2016). The Social Effects of Ethnic Diversity at the Local Level: A Natural Experiment with Exogenous Residential Allocation. *Journal of Political Economy*, 124(3), 696–733.

Altonji, J. G. & Blank, R. M. (1999). Race and gender in the labor market. *Handbook of Labor Economics*, 3, 3143–3259.

Andreff, W. (2014). Building blocks for a disequilibrium model of a european team sports league. *International Journal of Sport Finance*, 9, 20–38.

Andreff, W. (2015). Governance of professional team sports clubs: Agency problems and soft budget constraints. *Disequilibrium Sports Economics*, Elgar: London, (pp. 175–228).

Angrist, J. D. & Pischke, J.-S. (2009). *Mostly harmless econometrics: An empiricist's companion*. Princeton university press.

Antecol, H., Bedard, K., & Stearns, J. (2016). *Equal but Inequitable: Who Benefits from Gender-Neutral Tenure Clock Stopping Policies?* IZA Discussion Papers 9904, Institute for the Study of Labor (IZA).

Arellano, M. & Bond, S. (1991). Some tests of specification for panel data: Monte carlo evidence and an application to employment equations. *Review of Economic Studies*, 58(2), 277–297.

Arrow, K. J. (1972). *Racial Discrimination in Economic Life*, chapter Some Mathematical Models of Race Discrimination in the Labor Market, (pp. 187–294). Lexington, Mass.: D.C. Heath.

Arrow, K. J. (1973). *Discrimination in Labor Markets*, chapter The theory of discrimination, (pp. 3–33). Princeton University Press.

Arrow, K. J. (1998). What has economics to say about racial discrimination? *Journal of economic perspectives*, 12(2), 91–100.

Ashenfelter, O. & Hannan, T. (1986). Sex Discrimination and Product Market Competition: The Case of the Banking Industry. *Quarterly Journal of Economics*, 101(1), 149–73.

Azmat, G. (2014). Gender diversity in teams. *IZA World of Labor*.

Baert, S., Cockx, B., Gheyle, N., & Vandamme, C. (2014). Is there less discrimination in occupations where recruitment is difficult? *Industrial and Labor Relations Review*, Forthcoming.

Bagues, M., Sylos-Labini, M., & Zinovyeva, N. (2017). Does the gender composition of scientific committees matter? *American Economic Review*, 107(4), 1207–38.

Bagues, M. F. & Esteve-Volart, B. (2010). Can Gender Parity Break the Glass Ceiling? Evidence from a Repeated Randomized Experiment. *Review of Economic Studies*, 77(4), 1301–1328.

Bagues, M. F. & Zinovyeva, N. (2015). The role of connections in academic promotions. *American Economic Journal: Applied Economics*, 124(5).

Bartik, T. J. (1991). *Who Benefits from State and Local Economic Development Policies?* Number wbsle in Books from Upjohn Press. W.E. Upjohn Institute for Employment Research.

Bayer, P., Fang, H., & McMillan, R. (2014). Separate when equal? Racial inequality and residential segregation. *Journal of Urban Economics*, 82(C), 32–48.

Bayer, P., Ferreira, F., & McMillan, R. (2007). A unified framework for measuring preferences for schools and neighborhoods. *Journal of Political Economy*, 115(4), 588–638.

Bayer, P., McMillan, R., & Rueben, K. S. (2004). What drives racial segregation? New evidence using Census microdata. *Journal of Urban Economics*, 56(3), 514–535.

Bayer, P. & Timmins, C. (2005). On the equilibrium properties of locational sorting models. *Journal of Urban Economics*, 57(3), 462–477.

Beaman, L., Chattopadhyay, R., Duflo, E., Pande, R., & Topalova, P. (2009). Powerful women: does exposure reduce bias? *The Quarterly Journal of Economics*, 124(4), 1497–1540.

Beaman, L., Duflo, E., Pande, R., & Topalova, P. (2012). Female leadership raises aspirations and educational attainment for girls: A policy experiment in india. *Science*, 335(6068), 582–586.

Becker, G. S. (1957). *The economics of discrimination*. University of Chicago Press.

