The Incentive Effects of Affirmative Action in a Real-Effort Tournament*

Caterina Calsamiglia[+] U. Autònoma de Barcelona

Jörg Franke[1] TU Dortmund

Pedro Rey-Biel[**] U. Autònoma de Barcelona

First draft: June 17, 2009

This version: February 22, 2010

Abstract

Affirmative-action policies bias tournament rules in order to provide equal opportunities to a group of competitors who have a disadvantage they cannot be held responsible for. Critics argue that they distort incentives, resulting in lower individual performance, and that the selected pool of tournament winners may be inefficient. In this paper, we study the empirical validity of such claims in a real-effort pair-wise tournament between children from two similar schools who systematically differ in how much training they received ex-ante in the task at hand. Our results show that when affirmative action measures were implemented performance was not reduced for either advantaged or disadvantaged subjects. Additionally, while affirmative action balanced the proportion of disadvantaged individuals winning their respective tournament, the average performance of the pool of winners only decreased slightly.

Keywords: Affirmative action, tournament, real-effort, experiment, sudoku.

JEL classification: C72; C91; J78; M52

---

* We are especially grateful to Antoni Calvó-Armengol for his support and encouragement on this project. We thank Jose Apesteguia, Miguel A. Ballester, Carmen Beviá, Jeanette Brosig, Antonio Cabrales, Marta García-Matos, Maureen Gleeson, Uri Gneezy, Guillaume Haeringer, Nagore Iriberri, Inés Macho-Stadler, Muriel Niederle, Carmit Segal, Tom Palfrey, Neslihan Uler and seminar audiences at many seminars and conferences. We are deeply grateful to the directors, faculty and administrative personnel at Agora, Aula, Colegio Alemán, Emili Juncadella, Pegaso and Sagrado Corazón for their understanding, permission and help with the experiments. We thank Pau Balart, Julen Berasaluce, Ignacio Fernández, Markus Kinateder, Tomasso Majer, Luca Merlino and Natalia Montinari for their help in running the experiments. Caterina Calsamiglia acknowledges financial support by a Ramón y Cajal contract of the Spanish Ministerio de Ciencia y Tecnología, through the Spanish Plan Nacional I+D+I (SEJ2005-01481, SEJ2005-01690 and FEDER), and through the "Grupo Consolidado de tipo C" (ECO2008-04756), the Generalitat de Catalunya (SGR2005-00626 and the Barcelona Economics Program of XREA), and the Consolider-Ingenio 2010 (CSD2006-00016) program. Jörg Franke acknowledges financial support from Ministerio de Educación y Ciencia (SEJ2005-01481/ECON) and FEDER, from Generalitat de Catalunya (2005SGR00454) and from Universitat Autònoma de Barcelona. Pedro Rey-Biel acknowledges financial support from Ministerio de Ciencia y Tecnología (ECON2009-07616 and Consolider-Ingenio CSD2006-00016), Barcelona GSE Research Network and of the Government of Catalonia (2009SGR-169).

+ Caterina Calsamiglia. Universitat Autònoma de Barcelona. Department d´Economia i d´Historia Econòmica. 08193, Bellaterra. Barcelona (Spain). E-mail: caterina.calsamiglia@uab.cat.

1 Jörg Franke, Department of Economics and Social Science, TU Dortmund, 44227 Dortmund, Germany. Tel.: (+49) 231 755 3246. E-mail: joerg.franke@tu-dortmund.de.

** Pedro Rey-Biel. Universitat Autònoma de Barcelona. Department d´Economia i d´Historia Econòmica. 08193, Bellaterra. Barcelona (Spain). Tel: (+34) 935812113. E-mail: pedro.rey@uab.cat.

1. Introduction

In many selection processes such as university admissions, job promotions and procurement auctions, competition helps identifying the highest-ability candidates, facilitate efficient allocation of talent, and provide incentives for individuals to improve their skills. This objective may not be achieved if some otherwise competitive candidates do not stand a fair chance to win the competition. For example, talented students from poor economic backgrounds may have attended high schools that receive less funding, which may affect their SAT performance and hence their university admission. Likewise, some individuals may belong to groups which historically have suffered discrimination, and have to overcome major obstacles in order to be on an equal footing to compete.

Affirmative Action policies (AA) have two main objectives: to guarantee that positions are fairly allocated in society and to allow for the creation and identification of talent. AA policies take proactive steps to provide equal opportunities to discriminated groups that have a potential disadvantage.[2] They are often implemented by biasing tournament rules in order to increase the probability of success of an otherwise disadvantaged group. For example, a fixed lump-sum bonus of 20 (out of 150) points was added to the score of minority applicants to the undergraduate program at the University of Michigan and a similar but "unofficial lift" scheme is used at many top universities.[3] In a different domain, public procurement auctions, bid preferences are granted in a multiplicative way. For example, road construction contracts in California are auctioned off by granting a 5% reduction of the submitted bid to small business enterprises.

The implementation of AA is usually accompanied by intense public debate focusing on whether such policies satisfy certain fairness criteria and on the possible incentive distortions they may create. Abstracting from fairness considerations, opponents of AA base their criticism

[2] Merriam-Webster Online defines affirmative action as "an active effort to promote the rights or progress of minority groups or other disadvantaged persons".

[3] This procedure was recently ruled to be unconstitutional by the Supreme Court, due in part to alleged distortionary effects on incentives that such compensation may create. State funded universities such as California, Florida and Texas have also applied similar policies in the past. A number of papers have analyzed the effects of such banning on the efficiency of race-blind policies that result from colleges preferences for diversity. See, among others, Chan and Eyster (2003), Fryer and Loury (forthcoming). The idea behind these papers is that banning affirmative action forces colleges with a taste for diversity to target minorities indirectly through randomly admitting students with relatively lower test scores or other characteristics correlated with race, which leads to important inefficiencies.

on two grounds. First, it is argued that disadvantaged individuals may see affirmative action policies as substitute for their own effort, lowering their performance, while advantaged individuals may be discouraged by the perceived unfairness of such policies.[4] Second, opponents argue that the pool of selected individuals, that is, the pool of winners, may be of poorer quality since lower-performing individuals may now be selected.[5] On the other hand, advocates of AA argue that leveling the playing field in a competitive environment may have positive effects on performance because AA reduces the asymmetry in capacities to compete, which increases competitive pressure and therefore enhances performance.

Both positions fail to base their views on solid empirical evidence since good data is scarce and very difficult to obtain. In this paper, we present results from a pair-wise real-effort tournament in which there exists a naturally induced source of disadvantage for one group of competitors, and where two different types of AA policies, lump sum and proportional bonuses, are implemented to compensate for it.

We designed pair-wise tournaments among children from two similar schools which differ in how experienced their students are in solving simple numerical puzzles known as "sudokus". Students in one school ("experienced") are taught how to solve sudokus as part of their regular math classes, while students in the other school ("non-experienced") are not.[6] The schools are very similar in all other relevant respects: both are private, located in the same upper-middle class neighborhood, are fully bilingual and have good records in national math and science competitions. Therefore, the difference in experience can be regarded as an exogenous source of disadvantage imposed on two similar groups, since sudoku practice in the experienced school was not in use at the time most parents made their schooling decision.

As expected, experience in the task provided an advantage in the competition, as verified by a comparison of the performance of subjects in the two schools in a baseline treatment in which they are not informed of the previous experience in the rival school and in which no affirmative action is implemented. We first study whether knowing that such an asymmetry in experience exists affects the performance of both experienced and non-experienced individuals. To do so we

[4] See, for example, the introductory remarks in Sowell (2004) and the discussion in Fryer and Loury (2005b) of "Myth No. 3: Affirmative action undercuts investment incentives".
[5] Loury and Garman (1993) and Kane (1998) provide mixed evidence on the possible mismatch that AA may generate on the allocation of students to institutions. Lott (2000) finds that AA in the recruitment of police officers has increased diversity but also crime rates mainly because of the decrease in standards of recruited individuals.
[6] Coate and Loury (1993) show how discrimination may arise in two symmetric groups as a self-fulfilling prophecy. We take such asymmetry as given.

compare results in our baseline treatment with another treatment where the difference in experience is made salient (and still no affirmative action is implemented). Finally, in treatments where subjects are aware of the asymmetry in experience, we implemented two types of compensation—lump-sum and proportional bonuses—designed to equalize the probability of non-experienced students beating their experienced rivals.[7] We then study how performance by students from both schools is affected by the implementation of AA and whether the average output of the new pool of tournament winners differs from the one obtained without any form of compensation.

