Mismatch in the workplace: Exploring racial gaps in labor market outcomes under the mismatch hypothesis using the After the JD study.

Christopher Smith *

June 9, 2023

Field and Classification

Primary [Secondary] Field: Empirical Labor [Income Inequality]
JEL Classifications: J3, J4, J7

Abstract

This paper explores the implications of the mismatch hypothesis in the context of the labor market using a survey on newly licensed US lawyers called the After the JD Study. Using a triple difference approach, I measure the impact of diversity quotas on marginal minority workers' future salaries, promotion rates, and leaving rates for occupation and job. With middling statistical power, my findings are in line with the mismatch hypothesis in that beneficiaries of the diversity quota policy are made

*Christopher Smith, Purdue University. Email: smit3601@purdue.edu. I would like to thank the Purdue University Research Center in Economics (PURCE) for the research grant which made this paper possible. I would also like to thank Tim Bond and Victoria Prowse for their expert advice and guidance. Without them this project would surely not exist. I also thank Ben Artz, Joshua Foster, Dave Fuller, and Marianne Johnson for their invaluable comments.

ex-ante worse off. My findings are also in line with recent literature on diminishing racial outcome gaps by skill.

1 Introduction

In this paper, I explore mismatching effects due to quota-based diversity policies at the firm on the outcomes of beneficiary workers, such as promotion rates, future salary, job tenure and career retention. Using a unique data set on newly licensed lawyers in the United States called the After the JD Study (AJD throughout), I develop a triple difference econometric model (DDD throughout) that allows me to obtain causal estimates of the impact of a diversity quota on the beneficiary group. The beneficiary group of an diversity quota policy are those minority workers who would not have been hired by a prestigious firm in the absence of the policy but are hired in its presence. With varying statistical strength, I find that beneficiaries of the policy are largely negatively impacted by it. Specifically, they have lower future salaries, lower promotion rates, are more likely to leave their profession, and endure shorter tenure. My results align with the theoretical and empirical predictions of Cavounidis, Lang, and Weinstein (2019) where the minority work force is heavily monitored and churned. My results are also consistent with the theoretical predictions of Fryer (2007) and Lehmann (2011) where workers are less likely to be promoted due to supervisor beliefs about the minority worker caused by statistical discrimination. Finally, my work is in line with the mismatch hypothesis where beneficiaries of affirmative action are actually made worse off due to a poor match with the newly attainable university (the seminal paper is Sander (2004)). Indeed, this paper uncovers what appears to be an extreme form of mismatch where, on the extensive margin, low skill minority workers impacted by the policy are more likely to leave their job and more likely to leave their profession. Furthermore, this paper shows that there are racial gaps in outcomes even for highly educated workers (lawyers).

In contrast to my paper which conditions on labor market characteristics and outcomes, *the mismatch hypothesis* states that university students who should benefit from an affirmative action policy are actually made ex ante worse off because they are mismatched with the newly attainable and more prestigious university (Sander 2004). More precisely, beneficiary students, those who would not have qualified to attend a prestigious university in the absence affirmative action, incur a cost due to a poor fit with the prestigious university. According to the mismatch hypothesis, this cost

dominates the benefit from attending the prestigious university in the presence of affirmative action. Contrary to the goal of affirmative action, beneficiaries are actually made worse off. Ex ante mismatch, also known as local mismatch, is possible under fully rational students and arises due to an asymmetry in information (Arcidiacono et al. 2011). These asymmetries are such things as admissions data and qualification minimums that the university knows but the student does not. In my paper these are a prestigious firm's diversity quotas and private information about a worker's quality. Note that while the worker has access to the firm's quota information, their signal is nonetheless muddled as they are unsure if they are the beneficiary of this hiring policy or if they would have qualified either way. Subsequent literature on the mismatch hypothesis focuses on educational success and initial career earnings as outcomes. This literature remains contentious and divided in whether the mismatch hypothesis holds. Some studies find significant, negative impacts of affirmative action due to mismatch (Sander 2004, 2005, and Williams 2013, e.g.), while others state no statistically significant impact (Ho 2005, Rothstein and Yoon 2009, e.g.). My goal is to explore this question in the context of the labor market exclusively as well as uncover some stylized facts about a highly educated sample of the labor force in the US.

My strategy consists of three important pieces: (i) a unique data set, (ii) the structure of law firm prestige rankings, and (iii) a novel DDD specification. (i) The AJD is a rich, longitudinal study of newly minted lawyers in the US. This study contains demographic and family characteristics as well as undergraduate and graduate school performance. Furthermore, the study contains an extensive questionnaire on both quantitative characteristics such as salary, promotion, and firm size and qualitative characteristics such as job, education, and personal life satisfaction. The data set follows lawyers from approximately their first year to their twelfth year of employment after passing the Bar exam.[1] Because of this richness, I am able to identify low and high skill workers as well as firms that do and don't practice some form of affirmative action. (ii) Law firms are structured such that incoming attorneys are expected to

[1]The data used in this paper comes in 3 waves, each spaced about 6 years apart and I use the first two waves. The first wave was sent out between 1998 and 2000. The second took place between the years 2006 and 2008. The final wave was completed between 2011 and 2012.

compete for a partnership (promotion) in their first 5 to 10 years of employment. If they do not land a partnership they will either be laid off or resign. Law firms are also extremely competitive both within the firm and outside. That is, prestigious firms are also extremely large and compete for placement in rankings (vault.com for example). There are many characteristics by which law firms are ranked but diversity is one of utmost importance.[2] Prestigious and large firms are incentivized to boost minority representation in order to perform well in these rankings as prestige brings wealthier clients.[3] (iii) Finally, I am able to compare lawyers who are identical in their observables but differ in their decisions to attend a prestigious or non-prestigious law firm by exploiting a DDD of outcomes. The first difference is that in the outcomes for low-skill minorities versus low-skill non-minorities. The second difference is that between workers who work at large firms versus those who work at non-large firms. This DDD, under a set of assumptions to be detailed later, identifies the effect of being a beneficiary of the affirmative action policy. This procedure differences out much of the endogeneity between the two racial groups if it is present.[4]

The paper proceeds as follows: Section 2 reviews the relevant literature both on the mismatch hypothesis as well as employer-employee mismatch and statistical discrimination. Section 3 describes the dataset structure and relevant summary measures. Section 4 details the econometric model and identification strategy. Section 5 presents and discusses results. Section 6 concludes.

2 Literature

The mismatch hypothesis, a term coined by Sander (2004) using the Law School Admissions Council data set called the Bar Passage Study (BPS), suggests that some minority students who are lifted up by policies like affirmative action to attend higher ranked universities, are actually, ex ante, worse off than they would have been without the policy. The line of thinking is that if a student qualifies for some university,

[2]

[3]See the top ten US firms as ranked by vault.com for example. I refrain from naming firms in this paper for obvious reasons.

[4]I address some threats to identification later, barring in mind concerns in Lang and Kahn-Lang (2020)

then they are well matched, meaning the quality of the university and the student are equivalent. However, if the student arbitrarily gets in to a higher quality university then they might underperform because they are not of equivalent quality and match. Mismatch is also sometimes referred to as a quality versus fit tradeoff (Arcidiacono and Lovenheim (2016)). Regardless of name, the mismatch hypothesis posits that affirmative action arbitrarily distorts labor market (or educational market) equilibrium which has a net negative impact on minority outcomes.

Mismatch in the context of affirmative action is a controversial and difficult topic to study. There is conflicting empirical evidence whether mismatch is even present. The primary data sets used to study this topic are focused on university students and lawyers (e.g. the Bar Passage Study and the NLSY72). Many argue that Sander (2004) suffers from selection on unobservables bias and mis-specification of the econometric model. Some early papers which try to rectify this are Ayres and Brooks (2005) and Ho (2005), both of which take a "selection" approach. That is, they estimate the effects of mismatch by comparing students who are identical in terms of observables but differ in their decision to go to a selective school or not. These papers, among others (Chambers et al. (2005) and Barnes (2007)) struggle to show mismatch in the BPS data set. Clearly, as pointed out in Arcidiacono et al. (2011) and Rothstein and Yoon (2009), there are significant issues with the previous papers' approaches. Namely, students who select into the more selective university are likely to carry better unobservable characteristics. For this reason, Rothstein and Yoon (2009) argue that due to the nature of the question and the non-experimental structure of the available data, mismatch is not estimable. Rothstein and Yoon take a conservative approach and try to estimate mismatch using two methods. One compares a student with those in her own tier to estimate a between students' difference in differences mismatch parameter. The other explores a two stage least squares estimator with race as the instrument. They find that, using both approaches, one cannot conclude that mismatch is present in the data. Unfortunately, this still does not address the theoretical question of whether mismatch is occurring. In contrast, Williams (2013) uses a similar approach with the Bar Passage Study and finds statistically significant evidence to suggests that minority students are less likely to graduate from law school and less likely to pass the bar examination. Williams uses a standard distance match-

ing parameter in order to measure the effect of the "fit" between a university and a student. The issue with this measure is it does not necessarily detect mismatched individuals. It may simply pick out the bad students. In other words, it is ambiguous why they have a poor fit and so impossible to state that negative outcomes are from mismatch.