Becker, S. O., Fetzner, T., et al. (2016). *Does Migration Cause Extreme Voting*. Technical report, Competitive Advantage in the Global Economy (CAGE).

Beggs, S., Cardell, S., & Hausman, J. (1981). Assessing the potential demand for electric cars. *Journal of Econometrics*, 17(1), 1–19.

Berry, S., Levinsohn, J., & Pakes, A. (1995). Automobile prices in market equilibrium. *Econometrica: Journal of the Econometric Society*, (pp. 841–890).

Bertrand, M. (2011). New perspectives on gender. In *Handbook of Labor Economics*, volume 4 (pp. 1543–1590). Elsevier.

- Bertrand, M., Black, S. E., Jensen, S., & Lleras-Muney, A. (2014). *Breaking the glass ceiling? The effect of board quotas on female labor market outcomes in Norway*. Technical report, National Bureau of Economic Research.
- Biddle, J. E. & Hamermesh, D. S. (2013). Wage discrimination over the business cycle. *IZA Journal of Labor Policy*, 2(1), 1–19.
- Black, D. A. (1995). Discrimination in an equilibrium search model. *Journal of labor Economics*, 13(2), 309–334.
- Blanchard, O. J. & Katz, L. F. (1992). Regional Evolutions. *Brookings Papers on Economic Activity*, 23(1), 1–76.
- Boring, A. (2017). Gender biases in student evaluations of teaching. *Journal of Public Economics*, 145(C), 27–41.
- Borjas, G. J. (2003). The labor demand curve is downward sloping: Reexamining the impact of immigration on the labor market. *The Quarterly Journal of Economics*, 118(4), 1335–1374.
- Borjas, G. J. & Monras, J. (2017). The labour market consequences of refugee supply shocks. *Economic Policy*, 32(91), 361–413.
- Bound, J. & Holzer, H. J. (2000). Demand Shifts, Population Adjustments, and Labor Market Outcomes during the 1980s. *Journal of Labor Economics*, 18(1), 20–54.

- Boustan, L. P. (2009). Competition in the Promised Land: Black Migration and Racial Wage Convergence in the North, 1940–1970. *The Journal of Economic History*, 69(03), 755–782.
- Bowlus, A. J. & Eckstein, Z. (2002). Discrimination and skill differences in an equilibrium search model. *International Economic Review*, 43(4), 1309–1345.
- Bowlus, A. J., Kiefer, N. M., & Neumann, G. R. (2001). Equilibrium search models and the transition from school to work. *International Economic Review*, 42(2), 317–343.
- Breda, T. & Hillion, M. (2016). Teaching accreditation exams reveal grading biases favor women in male-dominated disciplines in france. *Science*, 353(6298), 474–478.
- Breda, T. & Ly, S. T. (2015). Professors in Core Science Fields Are Not Always Biased against Women: Evidence from France. *American Economic Journal: Applied Economics*, 7(4), 53–75.
- Broder, I. E. (1993). Review of nsf economics proposals: Gender and institutional patterns. *The American Economic Review*, 83(4), 964–970.
- Burdett, K. & Mortensen, D. T. (1998). Wage differentials, employer size, and unemployment. *International Economic Review*, 39(2), 257–273.
- Buser, T., Niederle, M., & Oosterbeek, H. (2014). Gender, competitiveness, and career choices. *The Quarterly Journal of Economics*, 129(3), 1409–1447.

Cadena, B. C. & Kovak, B. K. (2016). Immigrants equilibrate local labor markets: Evidence from the great recession. *American Economic Journal: Applied Economics*, 8(1), 257–90.

Card, D. (2009). Immigration and inequality. *American Economic Review*, 99(2), 1–21.

Card, D. & Lemieux, T. (2001). Can Falling Supply Explain the Rising Return to College for Younger Men? A Cohort-Based Analysis. *The Quarterly Journal of Economics*, 116(2), 705–746.

Card, D., Mas, A., & Rothstein, J. (2008). Tipping and the Dynamics of Segregation. *The Quarterly Journal of Economics*, 123(1), 177–218.