Schotter and Weigelt (1992) studied the incentive effects of AA in a pioneering laboratory experiment where effort exertion in a tournament is modeled as an individual decision problem based on monetary costs. Subjects' exogenous disadvantage was induced by assigning different cost parameters for which individuals were later compensated by affirmative action. This procedure makes it possible to vary the size of the asymmetry, tailor the compensations to exactly level the playing field and compare it with cases in which compensations are higher (or lower) than the initial asymmetry. Their results indicate that AA can either boost or worsen performance depending on the sizes of the cost disadvantage and the compensation implemented. Our study complements their work by analyzing a real-effort tournament where the asymmetry between subjects exists ex-ante and is not induced by the experimentalist, and hence where compensation seems natural. This comes at a cost of not having an exact ex-ante measure of the size of the asymmetry, which forced us to rely on results from pilot experiments to roughly calculate the appropriate size of compensations which on average leveled the playing field.[8]

Subjects in our experiment were school children and the experiment was conducted in their respective schools. They were unaware that their choices were the object of a study since the experiment was presented as an extra-curricular activity of a type not uncommon in the schools we selected. Using children as subjects has additional advantages: they react very spontaneously in competitive situations; their performance is not affected by them questioning the underlying motivation of the experimentalist; and it is relatively easy to provide them with incentives. It has

[7] Calsamiglia (2009) shows that an appropriately designed AA policy should equalize rewards to effort whenever the set-up affects one of many factors determining individual final welfare. In this particular environment rewards to effort are equalized with proportional AA.

[8] Several recent experimental studies employ similar strategies based on naturally occurring differences in characteristics among social groups. Examples are Hoff and Pandey (2006), where social caste differences are exploited, as well as Gneezy et al. (2003), and Niederle and Vesterlund (2007), where the performance of women versus men and their respective propensity to compete in mixed-gender tournaments is analyzed.

also been shown that children react rationally and in line with economic theory (see Harbaugh et al. (2001) and Harbaugh and Krause (2000)). Finally, studying how children react to affirmative action is important since some social asymmetries may be ideally resolved at these early ages, before they are exacerbated.

The implementation of affirmative action policies aims to reduce the underlying heterogeneity between competitors and induce a more leveled playing field. There are a number of theoretical papers analyzing related policies, where either strong players are handicapped, or weak players favored. For instance, Lazear and Rosen (1981), show that a handicapping system induces efficient competition in a rank-order tournament between heterogeneous players. Che and Gale (1998), analyze bid caps for strong bidders in an all-pay auction framework. Also Myerson (1981) shows that an optimal, i.e. revenue maximizing auction, between asymmetric bidders implies favoring weak bidders. Theoretical papers that explicitly analyze affirmative action policies are Schotter and Weigelt (1992), Fu (2006), Franke (2008) and Balart (2009). They model affirmative action as a bias in favor of ex-ante disadvantaged players in an all-pay auction or contest set-up. The conclusion that can be drawn from most of these papers is that reducing the asymmetry in competitive advantage tends to enhance individual performance. However, whether the implementation of affirmative action policies in real applications is suitable to level the playing field is a rather open empirical question.

While there is a large empirical literature on tournaments (see Prendergast (1999) for a survey) and also on affirmative action (see Holzer and Neumark (2000) for a survey), the incentive effects of affirmative action policies remains rather unexplored, given the difficulty of obtaining good quality data. Some recent exceptions are Bertrand, et al. (forthcoming), which analyzes the effects of AA on Indian minorities, Miller and Segal (2008), which studies the long-term effects of affirmative action on the pool of hired law enforcement officers in the US, and some studies on bid preferences in public procurement auctions such as Krasnokutskaya and Seim (2007) and Marion (2007). Experimental research addressing similar issues include Niederle et al. (2008), where the effects of quotas on tournament participation of women are examined, as well as Freeman and Gelber (2010), where the effects of different prize structures and information about competitors' ability in a tournament framework are analyzed.

The experimental results of our study suggest that the implementation of AA policies does not necessarily have an adverse effect on the performance of affected individuals. First, we confirm that the asymmetry in experience is reflected in subjects' performance, since we observe that non-experienced subjects solve significantly less sudokus than experienced ones. Next, for non-experienced individuals we find that AA enhances their performance independently of their ability. For experienced individuals performance effects differ with ability. For those subjects with relatively low or average ability, performance is enhanced while for those with highest ability performance worsens. Our AA policies balanced the tournament, on average, since around half of non-experienced subjects in treatments where AA was implemented won their respective tournaments in all possible matches. Also, the average performance of all possible tournament winners selected through AA was moderately lower than the average performance of the winners who would have been selected without it, since the increased levels of efforts performed under AA partially compensates for the effect of selecting as tournament winners a higher proportion of non-experienced individuals. We also find that AA positively increases the confidence in winning of non-experienced subjects, while that of experienced subjects is unaffected. Finally, all subjects regard the initial asymmetry in experience as unfair. However, fairness perceptions vary with treatments. For instance, proportional AA, whose size depends on individual performance, is perceived as fairer than lump sum AA, although the actual compensation received in the proportional treatments was higher.

The rest of the paper is organized as follows. Experimental design and procedures are explained in Section 2. Section 3 presents the results. Section 4 sums up our conclusions. The Appendix contains an English translation of the instructions used in the experiment.

2. Experimental Design and Procedures

We conducted pair-wise tournaments among 336 school children, aged 10-13, from two similar non-religious, bilingual private schools located in the same upper-class neighborhood of Barcelona. Students at both schools have a systematic difference in experience of a specific real-effort task consisting in solving simple "sudokus". This ex-ante difference in experience is due to the fact that during regular math classes, students in the "experienced" school (E) are trained in solving sudokus (and in fact have to solve sudokus as part of their regular homework) while students at the "non-experienced" school (NE) are not.[9]

Sudoku is a logic-based number-placement puzzle. The objective is to fill a 9x9 grid so that each column, each row and each of the nine 3x3 boxes contains one-digit numbers from 1 to 9 only once. The puzzle setter provides a partially completed grid. We use a simplified 4x4 grid version in order to obtain sufficient variability in performance. We chose this task because the rules are simple, yet it requires substantive logical reasoning and concentration by the subjects. Additionally, performance is easy to measure and, crucially, depends on effort. Most importantly, both effort and ability play a role, so that non-experienced subjects still have a chance of winning, independently of whether they are favored by an affirmative action policy or not.[10] Figure 1 below shows one of the sudokus used in the experiment (a) and its solution (b).

	4		2
		3	
1			

(a) Unsolved Sudoku

3	4	1	2
2	1	3	4
1	2	4	3
4	3	2	1

(b) Correctly Solved Sudoku

Figure 1: An example of the real-effort task (sudoku).

Each student from E was randomly and anonymously matched with a student from NE in his or her same school year (4th or 6th grade). Each pair competed in a tournament which lasted

[9] An ex-post experimental questionnaire showed that some students from both schools were familiar with sudokus due to prior experience. We control for this ex-ante experience in our analysis by using a proxy for ability. Results from pilot experiments, in addition to the present one, show that subjects from NE were in fact disadvantaged in the competition (see section 3.1). The task was defined as "filling in a grid" and the word "sudoku" was never mentioned.

[10] In fact, the percentage of NE winners in their respective tournament was at least 13.3% (for experimental treatment "K" and 4th year students, where no affirmative action was implemented).

30 minutes.[11] Subjects had to correctly solve as many sudokus as possible in order to beat their matched rival. All subjects were handed the same answer sheet containing 96 sudokus randomly generated with the same level of difficulty by a computer program.[12] Each pair of subjects was competing for a 7€ (euro) voucher from a bookshop located in Barcelona.[13] In each pair, the student who had correctly solved more sudokus during a 30 minute period won the voucher. In the case of ties, the winner was decided randomly.