Arcidiacono (2005) is one of the only papers in this literature that explores labor market outcomes (initial wages out of college) in the context of mismatch. He uses the NLS72 data set to structurally estimate the effects of affirmative action in higher education and finds no evidence of changes in wages. Furthermore, he finds that affirmative action can lead to an increase in minority student enrollment at universities, a necessary, but unsurprising, condition for mismatch. Arcidiacano et al. (2011) define mismatch as an ex ante realization on the students' behalves, that they have made a mistake under the asymmetric environment. This is similar to the approach that I take in this paper, where there are equilibria that support non-affirmative action and non-minority mismatch. Along with a unique dataset gathered by Duke University, Arcidiacono et al. (2011) use a non-parametric estimation procedure developed by Kotlarski (1967) to address mismatch at the university level. This method requires a unique data set which contains ex ante and ex post observations. Unfortunately, this isn't something that most labor market data contains. They find evidence to support the necessary conditions needed for mismatch to occur at the university level. However, they do not conclude that there is mismatch.

Arcidiacano et al. (2014) explore affirmative action in the context of university-student fit. They use a quasi-experiment involving a law in California, which suddenly prevented universities from using race in admissions decisions. They use a Dale-Kruger selection-adjustment to estimate the pre and post-law change outcomes for minority students. They find positive evidence for mismatch. Those universities that are higher ranked tend to reward those highly ranked students. However, for low-ranked students at these high-ranked universities, the effect is actually detrimental. They find that when there is no affirmative action, students are more well-matched with their universities and do not experience as many negative outcomes as a result. This is very closely related to my paper and theoretical framework. In equilibrium

with no affirmative action, there are mismatches; however, this mismatching is exacerbated under affirmative action.

The primary focus of this paper is the mismatch hypothesis in the context of previous literature but statistical discrimination that arises out of affirmative action is an important theoretical implication. While I do not provide a theoretical framework, Fryer (2007) provides a robust foundation for this paper. Lower hiring standards for minorities leads to lower average quality during later phases of employment like promotion. This prior of quality leads to minorities not being promoted as often as their non-minority counterparts by managers at the company. Even more closely related to my paper in this regard and the source for which I discovered the AJD is Lehmann (2011) in which they extend Fryer's theoretical framework and explore lawyer early career task assignments and outcomes. They find that due to some statistical discrimination minorities are more likely to get "bad" tasks involving little interaction with firm partners. However, because they face no discrimination, non-minorities are more commonly assigned to tasks which allow them to interact with partners. This interaction leads to better promotion rates among non-minorities than minorities. My work differs from Lehmann's in that I am specifically interested in outcomes and quality, not task assignments. My paper is also founded in the mismatch hypothesis literature whereas Lehmann's is founded in the statistical discrimination literature. However, both this paper and Lehmann's use the AJD study.

One might also be interested in matching literature for this question considering the word mismatch is scattered throughout this paper. While the standard matching theory is concerned with heterogeneous firms and workers, and even endogenous choice, it is not equipped, to my knowledge, to manage discrimination topics (see, e.g., Abowde et al. (2003)). For example, the aforementioned paper by Abowde et al. does an excellent job at explaining wage differentials in industries. However, these differentials are caused by differences in underlying skill and human capital. In this paper, I am mainly concerned with an entirely exogenous difference, race. Of further note is the educational and skill mismatch as reviewed by Sattinger (2012) in which the returns to over or under education are explored. Again, this literature is not concerned with discrimination. Black et al. (2006) find that the race wage differential

tends to disappear at higher levels of education. Bjerk (2007) also finds evidence to suggest this claim. These are both important papers to keep in mind as the study herein is concerned with highly educated individuals. However, I am interested in exploring the previously mentioned bodies of literature with respect to racial gaps at a later time.

3 Data

3.1 Features of Law Practice in the US

There are two key features of Law practice in the US that I have only mentioned in passing thus far. The first feature provides framework for a necessary assumption of my identification strategy. That is, prestigious law firms have an incentive to boost their diversity score by increasing minority share of their workforce. This incentive is due to the benefits that a prestigious firm will gain if they rank high in law firm rankings. Indeed, almost all large firms employing more than 250 lawyers post diversity statements on their company websites committing to diversity practices. Examples of these types of policies are often focused on diversity share (i.e. 30% minority work force within the firm by some date). These policies often focus on new hires and don't emphasize retention. Evidence of this can be seen in the summary statistics in section 3.1. I explore this in the results section later. There are also national movements focused on boosting diversity in the work force, for example, The American Bar Foundation Resolution 113 was passed in August 2016 (Letts and Wolfe 2016). Indeed, the law profession is well aware of their trailing diversity numbers. Because of this, I assume that prestigious law firms practice some form of affirmative action. Of important note is the fact that all of the most prestigious firms employ a large number of lawyers (Nelson 1988). For example, the top 100 US law firms in Vault.com rankings each have a minimum of 250 lawyers (many in the thousands).[5] In this way, I can take a large firm to also be a prestigious firm.

[5]Vault.com is one of the most popular media sites to rank US law firms in many characteristics including diversity. A database for diversity numbers among law firms in the US can be found here: http://mcca.vault.com/searchresults.cfm. However, I do not formally analyze these numbers in this paper.

The next feature, while not critical for identification, is important for understanding outcomes for lawyers.

The second feature is law firm lockstep salary and promotion schemes. At most medium to large law firms in the US, associate attorneys with the same experience paid the same amount. The only time a lawyer will make more than her peers is if she is promoted to partner. Each law firm has a number of equity and non-equity partners. Larger firms have more partners than smaller firms. However, the ratio of partners to associates is far greater at small firms (e.g., 2:20 at a smaller firm and 20:750 at a larger firm). After 5-10 years at the firm, if an attorney has not been promoted to some form of partner, it is expected that she will politely resign and go work somewhere else. Because of this, it is possible that there are many lawyers who, after their first position, choose not to come back to law. This question of leaving law is addressed in the empirical section later.

3.2 The AJD

The data for this paper, the AJD study, is obtained by limited access from the American Bar Foundation (ABF) and contains detailed, non-identifying questionnaire information about 5,267 newly licensed U.S. lawyers. This longitudinal data set spans three waves, the first collected in 2000 and the last in 2012. Due to attrition, the final wave of the dataset contains 2,862 observations, about a third of which are usable in the final regressions of this paper. As such, I restrict this analysis to the first two waves. This should still provide sufficient variation in outcomes for lawyers as most lawyers will either be promoted or resign within the first 5-10 years of their time at a given firm. There are two main specifications in this paper, both of which use a different sub-sample of the first two waves. The first specification focuses on promotion rates for lawyers that are practicing in both waves. This sub-sample contains 2,199 observations, 715 of which are useful when running regressions. Similarly, the second specification is concerned with those lawyers who left law and contains lawyers that were private in the first wave and either private or not private in the second wave. This sub-sample contains 2,890 observations, 964 of which are useful for analysis. The comparison between the AJD sample of lawyers

in the US and the population sample from the American Bar Association (ABA) 1997 is done by Lehman (2011) who finds that the AJD sample matches up well with the population parameters found in the ABA survey. Because I am interested in outcomes for practicing lawyers and those who dropped out of law, I do not drop any observations except those due to attrition from the first to the second wave. Next, I will discuss some of the preliminary summary statistics for this data set.

3.3 Summary Statistics

Find summary statistics in Table 1 below which I have split into three sections, the whole workable sample, minority/non-minority sub samples, and large/non-large firm sub samples. Minority is considered to be non-white in race and a large firm is considered to employ greater than or equal to 250 lawyers. Each individual in the data set reports the number of lawyers that work at their current firm and I use this to determine whether they are at a large firm or not.[6] Note that I have data on the reported number of lawyers at each individual's firm. LS GPA has been re-coded so that the GPA reported here is on the standard 4.0 scale. LS US News Rank refers to the ranking category from 1 to 7 where 1 is a top 10 school, 2 is 11-20, 3 is 21-100, 4 is tier 3, 5 is tier 4, 6 is outside of the US, and 7 is unaccredited. Thus, lower values of LS US News Rank are better. LS class rank is categorical from 1 to 5 with smaller values being better. 1 is top 10%, 2 is top 11-25%, 3 is second quarter, 4 is third quarter, and 5 is the fourth quarter of the class. Law Review Leader indicates whether an individual was a leader in either a General Law Review or some other law review at their Law school. Law reviews are student run organizations in which students publish articles, notes, and other law commentaries. Log(salary) is just the log of annual salary at one's first position as a practicing lawyer.

Column 1 focuses on the sample as a whole. Note that 29% of the sample are in the minority category where Blacks, Asians, and Latinos have a similar share at 8-9% of the sample. Furthermore, there is about a 47% female share in my sample. LS GPA,

[6]The AJD contains each lawyer's past history in law practice. For example, if a lawyer switched jobs between their first job out of law school and the time of the survey, I can capture this by looking at their first job instead of the job at the current time of the first wave of the survey.

US News Rank, class rank, and law review leader are my measures of lawyer skill and I will use them to rank lawyers later.

The second and third columns split the sample into a minority and non-minority sub sample. Of primary concern are the differences in law school GPA, rank, and class rank. Non-minorities have better qualifications than minorities in the GPA and class rank category. In contrast, minorities, on average, go to better ranked law schools. This is understandable because many of the top public and private law programs in the US practiced some form of affirmative action during the time participants in this survey were attending law school (mid 1990s). For example, the whole state of California practiced affirmative action before the 1996 California Proposition 209 banning affirmative action in public institutions. Furthermore, 3 percentage points more non-minorities are law review leaders than minorities. Finally, note that minorities tend to have higher salaries and go more often to the larger (prestigious) firms.