Casas-Arce, P. & Saiz, A. (2015). Women and power: unpopular, unwilling, or held back? *Journal of Political Economy*, 123(3), 641–669.

Charles, K. K. & Guryan, J. (2011). Studying discrimination: Fundamental challenges and recent progress. *Annual Review of Economics*, 3(1), 479–511.

Chenu, A. & Martin, O. (2016). Le plafond de verre chez les enseignants-chercheurs en sociologie et démographie. *Travail, genre et sociétés*, (2), 133–156.

Combes, P.-P., Linnemer, L., & Visser, M. (2008). Publish or peer-rich? The role of skills and networks in hiring economics professors. *Labour Economics*, 15(3), 423–441.

Cutler, D. M., Glaeser, E. L., & Vigdor, J. L. (1999). The Rise and Decline of the American Ghetto. *Journal of Political Economy*, 107(3), 455–506.

Cymrot, D. J. (1985). Does competition lessen discrimination? some evidence. *The Journal of Human Resources*, 20(4), 605–612.

Davis, M. A. & Heathcote, J. (2007). The price and quantity of residential land in the United States. *Journal of Monetary Economics*, 54(8), 2595–2620.

Davis, M. A. & Ortalo-Magne, F. (2011). Household Expenditures, Wages, Rents. *Review of Economic Dynamics*, 14(2), 248–261.

Davis, M. A. & Palumbo, M. G. (2008). The price of residential land in large US cities. *Journal of Urban Economics*, 63(1), 352–384.

Diamond, R. (2013). *The Determinants and Welfare Implications of US Workers' Diverging Location Choices by Skill: 1980-2000*. Research Papers 3143, Stanford University, Graduate School of Business.

Duflo, E. (2000). *Schooling and labor market consequences of school construction in Indonesia: Evidence from an unusual policy experiment*. Technical report, National Bureau of Economic Research.

Dustmann, C., Schönberg, U., & Stuhler, J. (2016). The impact of immigration: Why do studies reach such different results? *Journal of Economic Perspectives*, 30(4), 31–56.

- Edo, A., Jacquemet, N., & Yannelis, C. (2015). Language skills and homophilous hiring discrimination: Evidence from gender and racially differentiated applications. *Review of Economics of the Household*, (pp. 1–28).
- Ellemers, N., Van den Heuvel, H., de Gilder, D., Maass, A., & Bonvini, A. (2004). The underrepresentation of women in science: Differential commitment or the queen bee syndrome? *British Journal of Social Psychology*, 43(3), 315–338.
- Feld, J., Salamanca, N., & Hamermesh, D. S. (2016). Endophilia or exophobia: beyond discrimination. *The Economic Journal*, 126(594), 1503–1527.
- Flood, S., King, M., Ruggles, S., & Warren, J. R. (2015). Integrated public use microdata series, current population survey: Version 4.0. [machine-readable database]. *Minneapolis: University of Minnesota*.
- Garcia-del-Barrio, P. & Szymanski, S. (2009). Goal! profit maximization versus win maximization in soccer. *Review of Industrial Organization*, 34(1), 45–68.
- Geuna, A. & Shibayama, S. (2015). Moving out of academic research: why do scientists stop doing research? In *Global mobility of research scientists* (pp. 271–303). Elsevier.
- Gilens, M. (1996). "race coding" and white opposition to welfare. *The American Political Science Review*, 90(3), 593–604.

Guerrieri, V., Hartley, D., & Hurst, E. (2013). Endogenous gentrification and housing price dynamics. *Journal of Public Economics*, 100(C), 45–60.

Guryan, J. & Charles, K. K. (2013). Taste-based or statistical discrimination: The economics of discrimination returns to its roots. *The Economic Journal*, 123(572), F417–F432.

Gyourko, J., Saiz, A., & Summers, A. (2008). A new measure of the local regulatory environment for housing markets: The wharton residential land use regulatory index. *Urban Studies*, 45(3), 693–729.

Habyarimana, J., Humphreys, M., Posner, D. N., & Weinstein, J. M. (2007). Why does ethnic diversity undermine public goods provision? *American Political Science Review*, 101(04), 709–725.