Our objective was to study: 1) the effect of providing information on competitors' previous experience with the task and 2) the effect of implementing affirmative action policies on subjects' performance and as a result, on the output generated by subjects selected as tournament winners. Thus, we randomly assigned similar numbers of subjects from each school to each of six treatments. In treatment NK no subject was informed about whether subjects from the other school were experienced or not in solving sudokus. In treatment K students at the NE school were told that students in the E school had previous experience in solving sudokus. Similarly, students in the E school were told that students at the NE school were not taught how to solve sudokus. In the remaining four treatments all subjects were informed about the existence of a difference in experience across schools and about the particular affirmative action policy applied to compensate NE subjects. In treatments LH (Lump-sum High) and LL (Lump-sum Low), all subjects knew that NE subjects were given a predetermined number of solved sudokus ex-ante: 20 in LH and 8 in LL. In treatments PH (Proportional High) and PL (Proportional Low), all subjects knew that NE subjects were given a number of solved sudokus proportional to the number of sudokus they correctly solved, one for every correctly solved one in the case of PH, and one for every two correctly solved ones in the case of PL.

Comparisons across treatments NK and K allow us to study the effects of information when no affirmative action policies are implemented. In parts of our analysis, we pool the data from all treatments where affirmative action is implemented (LL, LH, PL, PH) and refer to them generically as the "AA" treatment. Table 1 summarizes our treatment design. Comparisons

[11] We chose pair-wise tournaments versus multiple-prize tournaments with N player because the schools did not want us to establish intra-school competitions. Additionally, it allowed us to control the amount of information that each subject had on its rival's ability.
[12] The software used was "SuDoku Pro" by Dualogy Systems. The proportion of mistakes across all solved sudokus was similar. No subject was able to complete all 96 sudokus provided.
[13] Subjects were explicitly told that the voucher was redeemable for "books, collector's cards, toys, music or comics". Experiments took place at approximately the time the final *Harry Potter* book was published in Spain.

across treatments K and AA allow us to study the effect on performance of applying affirmative action policies once the asymmetry in experience is known.

Table 1: Description of Treatments		
Not Know	NK	Subjects unaware of others' experience
Know	K	Subjects aware of others' experience
Lump-sum High	LH	Subjects aware of experience and NE subjects receive a bonus of 20 correct sudokus bonus
Lump-sum Low	LL	Subjects aware of experience and NE subjects receive a bonus of 8 correct sudokus bonus
Proportional High	PH	Subjects aware of experience and NE subjects receive 1 correct sudoku bonus for every 1 correct
Proportional Low	PL	Subjects aware of experience and NE subjects receive 1 correct sudoku bonus for every 2 correct

The sizes of the affirmative action policies were determined using results from pilot experiments in similar schools. Since we were unaware of the exact size of the asymmetry between groups, we opted to choose two different sizes of each AA policy so that we could potentially observe how different sizes affect performance. At the same time, the objective was to design compensations that would on average equalize the chances of winning once the subjects reacted to the implemented affirmative action policies. As shown in section 3.4, the implemented policies roughly induced a "leveled playing field" ex-post on average (51% of subjects winning any possible pair-wise match were from the NE in the AA treatments).

Prior to conducting the experiments, we repeatedly met with faculty from both schools in order to guarantee their collaboration and pedagogical interest in the project. During these meetings we obtained information on subjects' gender, birth date, teaching group and school grades. We later assigned subjects to treatments in such a way that the groups were balanced in accordance with these pre-specified characteristics. Table 2 below shows descriptive statistics of subjects assigned to each treatment at each school.[14] Small variations across treatments were mainly due to absent students and latecomers.

Table 2: Description of the Subject Pool												
N = 336	Experienced						Non-experienced					
	NK	K	LH	LL	PH	PL	NK	K	LH	LL	PH	PL
% Female	41	43	48	53	50	48	46	46	48	39	48	47
% 6^{th} Year	48	43	45	43	46	48	50	46	59	48	48	59

[14] Average Grade is calculated using grades in all topics in the preceding term and it is slightly higher at the NE school than at the E school (3.55 vs. 3.44, significant at the 1% level). This difference is due to different grading systems across schools.

Average Grade (1=Worst,5=Best)	3.32	3 .31	3.47	3.46	3.14	3.35	3.44	3.44	3.65	3.65	3.54	3.57
Number of subjects	29	30	31	30	28	31	24	24	27	23	27	32

Subjects were unaware of their participation in an experiment. With the help of each school's faculty, subjects were told that this was an extracurricular activity, not dissimilar to previous ones carried out during the same school year. Participation was quasi-mandatory, which helped to avoid selection biases and simplified matters for the school. None of the subjects manifested opposition to participating.

Experiments were carried out on two separate but close dates in 2008. In each school experimental sessions took place at different times of the day for 4th and 6th graders for practical reasons.[15] Subjects were conducted to separate classrooms according to our predefined assignment. While students waited for the experimentalist, teachers conducted a specific and identical school activity (writing an essay) in order to keep the subjects calmed and equally uninformed about the experiment. The same experimentalist arrived at each of the classrooms at twenty-minute intervals and then sessions started.[16] Teachers were not present during the experimental sessions, in order to minimize their influence.

The experimental sessions lasted one hour. First, the experimentalist read out general instructions on how to solve sudokus (see "Pre-instructions" in the Appendix). Then, subjects had a five-minute practice round to solve sudokus with no incentives offered for performance and no mention of any competition. After this period, the experimentalist solved one of the practice sudokus in front of the students. Once questions were clarified, instructions for each of the treatments were read aloud. The instructions made it clear that each student was competing against an anonymous student from another comparable school and that students at the other school were systematically experienced (or not) in solving sudokus (for treatment NK this information was omitted). The difference in ex-ante experience was explicitly mentioned and was used to justify the implementation of the affirmative action bias in favour of the non-experienced group in the AA treatments (see the Appendix for the instructions). Tournament rules were explained giving numerical examples (specific to each treatment) for all potential

[15] We are unaware of cross-contamination between schools or between subjects from different school years at the same school. The timing of the experiments was carefully designed so as to avoid these problems.

[16] This was the reason different treatments were carried out at different time-intervals. Since the experiment deals with effort motivation and children may be easily influenced, it was crucial to have the same experimentalist conducting the sessions. The experimentalist rehearsed repeating exactly the same cues across sessions.

outcomes of the tournament, i.e., losing, winning, and tying. Moreover, aggregate information with respect to the number of sudokus (i.e., mean, minimum and maximum) that had been correctly solved by a comparable subject pool was provided. This information, identical for all subjects, was based on the results of our pilot experiments. The experimentalist also held up a 7€ voucher to increase the credibility of the prize offered to tournament winners. After that, subjects had thirty minutes in which to solve the sudokus in two separate handouts. After the first fifteen minutes, subjects were instructed to start working on the second handout, so that we could measure whether there were intra-session learning effects or whether these were over-ruled by fatigue.[17] Subjects were explicitly told that they could stop solving sudokus at any time and start any other activity, such as drawing, so long as they kept quiet and did not bother others.

After the thirty minutes had passed, the handouts were collected and a questionnaire about previous experience in solving sudokus, self-confidence and the perceived justice behind affirmative action policies was distributed. Once the questionnaires had been filled in, subjects continued with their regular classes. The experimentalists then randomly matched participants from both schools, determined the winners and deposited the vouchers at the schools, so that they could be distributed by school faculty.

3. Results

3.1 Descriptive Statistics

We start by taking a descriptive look at the data. Table 3 reports the average number of correct sudokus by treatment and school year (4^{th} or 6^{th} grade) in each of the schools (E and NE), as well as standard deviations. There is high heterogeneity in performance in all treatments and thus, standard deviations are large. Table 3 provides a first indication that subjects from the experienced school (E) solve, on average, more sudokus than subjects in the non-experienced school, a key hypothesis justifying our experimental design. Independently of treatment, subjects in the experienced school solved on average 35.88 sudokus, while subjects in the NE school solved 23.31. Using the number of sudokus solved in the five minute practice round as a measure

[17] We did not find substantive differences in performance between the two parts of the test, indicating that the effects of learning and fatigue possibly cancel out. Experienced subjects completed on average one more sudoku in the second part than in the first part (significant at the 1% level). Non-experienced subjects did not solve a significantly different number of sudokus in the two parts.

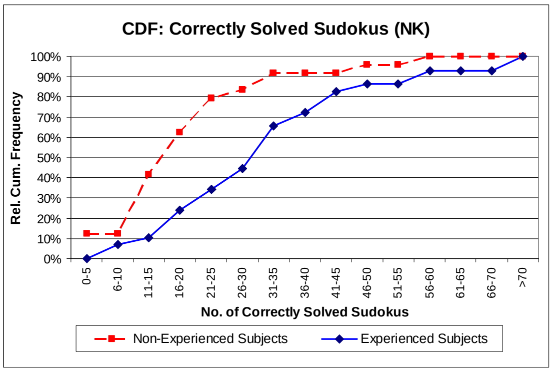
of individual ability, we find that experienced subjects of low ability, those who solve three sudokus in the practice round, solve a similar number of sudokus (25.61).in the tournament as the average of NE subjects

Table 3 also shows that age affected performance. The average performance of 4^{th} grade experienced subjects in all treatments is similar to that of 6^{th} grade non-experienced subjects.