The final two columns split the sample into a large and non-large sub sample. In line with Lehmann (2011) I consider a firm with more than 250 lawyers to be a large firm. The estimates in the remainder of this paper are robust to this definition (e.g., greater than 200 or greater than 300 lawyers both produce similar estimates). A large firm is considered to be more prestigious than a non-large firm as size is taken to be a proxy for prestige. Importantly, note the difference in the minority share at the large firm versus the non-large (.30 and .27 respectively). This is suggestive that preference is given to hiring minorities at large firms, noting that minorities are not, on average, better qualified than non-minorities as noted in the previous paragraph. Also notice that large firms pay a much higher salary than their smaller counterparts, in line with the assertion that large firms are more prestigious. With the data sufficiently explained, I discuss my identification strategy in the next section.

Table 1: Summary Statistics

	All	Sub-Sample			
		Minority	Non-Minority	Large Firm	Non-Large Firm
Minority	0.29	1.00	0.00	0.30	0.27
	(0.45)	(0.00)	(0.00)	(0.46)	(0.44)
Black	0.08	0.28	0.00	0.07	0.07
	(0.27)	(0.45)	(0.00)	(0.26)	(0.25)
Asian	0.09	0.31	0.00	0.11	0.08
	(0.28)	(0.46)	(0.00)	(0.31)	(0.26)
Native American	0.01	0.05	0.00	0.01	0.02
	(0.12)	(0.22)	(0.00)	(0.10)	(0.12)
Latino	0.08	0.28	0.00	0.08	0.08
	(0.27)	(0.45)	(0.00)	(0.27)	(0.27)
Other Race	0.04	0.12	0.01	0.03	0.04
	(0.19)	(0.33)	(0.07)	(0.18)	(0.21)
Large Firm	0.25	0.27	0.24	1.00	0.00
	(0.43)	(0.45)	(0.43)	(0.00)	(0.00)
Female	0.47	0.52	0.45	0.48	0.44
	(0.50)	(0.50)	(0.50)	(0.50)	(0.50)
Married	0.56	0.44	0.61	0.54	0.58
	(0.50)	(0.50)	(0.49)	(0.50)	(0.49)
LS GPA	3.26	3.16	3.29	3.46	3.23
	(0.34)	(0.35)	(0.34)	(0.28)	(0.34)
LS US News Rank	3.16	2.99	3.22	2.36	3.36
	(1.20)	(1.24)	(1.18)	(1.08)	(1.12)
LS Class Rank	2.54	2.91	2.44	1.98	2.64
	(1.11)	(1.16)	(1.07)	(0.90)	(1.11)
Law Review Leader	0.25	0.22	0.25	0.40	0.21
	(0.43)	(0.41)	(0.43)	(0.49)	(0.41)
Log(Salary)	11.17	11.19	11.16	11.68	11.10
	(0.56)	(0.62)	(0.53)	(0.32)	(0.53)
Tenure in Wave 2	5.03	4.64	5.18	4.65	5.00
	(3.20)	(2.94)	(3.30)	(2.71)	(3.04)
N	2890	826	2046	553	1650

4 Identification

4.1 Worker types and skill assignments.

For identification, I must sort lawyers into low and high type groups because the policy only affects a sub group of less qualified lawyers. To address this, I construct a skill characteristic for each lawyer, rank them all, and choose some cutoff of the distribution of skills. Lawyers below this cutoff are low type and those above are high type. I assume that the lawyers below this cutoff are at risk of being a beneficiary of affirmative action because they are more likely to have not gotten a job at the large firm in the absence of affirmative action (i.e. they are marginal). Each lawyer has a vector of qualifications:

$$q_i = \begin{bmatrix} \text{LS US News Rank} \\ \text{LS class rank} \\ \text{Law Review Leader} \\ \text{BAR passage state} \end{bmatrix}$$

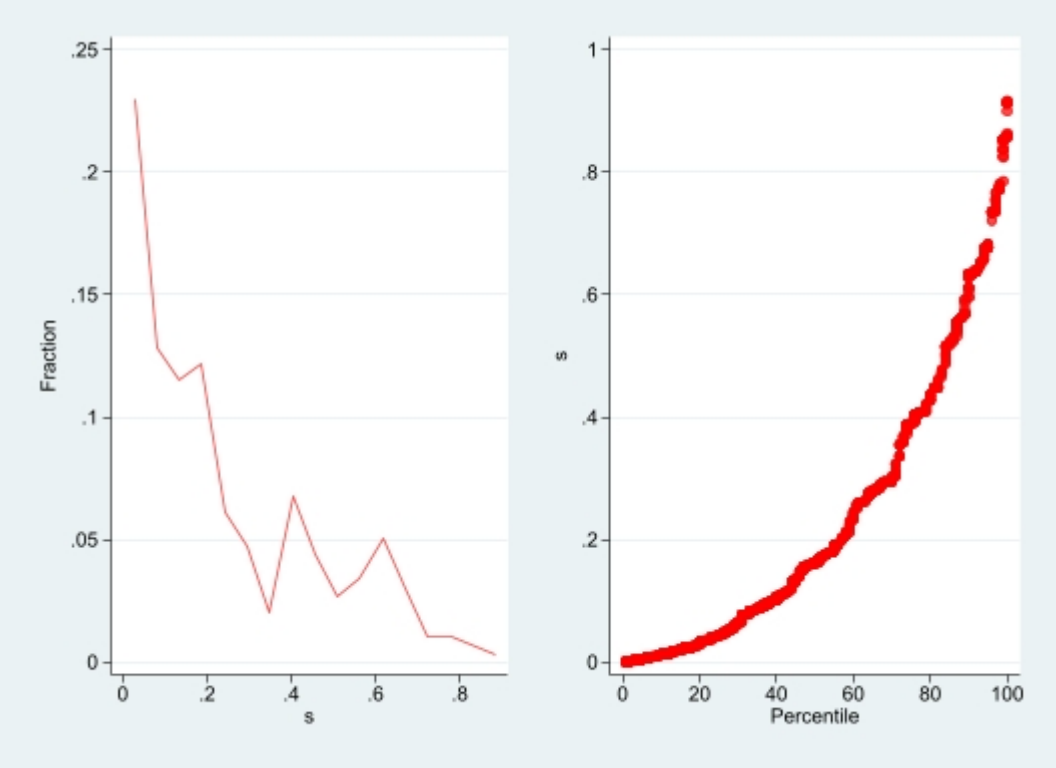
which can be used to rank lawyers according to their place in the distribution of these qualifications. [7] In order to compact these characteristics in a way that is conducive for ranking lawyers, I run the following probit and predict skill propensities for each lawyer based on their qualifications:

$$s_i = Pr(\text{Large}_i = 1 \mid q_i) = \Phi(\alpha_0 + \alpha_1 q_i) \tag{1}$$

Where α_1 is a vector of coefficients that estimate the effect of each of the characteristics in the qualification vector, q_i, on the likelihood of a lawyer working at large firm in their first position. Figure 1 below shows the distribution of s_i in the data as well as the percentiles and their associated s. Thus, I can rank all of the lawyers based on their s_i and choose some percentile cutoff and call the lawyers below that

[7] I include the BAR passage state in this list of characteristics because some states are known for having more rigorous exams, California and Delaware for example. I exclude LS GPA because, as noted in Ho 2005, GPA is not a good measure of law student quality within or outside of their school.

Figure 1: Skill (s) distribution and percentile break down



cutoff low type and those above high type. In my analysis, I allow this cutoff to vary because the percentile that contains the largest portion of beneficiaries is unknown to me. Thus, by exploring different cutoffs of the distribution, I obtain estimates that are robust to the definition of a low type lawyer. Next, I discuss the decision to use *Large* as the predictor for LS skills.

4.2 Law Firm Promotion to Partner

It is critical to note that the skill characteristic is increasing in one's qualifications for a large firm. Furthermore, promotions are far less likely to occur at large firms because there are fewer partners at a large firm relative to small firms. For this reason, level differences in promotion between large and non-large firms is not as informative as relative changes between groups, which is what I explore in this paper. Also it is important to distinguish between the extensive and intensive margin in this context. Some lawyers will drop out of law and to study them I use the full sample. It is possible to study who drops out and for what wage they switched, but I cannot study promotion rates for these people. However, for those who stay in law, I can study promotion patterns or whether they switch to a more or less prestigious firm. That is, I must be careful which group I use to study certain outcomes. With that I move on to the primary specification of the paper.

4.3 Baseline Dif in Dif

For the purposes of this paper, I assume that a large (prestigious) firm is the treatment: affirmative action. There are likely differences in outcomes for those at the large firm versus those at the non-large firm. Perhaps working at a large firm impacts one's outcomes positively or negatively whether that firm practices quasi affirmative action or not (e.g. Hoekstra 2009). If we employ a triple difference, we can eliminate this bias from the estimates. To arrive at such a model, I first write the difference in outcomes between the minority and non-minority group by type. To detect this, one might run the following difference in differences on a sub-sample of lawyers at large firms where affirmative action is practiced:

$$y_i = b_0 + b_1(Low_i) + b_2(Minority_i) + b_3(Low \times Minority_i) + e_i,$$
$$\text{and } e_i = u_i + \epsilon_i. \quad (2)$$

Low indicates that the lawyer's skill is below some percentile cutoff of the distribution of s and *Minority* indicates whether a lawyer is non-white in race. While unobserved heterogeneity between individuals is differenced out during this phase, there may

still be unobserved heterogeneity between the group, biasing the estimates. e_i can be broken into two parts. u_i is the unobserved difference in outcomes between the minority and non-minority groups by their type. ϵ_i is the independently distributed error term. For these estimates to be unbiased, $u_i = 0$. However, this is unlikely to be true if there are systematic difference in types by minority status that are not observed by the researcher. As such, a third difference must be used to eliminate unobserved heterogeneity between the minority and non-minority group.[8]

4.4 Triple Difference

With the new variable that indicates large firm, i.e., affirmative action, I can take a third difference from the specification in (2) and add in relevant controls to obtain the DDD used in this paper:

$$\begin{aligned} Outcome_{ir} = \beta_0 + \beta_1(Minority_i) + \beta_2(Low_i) + \beta_3(Large_i) \\ + \beta_4(Minority_i \times Low_i) + \beta_5(Minority_i \times Large_i) + \beta_6(Low_i \times Large_i) \\ + \beta_7(Minority_i \times Large_i \times low_i) + X_{ir} + \Gamma_r + \xi_i + \epsilon_{ir}. \quad (3) \end{aligned}$$

Minority and *Low* act as they did in (2) and *Large* indicates whether a lawyer's firm employs at least 250 lawyers. X_{ir} is a vector of controls: mother and father's education, marital status, and gender and Γ_r are regional fixed effects of which there are 26. [9] Importantly, I have differenced out part of u_i from before and am left with ξ_i which is the unobserved heterogeneity between low and high type minority workers who chose to work at a large firm. ϵ_{ir} is still the independently distributed error term, orthogonal to the covariates in (3). DDDs are uncommon and so I provide some insight as to what my particular specification does.