Haldane, A. G. & Turrell, A. E. (2018). An interdisciplinary model for macroeconomics. *Oxford Review of Economic Policy*, 34(1-2), 219–251.

Hardle, W. & Mammen, E. (1993). Comparing nonparametric versus parametric regression fits. *The Annals of Statistics*, (pp. 1926–1947).

Hausman, J. A. & Ruud, P. A. (1987). Specifying and testing econometric models for rank-ordered data. *Journal of Econometrics*, 34(1-2), 83–104.

Ianni, G. (2010). Scholar h-index calculator.

Johnson, S. K. & David, H. R. (2016). Women and minorities are penalized for promoting diversity. *Harvard Business Review*. Retrieved from <https://hbr.org/2016/03/women-and-minorities-are-penalized-for-promoting-diversity>.

Katz, L. F. & Murphy, K. M. (1992). Changes in relative wages, 1963–1987: supply and demand factors. *The Quarterly Journal of Economics*, 107(1), 35–78.

Kleven, H. J., Landais, C., & Saez, E. (2013). Taxation and international migration of superstars: Evidence from the european football market. *American Economic Review*, 103(5), 1892–1924.

Lang, K. & Lehmann, J.-Y. K. (2012). Racial discrimination in the labor market: Theory and empirics. *Journal of Economic Literature*, 50(4), 959–1006.

Levine, R., Levkov, A., & Rubinstein, Y. (2008). *Racial Discrimination and Competition*. Technical Report 14273, National Bureau of Economic Research.

Li, D. (2017). Expertise versus bias in evaluation: Evidence from the nih. *American Economic Journal: Applied Economics*, 9(2), 60–92.

Lincoln, A. E., Pincus, S., Koster, J. B., & Leboy, P. S. (2012). The matilda effect in science: Awards and prizes in the us, 1990s and 2000s. *Social Studies of Science*, 42(2), 307–320.

List, J. A. & Rasul, I. (2011). Field experiments in labor economics. *Handbook of Labor Economics*, 4, 103–228.

- Lkhagvasuren, D. (2012). Big locational unemployment differences despite high labor mobility. *Journal of Monetary Economics*, 59(8), 798–814.
- Manacorda, M., Manning, A., & Wadsworth, J. (2012). The Impact Of Immigration On The Structure Of Wages: Theory And Evidence From Britain. *Journal of the European Economic Association*, 10(1), 120–151.
- Manning, A. (2003). The real thin theory: monopsony in modern labour markets. *Labour Economics*, 10(2), 105–131.
- McFadden, D. (1973). Conditional logit analysis of qualitative choice behavior. *Frontiers in Econometrics*.
- Michaels, G., Rauch, F., & Redding, S. J. (2012). Urbanization and structural transformation. *The Quarterly Journal of Economics*, 127(2), 535–586.
- Naidu, S. (2010). Recruitment restrictions and labor markets: Evidence from the postbellum us south. *Journal of Labor Economics*, 28(2), 413–445.
- Naidu, S., Nyarko, Y., & Wang, S.-Y. (2016). Monopsony power in migrant labor markets: Evidence from the united arab emirates. *Journal of Political Economy*, 124(6), 1735–1792.
- Nimmesgern, H. (2016). Why are women underrepresented in stem fields? *Chemistry – A European Journal*, 22(11), 3529–3530.

Notowidigdo, M. J. (2011). *The Incidence of Local Labor Demand Shocks*. NBER Working Papers 17167, National Bureau of Economic Research, Inc.

Ottaviano, G. I. P. & Peri, G. (2012). Rethinking The Effect Of Immigration On Wages. *Journal of the European Economic Association*, 10(1), 152–197.

Palacios-Huerta, I. (2014). *Beautiful Game Theory: How Soccer Can Help Economics*. Princeton University Press.

Paola, M. D. & Scoppa, V. (2015). Gender Discrimination and Evaluators' Gender: Evidence from Italian Academia. *Economica*, 82(325), 162–188.

Parsons, C. A., Sulaeman, J., Yates, M. C., & Hamermesh, D. S. (2011). Strike three: Discrimination, incentives, and evaluation. *American Economic Review*, 101(4), 1410–1435.