Table 3: Average Number (and Standard Deviations) of Correct Sudokus by Treatment and School Year

	4^{th} Grade		6^{th} Grade		Overall	
	E	NE	E	NE	E	NE
NK	**28** (15.43)	**16.98** (8.01))	**38.93** (16.10)	**24.67** (15.44)	**33.27** (16.44)	**20.38** (12.80)
K	**29.88** (12.47)	**17.69** (10.70)	**43** (17.98)	**29.09** (13.43)	**35.57** (16.22)	**22.92** (13.13)
AA	**29.38** (13.78)	**19.26** (9.48)	**45.67** (12.04)	**28.08** (12.12)	**36.58** (15.53)	**24.04** (11.80)
LH	27.59 (12.26)	23.36 (9.19)	44.86 (11.51)	29.50 (14.43)	35.39 (15.02)	27 (12.73)
LL	27.59 (12.26)	19.42 (11.79)	51.54 (11.44)	26 (9.01)	37.97 (16.82)	22.56 (10.85)
PH	29.67 (12.38)	17.92 (9.05)	46.92 (11.09)	26.54 (11.16)	37.68 (14.52)	22.07 (10.84)
PL	32.94 (17.36)	17.07 (7.59)	40.27 (12.38)	29.16 (12.83)	36.48 (15.37)	24.25 (12.42)

Figure 2 below shows the cumulative distribution function (CDF) of the number of correct sudokus solved by students in the E and NE school for the two treatments where affirmative action policies are not implemented (NK and K). Note that the distributions have a large spread and range from 0 sudokus solved to more than 70. Stochastic dominance of the CDFs for the E school clearly shows that the lack of experience in solving sudokus is in fact a disadvantage for the NE subjects. Mann-Whitney tests comparing the inter-school number of correct sudokus in both of these treatments show significant differences at the 1% level (p-values of 0.002 for NK and of 0.004 for K).



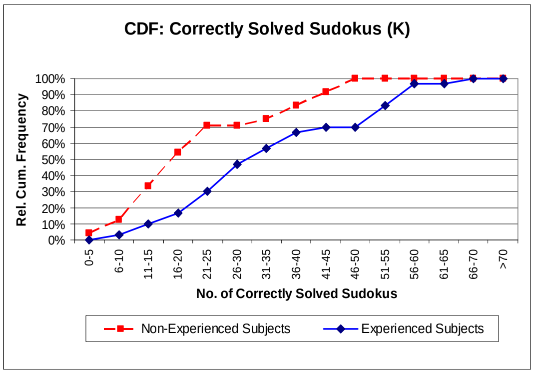


Figure 2: CDFs of the number of correct sudokus by E and NE in the NK and K treatments.

Intra-school comparisons across treatments are less clear-cut. Table 3 reports the number of correctly solved sudokus both for each of the unpooled data treatments (LH, LL, PH and PL), and for the pooled treatment (AA). Although standard deviations are very large, averages give a first indication that performance may be enhanced when providing information (K vs. NK treatments) and that affirmative action policies also enhanced performance (AA vs. K treatments, with the only negative comparisons being for E subjects in 4th grade and NE subjects in 6th grade). Figure 3 depicts the CDFs for the number of correct sudokus for the NK, K and AA treatments in each of the two schools. Visually, the CDF for the K treatment "almost stochastically dominates" the CDF for the NK treatment in both graphs, suggesting that the provision of information on the existence of a disadvantaged group does not decrease performance. Similarly, the CDFs for the AA treatment also lie mostly below the CDFs for the K treatment in both schools, suggesting that subjects faced with AA policies do not decrease their performance. Comparing the distributions of all treatments based on pair-wise Mann-Whitney tests does generally not result in significant differences at the standard levels, apart from the comparison of NK with LL for 6th year experienced subjects (average of 38.92 correct sudokus in NK and average of 51.54 in LL, p-value of 0.01). Additionally, a joint Kruskall-Wallis test for all treatments where affirmative action is implemented (LH, LL, PH and PL) does not show significant differences.

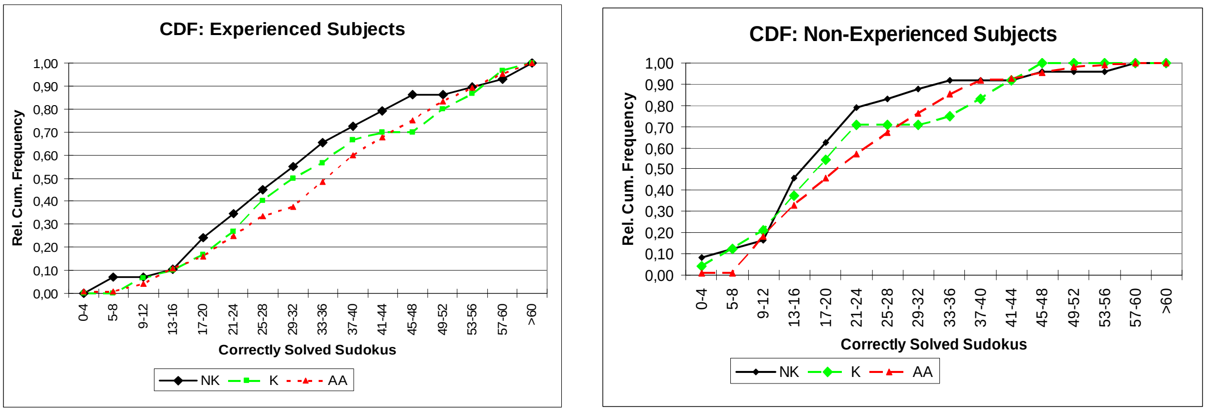


Figure 3: CDFs of the number of correct sudokus by school in the NK, K and AA treatments.

The fact that statistical tests are mostly non-significant can be attributed to large heterogeneity among subjects, which cannot be controlled for by using non-parametric methods. Therefore, in the next section we rely on regression analysis as it allows us to incorporate substantial additional information on the subject level to control for unobserved individual heterogeneity.

3.2 The Effects of Information on Performance

In this subsection we start by studying whether having information on the previous experience of rivals affected performance. In order to do so, we compare the number of correct sudokus in the K and NK treatments. Figure 3 indicates that performance in the K treatment is not lower than under the NK treatment. Although the CDF of performance in the K treatment in fact lies below the one of the NK treatment, both distributions are not significantly different. We here run OLS regressions with robust errors on performance, controlling for individual ability, age and gender.[18] Table 4 below shows linear regressions on the number of correct sudokus both for experienced and non-experienced subjects in treatments K and NK using a constant, a dummy variable to indicate the K treatment ("K"), the number of correct sudokus in the trial phase ("Pretest"), the average of grades obtained by each student in the preceding term ("Grade"), a

[18] The range and the variance of the dependent variable are sufficiently high such that we OLS regressions are appropriate. To check the robustness of our specification we repeated the whole empirical analysis using Poisson regressions. Results are similar with respect to sign and significance levels of the coefficients (the size of coefficients is not directly comparable between these two approaches as marginal effects in Poisson models are not constant)

dummy variable to indicate 4th or 6th grade ("Year") and a dummy variable for gender ("Gender").

Table 4: Correct Sudokus and Information		
	Experienced	Non-Experienced
	REG (1) Dep. Var: # Correct Sudokus	REG (2) Dep. Var: # Correct Sudokus
Constant	-9,68 (5.87)	1.59 (6.19)
K	-2.57 (2.65)	-0.06 (2.33)
Pretest (0=Min, 6=Max in E) (0=Min, 12=Max in NE)	7.22 (0.6)***	3.14 (0.95)***
Grade (1=Worst,5=Best)	3.11 (1.42)**	1.33 (1.89)
Year (0=4th,1=6th)	9.21 (2.49)***	6.61 (3.12)**
Gender (0=Male,1=Female)	0.69 (2.40)	1.70 (2.51)
# Observations	59	47
Adj. R^2	0.70	0.58

Notes: * denotes significance at the 10% level, ** denotes significance at the 5% and *** at the 1% level. Robust standard errors are in parenthesis.