Suppose I run (2) for two mutually exclusive groups of the sample, large firm (l) and

[8] I acknowledge that this is a point of debate as noted in Lang and Lehmann 2012.

[9] Note that many of the individual's characteristics are captured by Low.

non-large firm (n) lawyers:

$$\begin{aligned} y_i^l &= b_0^l + b_1^l(Low_i) + b_2^l(Minority) + b_3^l(Low \times Minority_i) + u_i^l + \epsilon_i \\ y_i^n &= b_0^n + b_1^n(Low_i) + b_2^n(Minority) + b_3^n(Low \times Minority_i) + u_i^n + \epsilon_i \\ &\text{where } u_i^h - u_i^l = \xi_i \end{aligned} \tag{4}$$

I obtain estimates of b_3 for both groups, b_3^h and b_3^l. At the mean, the difference between the two above equations in (4) will give (3). With respect to our coefficient of interest, this difference gives $\beta_7 = b_3^h - b_3^l$. It tells us if there is a difference in outcomes for low type minorities at the large firm and low type minorities at the non-large firm. Regarding non-minorities, this tells us whether local mismatch is happening more for minorities than it is for non-minorities. It is quite easy to visualize this using sample means and no controls. Please see Figure 2 for a graphical representation of this effect where I define *Low* as being below the 50^{th} percentile of the skill distribution.

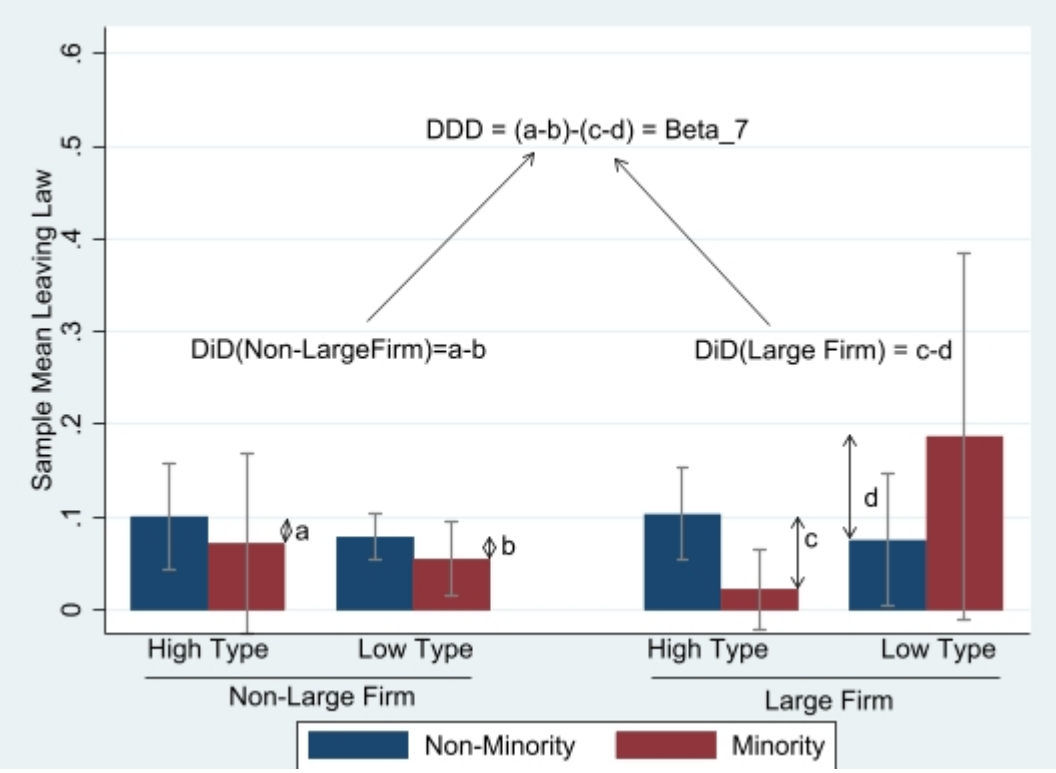
The consequences of this figure will be discussed in more detail later but for now, note the left side of the figure (above "Non-Large Firm"). we see there is a gap in outcomes for minorities and non-minorities of both the high and low type at the non-large firm. This gap does not really differ depending on whether a lawyer is high or low type. However, this difference is strikingly different on the right side of the figure, for large firms. We see a large jump in leaving rates for low type minority lawyers compared to their non-minority counterparts. My specification analyzes this gap with proper controls and the elimination of unobserved differences in the treatment and control groups. With the specification explained, I will address the assumptions necessary for this analysis to produce unbiased estimates of the impact of affirmative action in the local sense.

4.5 Identification Assumptions

As I suggest previously, identification of local mismatch due to affirmative action (the mismatch hypothesis) requires three assumptions:

1. Large firms are taken to be a proxy for prestigious firms. Prestigious firms are incentivized and do practice quasi-affirmative action in order to improve

Figure 2: Triple difference of Leaving Law



diversity scores and, consequently, prestige ranking.

2. Only low type minorities who choose to work at a large firm can be affected by the treatment, affirmative action.

3. There is no difference in the distribution of skills by minority status. As such, a low type minority will make the same decisions as a low type non-minority.

I draw these assumptions from the previous sections where I discuss features of the data, law practice and law firm structure, as well as the econometric features of a DDD. Number 3 is the critical identification assumption because it biases the estimates in an unknown direction. That is, we do not know what direction ξ_i in (3) would shift the mean outcome for the different types of workers. With the assumptions thoroughly explained and the data established, I now move into a discussion of my empirical findings.

5 Empirical Results

In the following section, all outcome variables are observed at the beginning of the second wave of the study and independent variables are observed either at or before the first wave of the study. Importantly, *Large* is determined by each lawyer's first job and not their current job in the first period. A small portion of lawyers in the data set switched jobs between passing the bar in 1998 at the earliest and participating in the survey in 2000. Fortunately, the survey collects data on previous positions so I can easily find key characteristics of these individuals' first jobs. Next, I describe the outcome variables of interest in detail. Then I state and justify the definition of *Low*. The final three subsections report my results.

5.1 Outcomes

In this paper, I study three outcomes on the intensive margin and one on the extensive margin, promotion and future salary, and leaving law, respectively.[10] *Promotion* is

[10]Lawyers on the intensive margin are those who are practicing lawyers in both waves of the AJD. Those on the extensive margin are lawyers who dropped out of law practice between the first and

one of the most popular measures of success among lawyers and it asks whether a lawyer went from a non-partner position at their first job to a partner position in Wave 2. *Future Salary* tells something about the prestige of the firm because prestigious firms pay more. I define *Future Salary* to be the log of salary at a lawyer's job in Wave 2. Recall, *Large* is a decision and wage growth is an outcome of that decision. Thus, wage growth asks the question, do certain lawyers who go to the prestigious firms end up downgrading to less prestigious firms or do they continue climbing the prestige ladder? The final outcome on the intensive margin is **Tenure** which is the number of years a lawyer has been at their current job at the time of the second wave (2007). On the extensive margin, I study leaving rates. That is, who are those lawyers that go from practicing law in the first wave of the study to not practicing law in the second wave of the study? If there is significant leaving my estimates for promotion and wage growth for any groups that are leaving will likely be biased towards zero. I extrapolate more on this topic in the following sub sections.