Peri, G., Shih, K., & Sparber, C. (2015). *Foreign and Native Skilled Workers: What Can We Learn from H-1B Lotteries?* Technical report, National Bureau of Economic Research.

Piyapromdee, S. (2014). *The impact of immigration on wages, internal migration and welfare*. Technical report, Working paper.

Preston, I. & Szymanski, S. (2000). Racial discrimination in english football. *Scottish Journal of Political Economy*, 47(4), 342–363.

- Price, J. & Wolfers, J. (2010). Racial discrimination among nba referees. *Quarterly Journal of Economics*, 125(4), 1859–1887.
- Raphael, S. & Riker, D. A. (1999). Geographic Mobility, Race, and Wage Differentials. *Journal of Urban Economics*, 45(1), 17–46.
- Rathelot, R. & Safi, M. (2014). Local ethnic composition and natives' and immigrants' geographic mobility in france, 1982–1999. *American Sociological Review*, 79(1), 43–64.
- Roback, J. (1982). Wages, rents, and the quality of life. *The Journal of Political Economy*, (pp. 1257–1278).
- Robinson, P. M. (1988). Root- N-Consistent Semiparametric Regression. *Econometrica*, 56(4), 931–954.
- Roodman, D. (2009). A note on the theme of too many instruments. *Oxford Bulletin of Economics and Statistics*, 71(1), 135–158.
- Rosen, S. (1974). Hedonic prices and implicit markets: product differentiation in pure competition. *Journal of political economy*, 82(1), 34–55.
- Saiz, A. (2010). The Geographic Determinants of Housing Supply. *The Quarterly Journal of Economics*, 125(3), 1253–1296.

Sapienza, P. & Zingales, L. (2013). Economic experts versus average americans. *American Economic Review*, 103(3), 636–42.

Schaffner, B. F., MacWilliams, M., & Nteta, T. (2018). Explaining white polarization in the 2016 vote for president: The sobering role of racism and sexism. *Political Science Quarterly*.

Schelling, T. C. (1969). Models of segregation. *The American Economic Review*, 59(2), 488–493.

Schelling, T. C. (1971). Dynamic models of segregation. *Journal of Mathematical Sociology*, 1(2), 143–186.

Serrato, J. C. S. & Wingender, P. (2011). *Estimating the incidence of government spending*. Technical report, Mimeo, 2011. 1.

Staiger, D. & Stock, J. (1997). Instrumental variables regression with weak instruments. *Econometrica*, vol.65(no.3), 557–586.

Staines, G., Tavis, C., & Jayaratne, T. E. (1974). The queen bee syndrome. *Psychology Today*, 7(8), 55–60.

Starr, E., Bishara, N., & Prescott, J. (2015). Noncompetes in the us labor force. *SSRN Working Paper*.

Stigler, G. J. & Becker, G. S. (1977). De gustibus non est disputandum. *The American Economic Review*, 67(2), 76–90.

Suchow, J. (2016). Dissertate. LaTeX Template.

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 economic design of sporting contests. *Journal of economic literature*, 41(4), 1137–1187.

Thurstone, L. L. (1927). A law of comparative judgement. *Psychological Review*, 34, 278–286.

Topel, R. H. (1986). Local Labor Markets. *Journal of Political Economy*, 94(3), S111–43.

Verardi, V. & Debarsy, N. (2012). SEMIPAR: Stata module to compute Robinson’s (1988) semi-parametric regression estimator. Statistical Software Components, Boston College Department of Economics.

Vernos, I. (2013). Research management: Quotas are questionable. *Nature*, 495(7439), 39–39.

Williams, W. M. & Ceci, S. J. (2015). National hiring experiments reveal 2: 1 faculty preference for women on stem tenure track. *Proceedings of the National Academy of Sciences*, 112(17), 5360–5365.

Wilson, G. & Nielsen, A. L. (2011). “color coding” and support for social policy spending: Assessing the parameters among whites. *The ANNALS of the American Academy of Political and Social Science*, 634(1), 174–189.

Zeltzer, D. (2015). Gender homophily in referral networks: Consequences for the medicare physician pay gap. *Working Paper*.