The results of both regressions confirm that there is no significant effect of providing information on performance.[19] Coefficients for the K treatment are negative in both cases but not significantly different from zero (and actually very close to zero for non-experienced subjects). Results of both regressions confirm that most of the other controls, apart from gender, are important in explaining performance and have the expected sign. Therefore we conclude:

Result 1: *Knowledge of the existence of an asymmetry in experience did not decrease performance by experienced or non-experienced subjects.*

3.3 The Effects of Affirmative Action on Performance

[19] This fact has been documented in other recent real-effort tournament. A recent example is Freeman and Gelber (2009), where performance is not substantially altered in their single prize tournament when competitor's past performance is revealed.

We now study how performance in the tournament was affected by the implementation of AA policies using suitable controls for individual heterogeneity. To that end we here compare performance in the K and AA treatments. Figure 4 plots the average difference in the number of correctly solved sudokus both for experienced and non-experienced subjects when classified by their ability, proxied by the number of sudokus they correctly solved in the practice round both for experienced and non-experienced subjects.[20] Although differences in averages are not significant under standard levels, the graph on the left suggests that experienced subjects react differently to affirmative action. Differences actually become negative for subjects who solve six sudokus in the practice round. However, non-experienced subjects seem to react positively to AA independently of their ability.[21] This may be explained by the fact that non-experienced subjects have less knowledge both about their relative performance and that of their rivals, since they are less familiar with the task.

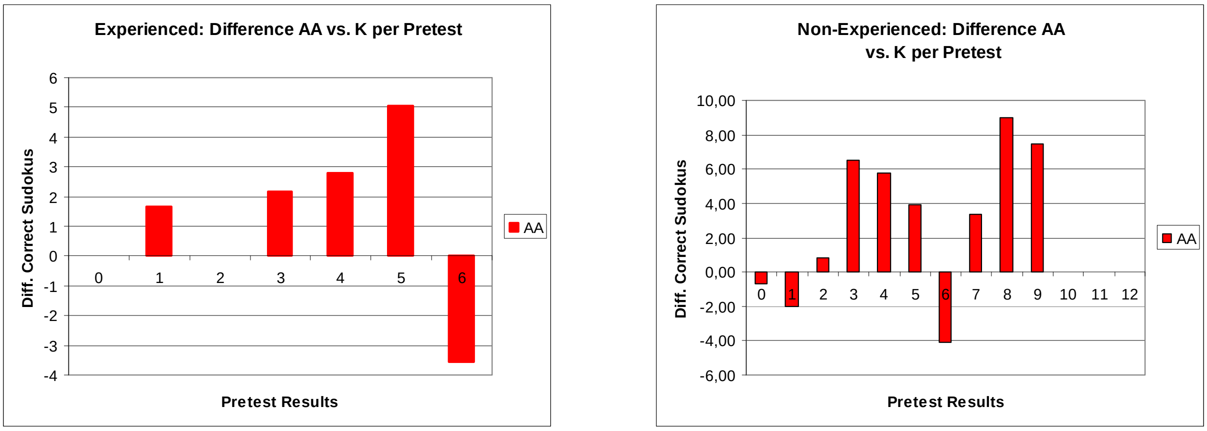


Figure 4: Differences in the number of correct sudokus by results in Pretest.

In order to assess the effect of affirmative action on performance when controlling for individual heterogeneity, we run separate regressions for each school with the number of correctly solved sudokus as a dependent variable (Table 5). Our baseline treatment is K, where

[20] Experienced subjects were provided with only 6 sudokus as part of their practice round. 40% of experienced subjects solved all six sudokus correctly. For this reason, our measure of ability for experienced subjects is cut off at 6, since it includes individuals who would possibly have solved more than 6 sudokus. We thus expect our estimated parameters to be smaller and less significant (due to higher variance). For the subsequently run experimental sessions with non-experienced subjects, we extended the number of trial sudokus to 12. By truncating these data artificially in the same manner as for experienced subjects, we were able to verify the conjecture that when there are fewer sudokus in practice rounds results become slightly less significant and weaker in absolute size without altering the qualitative results.

[21] The negative sign for subjects who solve six sudokus in the practice round is due to two exceptionally high performers in the K treatment.

subjects are aware of the existing disadvantage but no AA policy is implemented.[22] Explanatory variables are the AA treatment dummies (both pooled and unpooled in two separate regressions) and the other controls used in regressions (1) and (2). Given the results suggested by Figure 4, we included an interaction term ("AA*Pretest") for experienced subjects, such that different reactions by individuals with different individual ability could be accounted for.[23] The inclusion of the interaction term implies that a representative subject of the base group under K has low ability (zero correct sudokus in the five minute trials). In table 5 the results from OLS regressions with robust and clustered errors are presented for both E and NE schools, pooling the AA treatments in REG (3) and (5) and unpooled in REG (4) and (6).

Similarly to regressions in the previous subsection, our proxy for unobserved ability as measured by results in the practice rounds ("Pretest"), school grades ("Grade"), and being in sixth grade instead of fourth grade ("Year") have all positive and significant effects,[24] while the effect of "Gender" is not significant.[25]

We now focus on experienced subjects (E). REG (3) shows that when pooling all affirmative action treatments, the coefficient for AA has a positive and significant impact (at the 5% level), i.e., experienced subjects in the base group (with low ability) statistically solve 9.14 more sudokus when they compete with subjects favored by an affirmative action policy. However, the higher the ability of the experienced subject (measured by "Pretest"), the lower the increase in AA performance, since the interaction term ("AA*Pretest") is negative and significant (at the 5% level).[26]

Table 5: Correct Sudokus and Affirmative Action

	Experienced		Non-Experienced	
	REG (3) Dep. Var:	REG (4) Dep. Var:	REG (5) Dep. Var:	REG (6) Dep. Var:

[22] All results are maintained using baseline treatment K but including data from the NK treatment as an additional dummy variable.
[23] Similar results are obtained when creating a dummy variable in order to account for the truncation of our measure of ability for experienced subjects (0 for less than six sudokus in Pretest, 1 for more than 6 sudokus).
[24] "Grade" is not statistically significant for NE subjects.
[25] There exists an important literature analyzing how male and female individuals react differently to competition (see Gneezy et al. (2003), Gneezy and Rustichini (2004) and Niederle and Vesterlund (2007)). Subjects in our experiment did not know whether they were competing with a rival of the opposite gender, which may explain our result.
[26] The statistical effect of AA on subjects with higher ability can therefore be calculated as "AA"+"AA*Pretest".

	# Correct Sudokus	# Correct Sudokus	# Correct Sudokus	# Correct Sudokus
Constant	-14.22 (3.74)***	-14.47 (3.84)***	3.01 (3.58)	3.16 (3.62)
AA	**9.14** **(3.77)****	-	**3.98** **(1.66)****	-
AA*Pretest	**-1.80** **(0.78)****	-	-	-
LH	-	11.41 (6.09)*	-	4.71 (2.02)**
LL	-	-2.30 (4.63)	-	3.79 (2.01)*
PH	-	13.37 (4.25)***	-	3.64 (1.84)*
PL	-	4.59 (6.11)	-	3.76 (2.17)*
LH*Pretest	-	-2.35 (1.08)**	-	-
LL*Pretest	-	0.38 (0.92)	-	-
PH*Pretest	-	-2.20 (0.86)**	-	-
PL*Pretest	-	-1.13 (1.49)	-	-
Pretest (0=Min, 6=Max in E) (0=Min, 12=Max in NE)	7.12 (0.54)***	7.08 (0.56)***	4.25 (0.60)***	4.23 (0.62)***
Grade (1=Worst,5=Best)	3.65 (0.80)***	3.83 (0.75)***	0.03 (0.87)	0.03 (0.89)
Year (0=4th,1=6th)	10.62 (1.63)***	10.22 (1.56)***	4.46 (1.61)***	4.42 (1.65)**
Gender (0=Male,1=Female)	0.82 (1.66)	0.77 (1.64)	0.46 (1.35)	0.43 (1.37)
# Observations	150	150	132[a]	132[a]
Adj. R^2	0.66	0.67	0.65	0.65

Notes: * denotes significance at the 10% level, ** denotes significance at the 5% and *** at the 1% level.

Robust standard errors, clustered by treatment and class are in parenthesis.