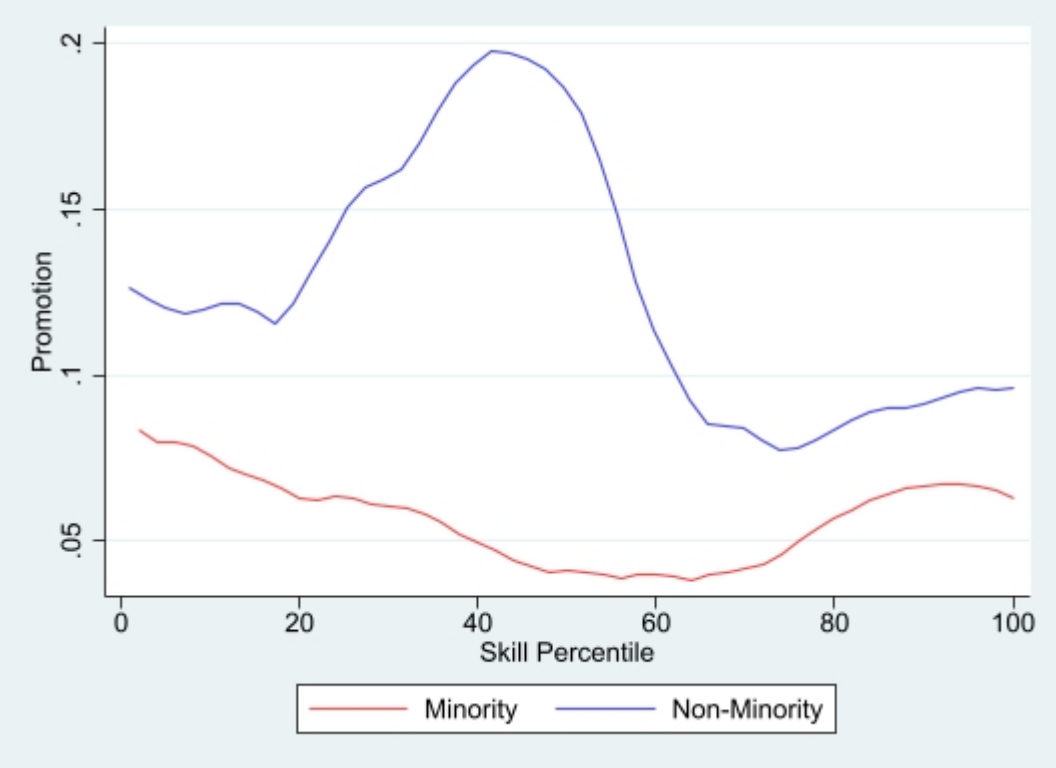
5.2 *Low* Definition

Up until now, I have not explicitly defined *Low*. In an effort to be both concise and robust, I choose a few definitions for *Low* for the following sections. To that end, I plot promotion rates over the skill percentile by minority status. See Figure 3 below for this plot with respect to promotion rates among practicing lawyers.[11] There is a jump in the mean promotion rates for non-minorities around the 40 to 50$^{\text{th}}$ percentile of skills. My regressions will focus on this part of the distribution and *Low* will be defined according to these percentiles of the distribution.

the second wave of the study. I leave out lawyers who did not practice law in the first wave in this study.

[11] I use `lpoly` in STATA which plots a weighted regression of the promotion indicator on skill percentile. I use the default bandwidth which is a rule of thumb method of obtaining optimal bandwidth

Figure 3: Probability of Promotion by Skill Percentile



5.3 Promotion Rates

In Table 2 below I present estimates of (3) with promotion rates as the dependent variable. Note that this section is limited to the sub-sample of lawyers that were practicing in both the first and second wave of the AJD. Each column reports estimates for a different percentile cutoff of the skill distribution within the range suggested by Figure 2. For example, in the first column, *Low* indicates someone in the bottom 40^{th} percentile of the distribution (and so on). The top row of the table presents β_7 from my main specification, the difference in the impact of affirmative action on the low-type minority lawyers and the low-type non-minority lawyers. The mismatch hypothesis suggests that this coefficient ought to be negative (affirmative action hurts beneficiaries). Interestingly, I note a marginal effect of somewhere between -12.1% and -6.3%. While these estimates are economically significant and somewhat stable across the different definitions of *Low*, they are not statistically significantly different

from zero. In other words, the estimates are consistent with the mismatch hypothesis but, due to large confidence intervals, we cannot say whether the mismatch hypothesis holds here. Promotion is not the only metric of success for a lawyer though. For example, an individual might not become promoted but they might switch to working at a more prestigious firm. This is measured by the change in salary from the first job to that in Wave 2 which I explore next.

Table 2: Triple DiD of Promotion Rates

	(1) ≤ 40 %ile	(2) ≤ 45 %ile	(3) ≤ 50 %ile	(4) ≤ 55 %ile	(5) ≤ 60 %ile
Large=1	-0.075*	-0.066	-0.049	-0.043	-0.059
	(0.041)	(0.041)	(0.042)	(0.042)	(0.045)
Minority=1	-0.098**	-0.080*	-0.083*	-0.049	-0.046
	(0.044)	(0.045)	(0.047)	(0.051)	(0.059)
Large=1×Minority=1	0.080	0.065	0.070	0.043	0.052
	(0.064)	(0.065)	(0.067)	(0.069)	(0.076)
Low=1	-0.037	-0.012	0.032	0.058	0.017
	(0.046)	(0.045)	(0.046)	(0.046)	(0.047)
Large=1 × Low=1	0.022	0.012	0.004	0.037	0.034
	(0.118)	(0.121)	(0.111)	(0.111)	(0.099)
Minority=1 × Low=1	0.071	0.033	0.034	-0.022	-0.024
	(0.075)	(0.072)	(0.070)	(0.069)	(0.073)
Large=1 × Minority=1 × Low=1	-0.098	-0.079	-0.108	-0.109	-0.137
	(0.148)	(0.149)	(0.137)	(0.138)	(0.133)
Female	-0.092***	-0.091***	-0.091***	-0.090***	-0.090***
	(0.029)	(0.029)	(0.029)	(0.029)	(0.029)
Married	0.020	0.022	0.024	0.026	0.024
	(0.032)	(0.032)	(0.032)	(0.032)	(0.032)
Constant	0.195**	0.183**	0.165**	0.152*	0.168**
	(0.080)	(0.080)	(0.082)	(0.082)	(0.085)
Parent Education Controls	Yes	Yes	Yes	Yes	Yes
Region FE	Yes	Yes	Yes	Yes	Yes
Observations	803	803	803	803	803
Adjusted R^2	0.062	0.061	0.062	0.065	0.062

NOTE– The dependent variable, Promotion, indicates that a lawyer went from a non-partner position in the first wave to a partner in the second wave (equity or non-equity partner). The sample is restricted to private and practicing lawyers in the second wave of the AJD. Minority indicates non-white race. Large indicates greater than or equal to 250 lawyers at the current firm. Low indicates at or below the percentile indicated at the top of each column. Percentiles are given by the distribution of lawyer quality where quality is derived from a probit of Large on important covariates: LS GPA, LS Class Rank, LS US News Rank, Bar passage state, and age. GPA is categorical from 1 to 8 (1:Under 2.25, 2: 2.25-2.49, 3: 2.50-2.74, 4: 2.75-2.99, 5: 3.00-3.24, 6: 3.25-3.49, 7: 3.50-3.74, 8: 3.75-4.00). Rank refers to the US News Law school Ranking and is categorical from 1 to 7 (1: Unaccredited, 2: Outside US, 3: Tier 4, 4: Tier 3, 5: Top 21-100, 6: Top 11-20, 7: Top 10).
*** p<0.01, ** p<0.05, * p<0.1 Robust standard errors clustered on the individual level are in parentheses.
† Low is defined as being at or below the percentile as indicated by the column headers.

5.4 Future Salary

Recall that the sample in this section and the previous section are restricted to practicing lawyers in both the first and second wave of the AJD. Using the same definitions of *Low*, I present Table 3 below which is (3) with the log of Wave 2 salary as the dependent variable. Although the estimates of β_7 appear striking at first glance, ranging between -0.715 and 0.306, they are highly insignificant and so interpreting them as the truth would be inappropriate. Of particular concern is the switch from positive to negative coefficients between the 40^{th} and 45^{th} percentiles. These results suggest that there is a massive racial gap in future earnings for low type lawyers at large firms. This could suggest that beneficiaries are quitting their jobs at large firms and going to smaller firms, thus making a lower salary. Indeed, this is the only possible answer if lawyers are staying within the law profession because one cannot get a pay decrease at the same firm and a worse firm will have a lower salary. Due to the inconsistency and size of these coefficients, one may wonder about those workers on the extensive margin. If the mismatch hypothesis is true in the local sense, it could be the case that mismatched minority workers are so much worse off by the policy that they dropped out of the profession entirely. The next section explores this possibility.[12]

[12]Please see the table in the appendix for a naive regression of Log(Salary) on minority×leftLaw with controls. This is a naive check to see if there are any racial differences in the outcomes for those who dropped out of law.

Table 3: Triple DiD of Wave 3 Log Salary

	(1) ≤ 40 %ile	(2) ≤ 45 %ile	(3) ≤ 50 %ile	(4) ≤ 55 %ile	(5) ≤ 60 %ile
Large=1	-0.151**	-0.130**	-0.146**	-0.089	-0.103
	(0.062)	(0.066)	(0.069)	(0.068)	(0.077)
Minority=1	-0.151	-0.171	-0.188	-0.177	-0.144
	(0.100)	(0.107)	(0.123)	(0.132)	(0.149)
Large=1×Minority=1	0.079	0.212*	0.230*	0.216	0.173
	(0.153)	(0.119)	(0.134)	(0.142)	(0.160)
Low=1	0.040	0.057	0.019	0.125	0.095
	(0.082)	(0.082)	(0.088)	(0.077)	(0.084)
Large=1 × Low=1	0.118	0.003	0.023	-0.066	-0.033
	(0.120)	(0.128)	(0.115)	(0.107)	(0.110)
Minority=1 × Low=1	-0.043	0.007	0.031	0.003	-0.040
	(0.124)	(0.128)	(0.140)	(0.144)	(0.159)
Large=1 × Minority=1 × Low=1	0.047	-1.133	-1.146	-1.126	-0.561
	(0.287)	(0.732)	(0.731)	(0.733)	(0.612)
Female	-0.109**	-0.103**	-0.105**	-0.100**	-0.105**
	(0.048)	(0.047)	(0.047)	(0.046)	(0.046)
Married	-0.043	-0.048	-0.050	-0.046	-0.044
	(0.044)	(0.044)	(0.042)	(0.044)	(0.044)
Constant	0.438***	0.412***	0.428***	0.372***	0.382***
	(0.129)	(0.127)	(0.126)	(0.113)	(0.120)
Parent Education Controls	Yes	Yes	Yes	Yes	Yes
Region FE	Yes	Yes	Yes	Yes	Yes
Observations	709	709	709	709	709
Adjusted R^2	0.055	0.075	0.073	0.082	0.066