(a): For one non-experienced subject "Grade" was not available. Another subject arrived late and did not participate in the practice rounds. Such observations are omitted from REG (3) and REG (4).

The unpooled analysis of the AA treatments in REG (4) shows significant positive coefficients for the high values of both affirmative action policies (at the 10% level for LH and

5% level for PH), while coefficients are not significant for the low values (LL and PL).[27] This suggests that the more intensive AA treatments, i.e. LH and PH, are the main contributors to the described incentive effects. This is confirmed by the negative and significant results (at the 5% level) of interacting Pretest with the high policies of affirmative action ("LH*Pretest" and "PH*Pretest").We thus conclude:

> Result 2: *Affirmative Action policies enhanced the performance of experienced subjects with the exception of subjects of highest ability. Most of the effect can be attributed to treatments where the compensation was high, i.e., PH and LH.*

We now focus on non-experienced subjects (NE). REG (5) shows that when pooling all affirmative action treatments, the coefficient for AA has a positive and significant effect (at the 5% level), i.e., non-experienced subjects solve 3.93 more sudokus when they are favored by an affirmation action policy. REG (6) suggests that the incentive enhancing effects of affirmative action cannot be attributed to a specific type of policy. Coefficients for all AA treatments (LH, LL, PH and PL) are similar in size and significance levels (5% in the case of LH, 10% in the others). Hence, there seems to be a performance enhancing effect of affirmative action policies on non-experienced subjects which is independent of its form or size.[28] The lack of sensitivity to the specific design, may again be the result of their lack of familiarity with the task which may reduce their capacity to assess the relative size of compensations. We thus conclude:

> Result 3: *Affirmative Action policies enhanced the performance of non-experienced subjects independently of the size and type of the implemented policy.*

3.4 The Effects of Affirmative Action on the Selection of Tournament Winners

[27] Similar regressions pooling the LH and PH treatments under one "high" variable and LL and PL under a "low" one show the same results.

[28] Similar linear regressions using the total number of solved sudokus as regressor, independently of them being correct or not, lead to the same conclusions and are available upon request.

An important concern that the implementation of affirmative action raises is that the selected pool of candidates is of lower ability because of the higher proportion of selected disadvantaged individuals who may perform poorly. There are two different approaches to answering this question, crucially depending on what the objective of these policies is. First, the objective may be to select individuals according to some ability that is unobserved but equally distributed among advantaged and disadvantaged individuals. If ability is positively correlated with performance, but differently so for the two groups, selecting a similar proportion of the best performing individuals in both groups should lead to selecting the highest ability individuals overall. Computing the proportion of all possible tournament winners[29] we find that only 23% of the non-experienced subjects win their respective tournament in the K treatment. When AA is implemented this percentage is increased to 51%, reflecting that the tournament is now at selecting the upper tails of the ability distributions. This result shows that the implemented AA policies leveled on average the playing field.

Second, if the objective of the tournament is to select the highest performing individuals, then the average performance of tournament winners may be lower under AA, since a higher number of disadvantaged subjects are selected. However, the increase in overall performance illustrated in the last section suggests that this reduction may be smaller than expected. Comparing the average number of correct sudokus solved by all possible tournament winners shows that in our experiment both forces are important. Average performance in the AA treatments was 2.93% lower than in the K treatment, although not significantly so. But when controlling for age there is a statistically significant decrease is of 6.46% for 4th graders and of 8.17%f or 6th graders.

3.5 The Effects of Affirmative Action on Expected Winning Probabilities

We here look at how subjects' expectations about winning their respective tournament were affected by affirmative action. This issue is important to study whether affirmative action undermined self-confidence. In question 6 of the questionnaire, subjects ranked their expectation of winning the tournament against their respective rival on an ordinal scale from 1 ("Definitely Not") to 5 ("Definitely"). As there was no information about the identity and characteristics of

[29] To find the expected tournament winners we computed the mean among all possible matches within each treatment. Note that the particular match used to reward subjects in our experiment was just one random realization of this process.

the respective opponent (with the exception of ex-ante experience in the AA treatments and in treatment K) we use these answers as a measure of confidence in winning. PROB (7) and PROB (8) in Table 6 show ordered probit regressions using our measure of confidence in winning as dependent variable and "Pretest" and the treatment dummy for affirmative action ("AA") as regressors. Understandably, we find that both E and NE high ability subjects have higher confidence in winning their respective tournament as "Pretest" has a positive and significant coefficient at the 1% level in both regressions. More importantly, we find that while for experienced subjects the presence of AA does not significantly affect reported confidence, it significantly increases the confidence of non-experienced subjects at the 5% level. These results are consistent with the experienced subjects not feeling frustrated by the introduction of affirmative action while, at the same time, AA correctly increasing the expectations of the non-experienced subjects of winning their respective tournament

Table 6: Expected Winning Probability, Affirmative Action and Ability

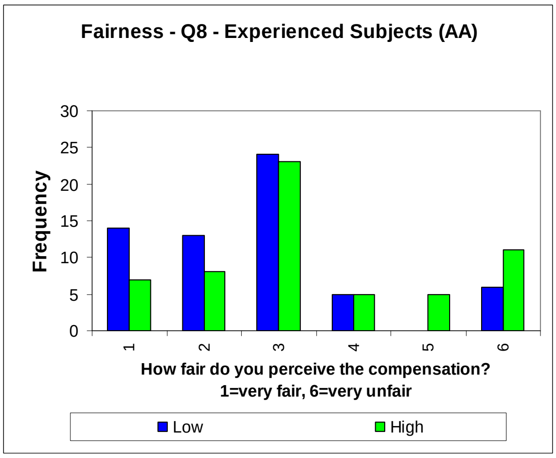
	Experienced	Non-Experienced
	PROB (1) Dep. Var.: Win Prob.	PROB (2) Dep. Var.: Win Prob.
AA	-0.13 (0.16)	0.42 (0.21)**
Pretest	0.22 (0.06)***	0.12 (0.04)***
# Observations	179	148
Pseudo R^2	0.038	0.033

Notes: * denotes significance at the 10% level, ** denotes significance at the 5% and *** at the 1% level. Robust standard errors, clustered by treatment and class are in parenthesis.

3.6 The Effects of Affirmative Action on Perception of Fairness

We here explore how affirmative action policies affected subjects' judgements about the fairness of the competition. This question is important to determine subjects' attitude towards the competition. The analysis of the tournament performance suggests that the incentive effects of AA are stronger for high levels of compensation. We now show that experienced subjects perceive the different AA policies as substantially different with respect to their inherent fairness based on the analysis of the responses to question No. 8 of the post-experimental questionnaire. In this question subjects were asked for their perceived fairness of the implemented bonus in their treatment, where responses could vary between 1 (very fair) and 6 (very unfair).

We start by looking at experienced subjects. Figure 5 suggest that experienced subjects perceived the high treatments (LH and PH) as more unfair than the low treatments (LL and PL), and the lump-sum treatments (LL and LH) as more unfair than the proportional treatments (PL and PH).



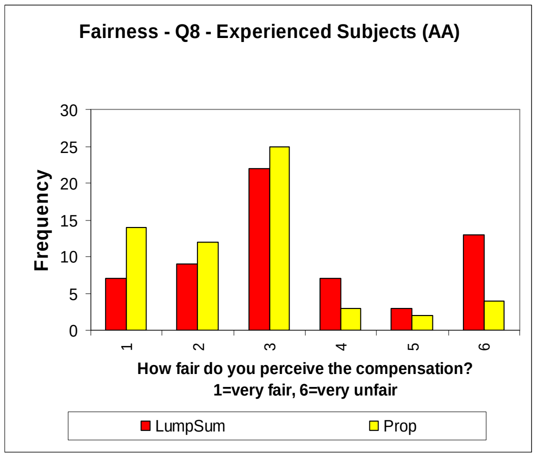


Figure 5: Fairness perception by experienced subjects for AA treatments.

The results of a non-parametric two-sided Mann-Whitney test confirm that the two distributions are significantly different from each other (p-value of 0.01 for low vs. high treatments and p-value of 0.00 for lump-sum vs. proportional treatments).[30] This is quite remarkable given that on average the compensation received by subjects in the proportional treatments was higher than those in the lump sum treatment. Compensations in the proportional treatments were on average 12 additional sudokus in the PL treatment and 22 additional sudokus

[30] The LH treatment was perceived as being significantly more unfair than any other AA treatment (p-values of 0.04 for LL vs. LH, 0.01 for PL vs. LH, and 0.02 for PH vs. LH for a two-sided MW-test).

in the PH treatment, both higher than 8 and 20, the compensations received in LL and LH respectively. The fact that the compensation size depends on performance seems to be fairer to them, although it decreases their chances of winning more than lump sum bonuses.[31]

Similar results hold for the comparison of treatments NK vs. K. First, the left side of figure 6 suggests that experienced subjects perceived the tournament to be fairer when they are not aware there is an asymmetry in experience. Non-parametric tests also confirm that treatment K is perceived as being significantly less fair than treatment NK (p-value of 0.02 for two-sided MW-test). Quite remarkably, subjects seem to evaluate the fairness of a treatment by abstracting, at least partially, from their self-interest.

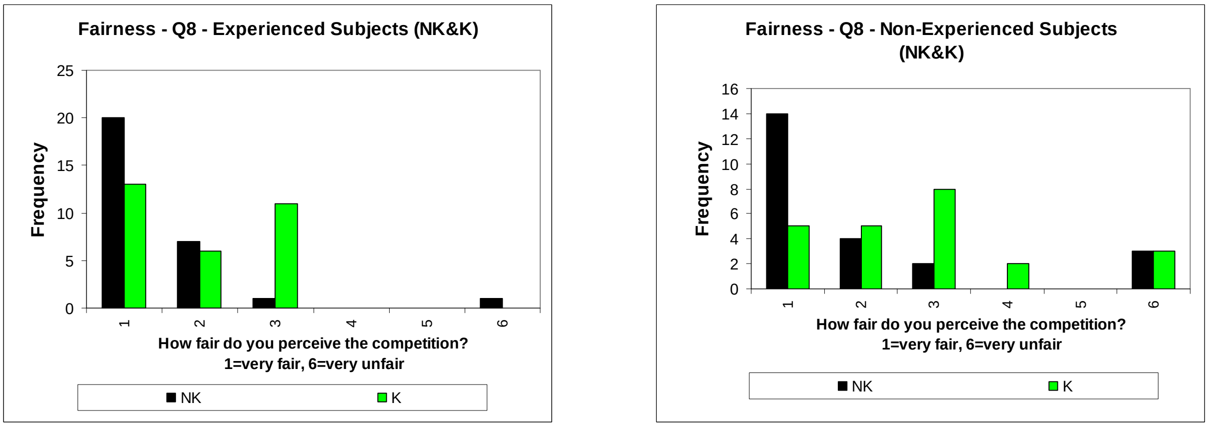
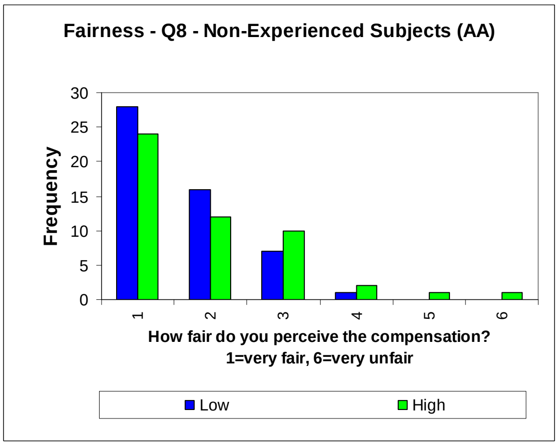


Figure 6: Fairness perception by experienced (left) and non-experienced subjects (right) for K and NK

Results for non-experienced subjects point in the same direction. For instance, the fairness of treatment K is perceived as being significantly more unfair than treatment NK (p-value of 0.013 for two-sided MW-test), which is also suggested in the right part of figure 6. The responses of the non-experienced subjects for the AA treatments suggest that the compensation that is granted through AA is mostly perceived as being sufficiently fair. A more detailed analysis for the high versus low compensations as well as the lump-sum versus proportional treatments indicates similar perceptions as for the experienced subjects: low (proportional) treatments seem to be perceived as slightly fairer than the high (lump-sum) treatments although non-experienced students actually benefit from high treatments. However, these distributions are not significantly different from each other, which may be a result of non-experienced subjects

[31] As often suggested in the literature on positive fairness. See Konow (2000) for a survey.

being less able to assess the appropriateness of the compensation because of lack of exposure to the task.[32]



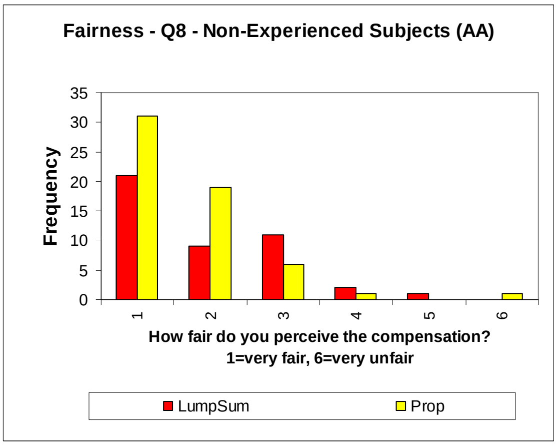


Figure 7: Fairness perception by non-experienced subjects for AA treatments.

4. Conclusion

This paper presents evidence on the effects in individual performance of implementing affirmative action policies in a tournament where capacities to compete are asymmetric. Theory in such settings predicts that preferential treatment of disadvantaged individuals may help or harm incentives to invest depending on the details of the setting. We here show a setting in which compensating disadvantaged individuals through AA policies in a tournament leads to enhanced performance by a large fraction of participants and to a small decline in the average performance of selected winners, while balancing the pool of selecting winners. Our results thus imply that there exist circumstances under which affirmative action policies are beneficial with respect to the incentives provided to all participants. They also suggest that different AA designs may significantly affect incentives and fairness perception substantially.

This paper provides a piece of evidence suggesting that AA can be beneficial in situations where there exists asymmetries in capacities to compete which are not individuals' responsibility. At the same time it leaves a large number of important questions still open, such as the robustness of these results to other settings, other sources of asymmetry not affecting

[32] The only exception is again the LH treatment which is perceived as being significantly less fair than all other AA treatments (p-values of 0.07 for LL vs. LH, 0.07 for PL vs. LH, and of 0.06 for PH vs. LH for a two-sided MW-test).

capacities to compete directly, or the long run effects of these policies. We plan to address these issues in future research.

5. References

Balart, P., 2009. "Equality of Opportunities and Welfare in a Model of Job Allocation". Universitat Autònoma de Barcelona. Mimeo.

Bertrand, M., Hanna, R. and Mullainathan, S."Affirmative Action: Evidence from College Admissions in India", NBER Working Paper #13926, forthcoming at *Journal of Public Economics*.

Calsamiglia, C., 2009. "Decentralizing Equality of Opportunity". *International Economic Review* 50(1): 273-290.

Coate, S., and Loury, S., 1993. "Will Affirmative-Action Policies Eliminate Negative Stereotypes?". *American Economic Review* 83(5): 1355-1360.

Chan, J., and Eyster, E., 2003. "Does Banning Affirmative Action Lower College Student Quality?". *American Economic Review* (93):858-872.

Che, Y., and Gale, I., 1998. "Caps on Political Lobbying". *American Economic Review* (88): 643-651.

Franke, J., 2008. "Does Affirmative Action Reduce Effort Incentives? A Contest Game Analysis". UAB and IAE Working Paper Series 711.07.

Freeman, R. B., Gelber, A. M., 2010. "Prize Structure and Information in Tournaments: Experimental Evidence". *American Economic Journal: Applied Economics* (2): 149-164.

Fryer, R. G. J., and G. C. Loury, 2005a. "Affirmative Action in Winner-Take-All Markets." *Journal of Economic Inequality* 3(3): 263-280.

Fryer, R. G. J., and G. C. Loury, 2005b. “Affirmative Action and Its Mythology”. *Journal of Economic Perspectives* 19(3): 147-162.

Fryer, R., and Loury, G., forthcoming “Valuing Identity: The Simple Economics of Affirmative Action Policies” *Journal of Political Economy*.

Fu, Q., 2006. “A Theory of Affirmative Action in College Admissions”. *Economic Inquiry* 44: 420-428.

Gneezy, U., Niederle M., and A. Rustichini, 2003. “Performance in Competitive Environments: Gender Differences”. *Quarterly Journal of Economics* 118(3): 1049-1074.