NOTE– The dependent variable, Log Wage in wave 2, is the log of salary of the lawyer in the second wave of the story. The sample is restricted to practicing lawyers in the second wave of the AJD. Minority indicates non-white race. Large indicates greater than or equal to 250 lawyers at the current firm. Low indicates at or below the percentile indicated at the top of each column. Percentiles are given by the distribution of lawyer quality where quality is derived from a probit of Large on important covariates: LS GPA, LS Class Rank, LS US News Rank, Bar passage state, and age. GPA is categorical from 1 to 8 (1:Under 2.25, 2: 2.25-2.49, 3: 2.50-2.74, 4: 2.75-2.99, 5: 3.00-3.24, 6: 3.25-3.49, 7: 3.50-3.74, 8: 3.75-4.00). Rank refers to the US News Law school Ranking and is categorical from 1 to 7 (1: Unaccredited, 2: Outside US, 3: Tier 4, 4: Tier 3, 5: Top 21-100, 6: Top 11-20, 7: Top 10).
*** p<0.01, ** p<0.05, * p<0.1 Robust standard errors clustered on the individual level are in parentheses.
† Low is defined as being at or below the percentile as indicated by the column headers.

5.5 Leaving Law

Up until now, I have ignored the possibility that low types are attriting out of the law labor force due to poor matches with firms, eliminating variation in outcomes. This attrition could cause the estimates of future salary and promotion to be weaker and even biased. To explore the question of whether this is the case I present Figure 3 below which is a local polynomial regression of the probability of leaving law practice on skill percentiles by minority status. Notice the spike in minority-leavers near the 60^{th} percentile. This spike is surprising and unexpected in the absence of affirmative action. However, such a spike is conceivable under some discriminatory policy. Minorities that go to the large firms because of the policy may be so poorly matched that they quit law altogether. To explore this econometrically, I present Table 4 below which explores the impact of local mismatch on leaving law for the entire workable sample of lawyers (both practicing and non-practicing in the second wave). I find massive positive and somewhat statistically significant values for β_7. That is, the racial gap in leaving rates between minorities and non-minorities is much larger at the large firm than at the small firms: anywhere from 20 to 57.1% according to my analysis. Focusing only on the bottom half of the distribution, the first 3 columns of Table 4, we see evidence to suggest that some extreme local mismatching is happening in the data.

Notice also the negative impact of being a low type lawyer at a large firm on leaving. That is, columns 1 and 2 of Table 4 suggest that a low type lawyer at a large firm is about 14% less likely to leave law than their high type counterparts. These estimates suggest that reverse mismatching may be occurring. More specifically, the positive spillover of the large firm is so strong that it dominates any negative impacts from being mismatched. An alternative explanation is that the very best lawyers are dropping out and starting their own businesses or going on to be professors at prestigious institutions. However, I do not explore this question in my paper.

Table 4: Triple DiD of leaving

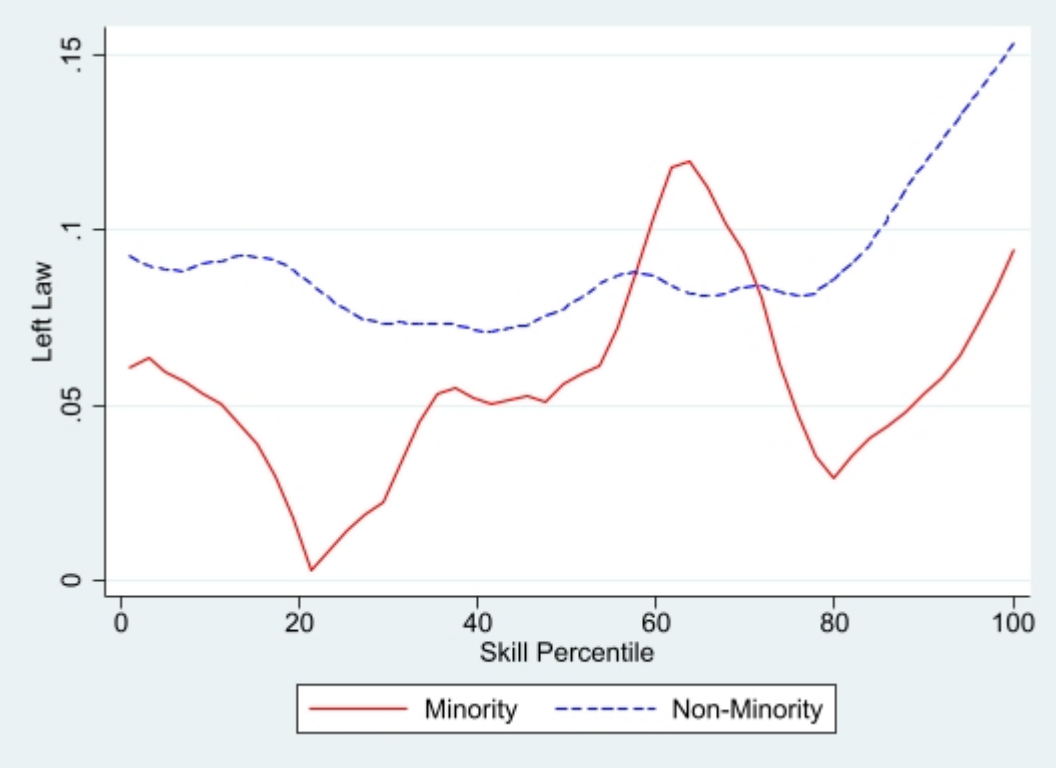
	(1) ≤ 40 %ile	(2) ≤ 45 %ile	(3) ≤ 50 %ile	(4) ≤ 55 %ile	(5) ≤ 60 %ile
Large=1	0.032	0.034	0.023	0.025	0.004
	(0.034)	(0.035)	(0.034)	(0.036)	(0.038)
Minority=1	0.025	0.008	0.016	0.021	0.034
	(0.045)	(0.044)	(0.050)	(0.062)	(0.072)
Large=1×Minority=1	-0.096	-0.082	-0.082	-0.089	-0.092
	(0.061)	(0.061)	(0.066)	(0.075)	(0.085)
Low=1	0.006	-0.001	0.001	-0.001	-0.024
	(0.028)	(0.025)	(0.025)	(0.026)	(0.028)
Large=1 × Low=1	-0.147***	-0.143***	-0.032	-0.044	0.021
	(0.054)	(0.047)	(0.093)	(0.079)	(0.076)
Minority=1 × Low=1	-0.130**	-0.087	-0.090	-0.086	-0.096
	(0.052)	(0.053)	(0.056)	(0.065)	(0.074)
Large=1 × Minority=1 × Low=1	0.595**	0.447**	0.324	0.333	0.216
	(0.256)	(0.216)	(0.220)	(0.218)	(0.188)
Female	0.045*	0.044*	0.043*	0.043*	0.043*
	(0.023)	(0.024)	(0.024)	(0.024)	(0.024)
Married	0.013	0.012	0.013	0.012	0.011
	(0.022)	(0.022)	(0.022)	(0.022)	(0.022)
Constant	0.081	0.083	0.086	0.086	0.102*
	(0.056)	(0.054)	(0.055)	(0.055)	(0.057)
Parent Education Controls	Yes	Yes	Yes	Yes	Yes
Region FE	Yes	Yes	Yes	Yes	Yes
Observations	954	954	954	954	954
Adjusted R^2	0.034	0.031	0.027	0.027	0.027

NOTE– The dependent variable, leaving law, indicates whether a lawyer went from practicing law in the first wave to not practicing in the second wave. The sample is restricted to only lawyers who were practicing at private firms in the first wave of the study. Minority indicates non-white race. Large indicates greater than or equal to 250 lawyers at the current firm. Low indicates at or below the percentile indicated at the top of each column. Percentiles are given by the distribution of lawyer quality where quality is derived from a probit of Large on important covariates: LS GPA, LS Class Rank, LS US News Rank, Bar passage state, and age. GPA is categorical from 1 to 8 (1:Under 2.25, 2: 2.25-2.49, 3: 2.50-2.74, 4: 2.75-2.99, 5: 3.00-3.24, 6: 3.25-3.49, 7: 3.50-3.74, 8: 3.75-4.00). Rank refers to the US News Law school Ranking and is categorical from 1 to 7 (1: Unaccredited, 2: Outside US, 3: Tier 4, 4: Tier 3, 5: Top 21-100, 6: Top 11-20, 7: Top 10).

*** p<0.01, ** p<0.05, * p<0.1 Robust standard errors clustered on the individual level are in parentheses.

† Low is defined as being at or below the percentile as indicated by the column headers.