Gneezy, U., and A. Rustichini, 2004. “Gender and Competition at a Young Age,” *American Economic Review Papers and Proceedings*: 377-381.

Harbaugh, W., and Krause, K., 2000. “Children's Altruism in Public Good and Dictator Experiments,” *Economic Inquiry* 38(1): 95-109.

Harbaugh, W., Krause, K. and Berry, T. R., 2001. “GARP for Kids. On the Development of Rational Choice Behavior”. *American Economic Review* 91(5): 1539-1545.

Hoff, K., and P. Pandey, 2006. “Discrimination, Social Identity, and Durable Inequalities”. *American Economic Review Papers and Proceedings* 96(2): 206-211.

Holzer, H., and D. Neumark, 2000. “Assessing Affirmative Action”. *Journal of Economic Literature* 38(3): 483-568.

Kane, T.J., 1998. “Racial and Ethnic Preferences in College Admissions” Ch. 12 in C. Jencks and M. Phillips (eds.) *The Black-White Test Score Gap.* Washington D.C., Brookings Institution Press: 431-456.

Krasnokutskaya, E., and K. Seim, 2007. "Determinants of the Participation Decision in Highway Procurement Auctions". University of Pennsyilvania. Mimeo.

Konow, J., 2003. "Which is the fairest one of all? A positive analysis of justice theories". *Journal of Economic Literature,*.41(4): 1186-1237.

Lazear, E..P., and S. Rosen, 1981. "Rank-Order Tournaments as Optimum Labor Contracts". *Journal of Political Economy* 1989(5): 841-864.

Lott, J., 2000. "Does a Helping Hand Put Others at Risk?: Affirmative Action, Police Departments and Crime ". *Economic Inquiry* 38(2)

Loury, L.D., and Garman, D., 1993. "Affirmative Action in Higher Education". *American Economic Review Papers & Proceedings* 83(2): 99-103.

Marion, J., 2007. "Are Bid Preferences Benign? The Effect of Small Business Subsidies in Highway Procurement Auctions". *Journal of Public Economics* 91: 1591-1624.

Miller, A., and Segal, C., 2008. "Do Temporary Affirmative Action Programs Produce Permanent Effects? A Study of Minority and Female Employment in Law Enforcement Agencies". Universitat Pompeu Fabra. Mimeo.

Myerson, R., 1981: "Optimal auction design". *Mathematics of Operations Research* 6: 58-73.

Niederle, M., C. Segal and Vesterlund, L., 2008. "How Costly is Diversity? Affirmative Action in Competitive Environments". Universitat Pompeu Fabra. Mimeo.

Niederle, M., and Vesterlund, L., 2007. "Do Women Shy Away From Competition? Do Men Compete Too Much?". *Quarterly Journal of Economics* 122(3): 1067-1101.

Prendergast, C., 1999. “The Provision of Incentives in Firms”. *Journal of Economic Literature* *37*(1): 7-63.

Schotter, A., and Weigelt, K., 1992. “Asymmetric Tournaments, Equal Opportunity Laws, and Affirmative Action: Some Experimental Results”. *Quarterly Journal of Economics* 107(2): 511-39.

Sowell, T., 2004. Affirmative Action around the World: An Empirical Study. *New Haven, Conn: Yale University Press.*

6. Appendix

Experimental Instructions

Below you can find a translation of the experimental instructions used in the experiment. Instructions for all treatments and schools were identical apart from the changes here indicated. Sentences in bold were not included in the Not Know treatment. The sentences in *bold and italics* were only included in the treatments with affirmative action (LH, LL, PH, PL). Words in (parenthesis) indicate changes between the experiences and non-experienced schools and changes in the type of the compensations (lump-sum or proportional). Sizes of the compensations varied as explained in section 2. Numerical examples varied in order to reflect changes in compensation sizes, but were created such as the results of both contestants were the same. A whole set of instructions is available upon request. Instructions were originally written in Spanish.

Pre-instructions

Your Code: ____

Thank you for participating. First, we are going to explain what you will be doing.

You have to fill in grids with the numbers 1, 2, 3 and 4.

To do this you have to use the following rules:

1. All boxes in a grid must be filled in with a number.
2. The same number can appear only once in each column (vertical).
3. The same number can appear only once in each row (horizontal).
4. The same number can appear only once in each square. Each grid is divided in 4 squares, marked in bold lines.
5. In each grid all numbers 1, 2, 3 and 4 must be in each column, each row, and each square.

Here are some examples:

3			
1			
3			
4			

This column is **completed wrongly**
because the 3 appears twice (rule 2)

3			
1			
2			
4			

This column is **completed correctly.**

2	4	3	4

This row is **completed wrongly**
because the 4 appears twice (rule 3

2	4	3	1

This row is **completed correctly.**

4	1		
1	3		

This square is **completed wrongly**
because the 1 appears twice (rule 4)

4	1		
2	3		

This square is **completed correctly.**

This is an example of a correctly completed grid.

4	1	2	3
2	3	4	1
3	4	1	2
1	2	3	4

Before starting you have 5 minutes to complete the following grids to check whether you have understood the rules. We will give you the correctly completed grids after the 5 minutes period.

	2		3
			4
2			
	3		

	1		2
3		4	

		1	4
2	1		

2			
	3	4	
		1	3

3	4		
	3		
			1

3		1	4
		4	3

	1		
			3
1	4		

3			
	1		
2		3	

		2	3
2	4		

		1	
	3		
			4
	4	3	

			1
	3	2	
3	4		

			1
		3	
			2
	2	1	

Please remain silent and on your seat without disturbing anyone during the whole practice.

Raise your hand after you have finished all grids and we will pick them up.

Good luck!

Instructions

Your Code: ____

You are randomly matched with another student, your matched participant, from another school similar to yours, who is completing the same grids as you are.

The students at the other school have (NOT) learned before how to solve these types of grids because it was (NOT) taught to them in their math classes.

You have now 30 minutes time to complete as many grids as possible with the numbers **1, 2, 3 and 4** on the formulaires that we are now going to distribute.

We will compare how many grids you have solved correctly with the number of correctly solved grids by your matched participant from the other school:

- If you have correctly solved more grids then you will earn a 7 EU voucher that you can redeem in "La Casa del Libro", where you can buy books, collector's cards, toys, music or comics.
- If you have correctly solved less grids then you will not earn the voucher.
- If you have correctly solved the same number of grids, then a toss of a coin will be used to determine who earns the voucher.

To compensate (the other students) for the fact that (they)/(you) have (less)/(more) practice (than you) we are going to give (them)/(you) (20 extra grids)/(1 grid more for each grid that (they)/(you) solve correctly).

For example *(example provided for the PH Treatment)*:

- If your matched participant correctly solves **12** grids, they count as 12 +12 = 24 grids. Therefore you will earn the voucher if you solve correctly 25 grids or more.
- If your matched participant correctly solves **30** grids, they count as 30 +30 = 60 grids. Therefore you will not earn the voucher if you solve correctly 59 grids or less.

- If your matched participant correctly solves **20** grids, they count as 20 +20 = 40 grids. Therefore, if you solve correctly 40 grids, a toss of a coin determines whether you earn the voucher.

The numbers of this example are chosen randomly and do not indicate how many grids a student can solve correctly.

We would like to inform you that we have studied the results of other students of your age from other schools who completed the same grids: The maximum number of grids that somebody managed to solve correctly in 30 minutes were 81 grids and the minimum was 0 grids. On average the students completed around 25 grids correctly.

Remember that only correctly solved grids count.

Wait to turn the answer sheet until we tell you to do so. You have 30 minutes. Good luck!

Your Code: ____

Thank you for your participation.

Final Questionnaire

Please answer the following questions:

1. How did you find today's task?

   Interesting   Entertaining   A bit long   Boring

2. How many grids like these have you tried before?

   None   Between 1 and 5   Between 6 and 20   Between 20 and 40   More than 40

3. If you have tried solving grids like these before, where did you do it? _________

4. How many grids do you think you have solved correctly today? ______________

5. How many grids do you think your partner of the other school has solved correctly?____

6. Do you think you are going to get the voucher?

   Definitely   Probably yes   I don't know   Probably not   Definitely not

7. Do you think it was a good idea to compensate the students of the other school that did not do grids like this before in school?

YES NO

8. The competition with the students of the other school from my perspective seemed to be:

Fair Rather Fair A bit Unfair Unfair Rather Unfair Very Unfair

9. Any other comment? ____