Figure 4: Probability of leaving law by skill Percentile

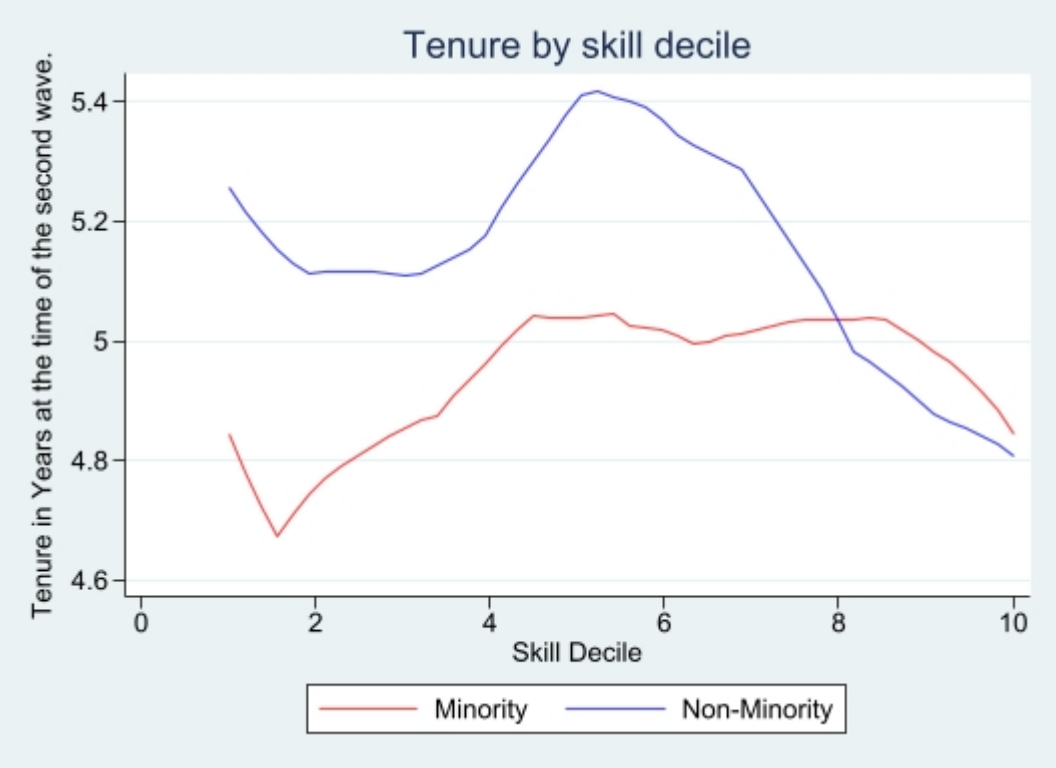


5.6 Tenure

It stands to reason that firms practicing quota diversity practices such as the ones displayed on many prestigious law firms' websites, lead to a gap in retention rates for minorities versus whites. For further exploration, I present Figure 5 below which shows the smooth polynomial average of tenure by skill decile. One immediately notices a racial gap at the lower deciles and convergence at the higher ones. This observation supports Cavounidis, Lang, and Weinstein 2019 in which minorities are churned more than their non-minority counterparts because they are more heavily monitored by superiors. Although it is possible that the monitoring decisions of supervisors cause these disparities, I claim that the quota-based diversity practice is what leads to this higher churning of the minority group than the non-minority group.

Table 5 below presents estimates of my primary specification with *Tenure* as the

Figure 5: Tenure by skill decile



dependent variable. Across all definitions of *Low*, we see that there is a negative yet statistically insignificant affect of the diversity quota on the beneficiary group (indicated by the bolded triple interaction term). Note also the near statistical significance and large gap in tenure for minorities, suggesting that, after controlling for race and skill interaction, there is still a racial gap in retention.

The results in this sub section are perhaps the most worrisome of those reported in this paper. Law firms may be implementing a quota policy focused on increasing diversity (a good thing) without a focus on retention (a bad thing), thus hurting the group they set out to help. From a pure policy standpoint, law firms should take special care when implementing their diversity hiring decisions and ensure that they have the resources to follow up on these hires for future success. Indeed many of the top law firms state a focus on lawyer retention and even partner goals. However, these data suggest they may not be succeeding in these goals.

While it is entirely possible that minorities are too qualified and are getting better job offers elsewhere, given the results in the previous sub sections, this is unlikely to be the case. Indeed this result is most striking as it suggests that there are racial gaps in outcomes even at very high levels of education. More research is needed as this could be an artifact of law as an occupation.

Table 5: Triple DiD of Tenure at current firm in second wave

	(1) ≤ 40 %ile	(2) ≤ 45 %ile	(3) ≤ 50 %ile	(4) ≤ 55 %ile	(5) ≤ 60 %ile
Large=1	0.056	0.058	-0.054	0.081	0.220
	(0.315)	(0.327)	(0.338)	(0.355)	(0.374)
Minority=1	-0.346	-0.362	-0.756*	-0.521	-0.489
	(0.390)	(0.411)	(0.417)	(0.488)	(0.533)
Large=1×Minority=1	0.109	0.020	0.471	0.186	-0.054
	(0.787)	(0.811)	(0.815)	(0.862)	(0.902)
Low=1	-0.248	-0.328	-0.357	-0.199	-0.027
	(0.354)	(0.346)	(0.357)	(0.362)	(0.382)
Large=1 × Low=1	0.275	-0.106	0.520	0.003	-0.376
	(0.878)	(0.737)	(0.729)	(0.657)	(0.625)
Minority=1 × Low=1	0.917	0.874	1.403	0.913	0.788
	(1.017)	(0.961)	(0.885)	(0.878)	(0.865)
Large=1 × Minority=1 × Low=1	-1.152	0.103	-0.999	-0.138	0.972
	(1.910)	(1.649)	(1.600)	(1.568)	(1.482)
Female	-0.118	-0.125	-0.130	-0.119	-0.112
	(0.246)	(0.247)	(0.245)	(0.245)	(0.244)
Married	0.546**	0.551**	0.544**	0.568**	0.560**
	(0.247)	(0.249)	(0.247)	(0.250)	(0.250)
Constant	4.652***	4.681***	4.736***	4.660***	4.569***
	(0.665)	(0.666)	(0.667)	(0.672)	(0.671)
Parent Education Controls	Yes	Yes	Yes	Yes	Yes
Region FE	Yes	Yes	Yes	Yes	Yes
Observations	820	820	820	820	820
Adjusted R^2	0.014	0.015	0.017	0.013	0.014

NOTE– The dependent variable, Tenure, is the number of years the lawyer has been at their current employer at the time of the second wave of the AJD. The sample is restricted to private and practicing lawyers in the second wave of the AJD. Minority indicates non-white race. Large indicates greater than or equal to 250 lawyers at the current firm. Low indicates at or below the percentile indicated at the top of each column. Percentiles are given by the distribution of lawyer quality where quality is derived from a probit of Large on important covariates: LS GPA, LS Class Rank, LS US News Rank, Bar passage state, and age. GPA is categorical from 1 to 8 (1:Under 2.25, 2: 2.25-2.49, 3: 2.50-2.74, 4: 2.75-2.99, 5: 3.00-3.24, 6: 3.25-3.49, 7: 3.50-3.74, 8: 3.75-4.00). Rank refers to the US News Law school Ranking and is categorical from 1 to 7 (1: Unaccredited, 2: Outside US, 3: Tier 4, 4: Tier 3, 5: Top 21-100, 6: Top 11-20, 7: Top 10).

*** p<0.01, ** p<0.05, * p<0.1 Robust standard errors clustered on the individual level are in parentheses.

† Low is defined as being at or below the percentile as indicated by the column headers.

6 Conclusion

The mismatch hypothesis states that minorities impacted by affirmative action may be worse off at a prestigious institution (law firm in this paper) due to a poor match. I explore this question in the context of the labor market, extending previous literature which focused on the educational market. To that end, I couple a novel econometric specification with a unique data set on newly licensed lawyers in the US in order to explore the mismatch hypothesis in the labor market. I use a DDD to identify the difference in the racial gap in outcomes for low type minorities and low type non-minorities at a large versus non-large firm. To do so, I utilize the structure of the US law labor market as well as unique measures for lawyer skill. By assigning a skill measure to each lawyer based on their observables, I am able to weakly identify beneficiaries who are lower skilled minorities and more likely to be impacted by the diversity quota. Because prestigious law firms are also very large, I can use the number of lawyers at a firm as a proxy for prestige. Furthermore, prestigious firms are incentivized to practice some form of affirmative action in order to improve their minority share as this boosts their place in the national rankings of law firms (many firms openly report these quotas on their websites).

Initially, I find that while racial differences in promotion rates and future salary are economically significant, they are not statistically significant, providing little evidence that diversity quotas negatively impact workers in the real world. However, upon closer inspection, I find that low-type minorities are exceptionally more likely (42.5 to 57.1%) to leave law if they work at a large firm than if they work at a non-large firm. This suggests an extreme form of mismatching in which lawyers who are impacted by the diversity quotas at the prestigious firm select out of law practice entirely. As such, these lawyers do not show up in the analysis for promotion rates and salary. Interestingly, I find that low types at large firms do tend to have better outcomes than if they were at the small firm. As mentioned previously, this benefit is greatly reduced for minorities.

Furthermore, it appears that minorities are churned more intensely than non-minorities, with about 25% shorter job tenure than their white counterparts, and this can be at-

tributed to the quota-based diversity practices at prestigious law firms. This coupled with the fact that the beneficiaries are leaving law at a dramatically greater pace than their counterparts gives a result very similar to the theoretical predictions in Cavounidis, Lang, and Weinstein 2019.

The warm glow for a low-type worker employed at a prestigious firm is inexplicably absent for minorities yet staunchly present for non-minorities. Affirmative action seems to treat the symptoms and not the underlying problem. Especially at large law firms in the US, special care should be taken for those lawyers who are hired to boost diversity rankings. Emphasis must be put on retention and not just hiring. Many law firms in the US have begun to focus on this in the last few years thanks to movements like the American Bar Association's Resolution 113. I suspect that my findings extend to other occupations but law practice is unique and allows for such an analysis for the reasons pertaining to law firm structure outlined previously. One might question whether affirmative action is the cause of the dramatic leaving among low-type minorities at large firms or this is attributed to some other economic impact. Even if this were the case, this paper presents startling stylized facts and suggests that some form of racial discrimination is occurring. These findings contrast the assertion in labor economics that racial gaps often diminish as education increases (Black et al. 2006 and Bjerk 2007). This paper also adds to the scarce research on racial gaps in outcomes as a function of skill as noted in Lang and Lehmann (2012). I find a convergence of outcomes for high skilled lawyers but an obvious gap for those with lower skills. Future research might marry the AJD with the skill mismatch literature, allowing for theoretical extension concerning racial discrimination. Furthermore, this framework may be well suited to address other racial discrimination topics yet unanswered.

7 Appendix Table 1

The following table explores whether or not lawyers who leave law switch to higher paying jobs by race. While all of the effects are positive, they aren't statistically significant so it is very hard to say whether those who leave law actually do make

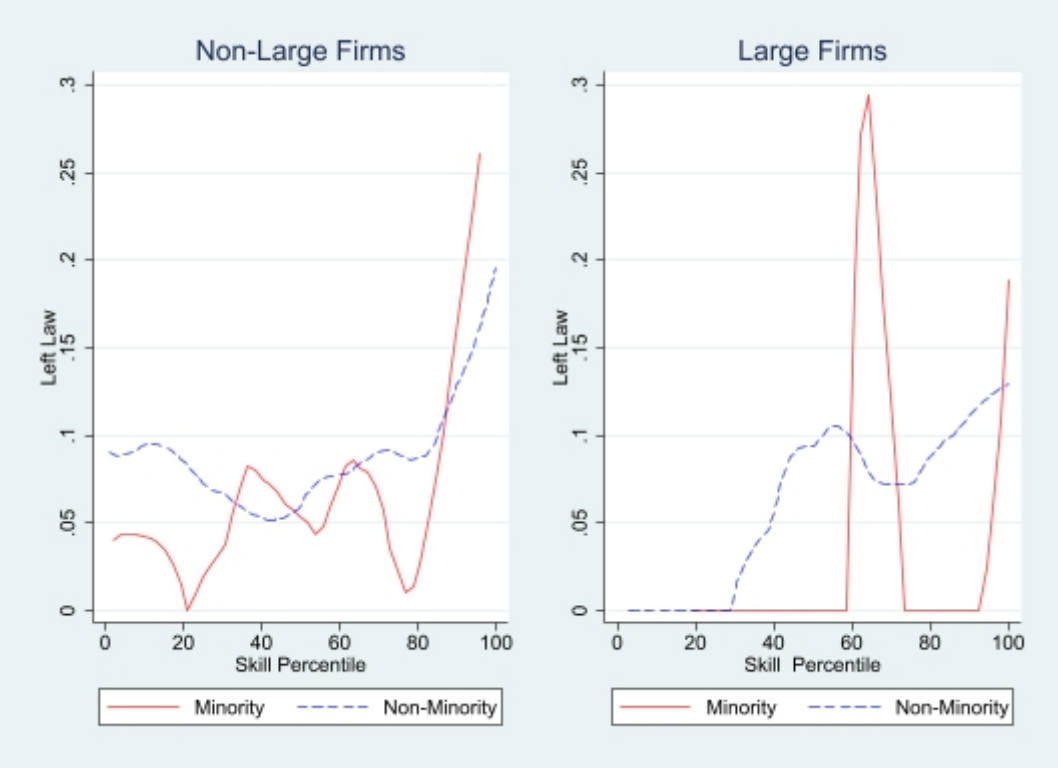
more than those who stay in law, all else equal. While not critical to the paper itself, this table tells us whether there is anything striking happening for those minorities who drop out of law compared to the non-minorities.

	Log(Salary)
Left Law	0.0377
	(0.45)
Minority	0.0164
	(0.35)
Left Law×Minority	0.0784
	(0.52)
Skills	0.924***
	(11.22)
Parents' Education	X
Regional FE	X
Constant	11.08***
	(129.42)
N	1180

t statistics in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Figure A1



8 Appendix Figure 1

One might worry that the graph in Figure 4 doesn't tell the full story. I present Figure A1 below which splits Figure 4 into two groups, large and non-large firms. We see on the left side that the minority and non-minority lpolys are very similar, consistent with the fact that the two groups should be equivalent in skills. However, on the right we see a massive inconsistency and spike in the minority lpoly, similar to the one in Figure 4. This suggest that the spike in Figure 4 can be attributed to minorities at the large firm. Though, due to low observation count, it is difficult to say that this is not an anomaly of the data set.

References

[1] Fryer, Ronald. 2007. Belief flipping in a dynamic model of statistical discrimination. *Journal of Public Economics*. 91(June):1151-1166.

[2] Coate, Stephen and Glenn Loury 1993. Will Affirmative Action Policies Eliminate Negative Stereotypes? *The American Economic Journal*. 83(5): 1220-1240.

[3] Moro, Andrea and Peter Norman. 2004. A general equilibrium model of statistical discrimination. *Journal of Economic Theory*. 114(1): 1-30.

[4] Arrow, Kenneth. 1973. The theory of discrimination. *Princeton University Press.*

[5] Phelps, Edmund 1972. The statistical theory of racism and sexism. *The American Economic Review*. 62: 659-661.

[6] Arcidiacono, Peter. 2005. Affirmative action in higher education: How do admission and financial aid rules affect future earnings? *Econometrica*. 73:1477-1524.

[7] Arcidiacono, Peter, Esteban Aucejo, Patrick Coate, and V Joseph Hotz. (2014)

[8] Arcidiacono, Peter, Michael Lovenheim. 2016. Affirmative action and the quality-fit tradeoff. *Journal of Economic Literature*. 54(1):3-51.

[9] Arcidiacono, Peter, Estaban Aucejo, Hanming Fang, Kenneth Spenner. 2011. Does affirmative action lead to mismatch? A new test and evidence. *Quantitative Economics*. 2:303-333.

[10] Black, Dan, Amelia Haviland, Seth Sanders, and Lowell Taylor. 2006. "Why do minority men earn less? A study of wage differentials among the highly educated." *Review of Economics and Statistics* 88(2): 300-313.

[11] Bjerk, David. 2007. "The differing nature of black-white wage inequality across occupational sectors." *Journal of Human Resources* 42(2):398-434.

[12] Sander, Richard 2004. A systematic analysis of affirmative action in American Law schools. *Stanford Law Review*. 57(2):367-483.

[13] Rothstein, Jesse and Albert Yoon. 2009. Mismatch in law school. Working Papers 79, Princeton University, Department of Economics, Center for Economic Policy Studies.

[14] Ayres, Ian. and Richard Brooks. 2005. Does affirmative action reduce the number of black lawyers. *Faculty Scholarship Series.* Paper 1231.

[15] Williams, Douglass. 2013. Do racial preference affect minority learning in law schools? *Journal of Empirical Legal Studies.* 10(2):171-195.

[16] Ho, Daniel. 2005. Why affirmative action does not cause Black students to fail the bar. *Yale Law Journal.* 114(8):1997-2004.

[17] Sander, Richard. 2005. A reply to the critics. *Stanford Law Review.* 57(6): 1963-2016.

[18] Lehmann, Jee-Yeon. 2015. Job assignment and promotion under statistical discrimination: evidence from the early careers of lawyers. MPRA Paper 33466, University Library of Munich, Germany.

[19] Letts, Eileen M. and David B. Wolfe. 2016. Letter. American Bar Association. *American Bar Association Diversity Inclusion 360 Commission Report to the House of Delegates: Resolution 113.* https://www.americanbar.org/content/dam/aba/administrative/racial˙ethnic˙diversity/aba/letter˙from˙the˙signatories˙general˙counsel˙implementation˙of˙aba˙resolution˙113.pdf

[20] Abowde, J. F. Karmarz, and D. Margolis. 2003. High wage workers and high wage firms. *Econometrica.* 67(2):251-333

[21] Korpi T. and Tåhlin M. (2009) ‘ Educational Mismatch, Wages, and Wage Growth: Overeducation in Sweden, 1974–2000 ’, Labour Economics 16 (2): 183 – 193.

[22] Bauer T. K. (2002) ‘ Educational Mismatch and Wages: A Panel Analysis ’, Economics of Education Review 21 (3): 221 – 229.

[23] Allen J. and Van der Velden R. (2001) ‘ Educational Mismatches versus Skill Mismatches: Effects on Wages, Job satisfaction, and on-the-Job Search ’, Oxford Economic Papers 53 (3): 434 – 452.

[24] Hoekstra, M. 2009. The Effect of Attending the Flagship State University on Earnings: A Discontinuity-Based Approach. *Review of Economics and Statistics*. 91(4): 717-724.

[25] Nelson, R. 1988. Partners with Power: The social Transformation of the Large law Firm. *University of California Press*. Book

[26] Lang, Kevin and Jee-Yeon. 2012. Racial Discrimination in the Labor Market: Theory and Empirics. *Journal of Economic Literature*. 50(4): 959-1006.

[27] Cavounidis, Costas, Kevin Lang, and Russel Weinstein. 2019. The Boss is Watching: How Monitoring Decisions Hurt Black Workers. *NBER Working Paper No. 26319*