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

This Version: June 16, 2009

Abstract

Affirmative action policies are often applied in tournament-like environments by biasing competition rules in order to provide equal opportunities to a group of competitors having a disadvantage they cannot be held responsible for. They are frequently criticized because they may distort incentives and result in lower performance by both advantaged and disadvantaged individuals and in an inefficient selection of tournament winners. We study the empirical validity of such claims in a pair-wise real effort tournament in which there exists a naturally induced source of disadvantage for one group of subjects. We find no evidence that affirmative action policies induce losses in performance of both advantaged and disadvantaged subjects. Instead, we find that performance is actually weakly enhanced and that, while balancing, the proportion of disadvantaged individuals winning their respective tournament (from 24% to 52%), the average performance of tournament winners does not significantly decrease.

Keywords: Affirmative action, tournaments, real effort, experiment, Sudoku.

JEL classification: C72; C91; J78; M52

---

* We are especially grateful to Antoni Calvo-Armengol for his support and encouragement on this project. We thank Miguel A. Ballester, Carmen Beviá, Jeanette Brosig, Antonio Cabrales, Marta García-Matos, Uri Gneezy, Guillaume Haeringer, Nagore Iriberri, Inés Macho-Stadler, Muriel Niederle, Carmit Segal, Tom Palfrey and seminar audiences at Rady School of Management, UC San Diego, Yale University, Universidad Pablo de Olavide, Università di Padova, Universidad de Navarra, Universität Trier, Universitat Autònoma de Barcelona, IMEBE (Alicante), ESA Tucson 2008 and the Conference in Memoriam of Antoni Calvó-Armengol for their comments. 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 research assistance. 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 Educacion y Ciencia (SEJ2006-00538 and Consolider-Ingenio CSD2006-00016), Barcelona GSE Research Network and of the Government of Catalonia (2005SGR-00836).

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

[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.es.

1. Introduction

Affirmative Action policies (AA) take proactive steps to bias competition rules in order to ameliorate a possible disadvantage of a discriminated group and as such, aim to provide equal opportunities of success to individuals having a competitive disadvantage they cannot be held responsible for.[2] Its implementation is usually accompanied by an intense public debate focusing on whether such policies satisfy certain fairness criteria and on the possible incentive distortions they may create. Abstracting from the fairness considerations of such policies we find that opponents of AA provide two types of reasons against it. First, advantaged individuals might perceive the preferential treatment of their (disadvantaged) opponents as unfair, which could lead to discouragement and thus, lower effort exertion. Disadvantaged individuals could anticipate this reaction by their opponents or perceive the AA policy simply as a substitute for own effort, leading also to lower effort exertion.[3] Second, since AA is frequently to select a group of successful competitors they argue, from an ex-post point of view, that the pool of selected individuals may be less efficient than the one obtained had AA not been implemented. 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 effort exertion. Surprisingly, although both positions fail to base their views on solid empirical findings, very little research in economics provides empirical evidence on whether AA improves or worsens performance (see Holzer and Neumark, (2000)). We here 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 distinctive types of AA policies are implemented to compensate for it. .

Many of the situations in which affirmative action policies are called for, such as university admissions, job promotions or procurement auctions can be understood as tournaments for a limited number of prizes. In such tournaments one group of individuals may share a specific characteristic, which is a source of disadvantage when competing for positions in a society. For example, students from poor economic backgrounds may have attended less funded high schools, which may affect their SATs performance and thus, university admission. Similarly, individuals may belong to a historically discriminated group, which may have to overcome severe obstacles in order to be in equal conditions to compete. Affirmative action policies correct asymmetries in such competitive environments by using biased tournament rules that increase the probability of success of the disadvantaged group. For example, a fixed lump-sum

[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] 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”.

bonus of 20 (out of 150) points was added to the score of minority applicants to the undergraduate program at the University of Michigan[4] and a similar but "unofficial lift" scheme is used at many top universities. Proportional types of affirmative action policies are observed in public procurement auctions where bid preferences are granted in a multiplicative way, e.g., road construction contracts in California are auctioned off by granting a 5% reduction of the submitted bid to small business enterprises.[5]

We analyze how AA affects performance in the context of a tournament where the asymmetry between otherwise similar students exists ex-ante.[6] We design pair-wise tournaments among children from two similar schools which differ in how experienced their students are in the real effort task on which the competition is based, i.e., solving simple numerical puzzles known as "sudokus".[7] 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. Such difference in experience can be considered as an exogenous source of disadvantage in the competition between the two schools since it existed before we carried out the tournament and it is most likely not the reason why parents chose one particular school. First we study whether knowing that such an asymmetry in experience exists affects the performance of both experienced and non-experienced individuals. For example, a psychological factor that might have adverse consequences for disadvantaged contestants is the so-called "stereotype threat" (Steele et al. (2002)), where the salience of discrimination in a competitive situation affects negatively the performance of discriminated individuals.[8] To study this question we had a baseline treatment in which subjects did not know there existed an asymmetry on experience with sudokus (and no affirmative action measures were used). Second, we implement two types of compensations, lump sum and proportional bonuses, designed to roughly equalize the probability of non-experienced students beating their experienced rival. We then study whether performance by students from both schools is affected

[4] This procedure was recently ruled as unconstitutional by the Supreme Court precisely partially due to the alleged distortionary effects on incentives to perform effort such compensation may create. State funded universities such as California, Florida and Texas also applied similar measures in the past.

[5] Other measures, such as quotas are also aimed to correct asymmetries by imposing that a certain percent of winners must belong to the disadvantaged group. This may create different types of effort distortions. The effects of quotas on tournament participation have recently been analyzed in Niederle, Segal and Vesterlund (2009).

[6] Coate and Loury (1993) show how discrimination may arise in two symmetric groups as a self-fulfilling prophecy. In their context affirmative action may or may not help solve the problem.

[7] 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 the digits from 1 to 9 only once. The puzzle setter provides a partially completed grid. We use a simplified 4x4 grid version for reasons explained below.

[8] This effect has been experimentally verified in unbiased competitive environments where no affirmative action is applied. In Hoff and Pandey (2006), for instance, the performance of discriminated caste members in rural India declined in mixed caste tournaments if their caste membership was revealed. These considerations would imply less effort exertion by discriminated individuals when the asymmetry is made salient.

and whether the output of the new pool of tournament winners differs from the one obtained without such compensations.

The closest theoretical papers explicitly addressing the incentive effects of affirmative action are Fu (2006), Balart (2009) and Franke (2008). They model affirmative action as a bias in favor of ex-ante disadvantaged players in an all-pay auction, or contest set-up. Similarly to the simple model we present below, the conclusion that can be drawn from these papers is that reducing the asymmetry in competitive advantage tends to improve individual performance. The intuition is straight forward: appropriately designed affirmative action (that equalizes rewards to effort) reduces the asymmetry between players, which increases competitive pressure and hence incentives for effort exertion.[9] Also, Fryer and Loury (2005) show that optimal affirmative action in winner take all tournaments should involve handicapping. This intuition is not only restricted to the affirmative action framework and similar results were first established for rank-order tournaments in Lazear and Rosen (1981), or for optimal auctions in Myerson (1981), where it is shown that favoring weak players might induce efficient allocations or maximization of expected revenue, respectively.

There exists a large empirical literature on tournaments and specifically on how the size of prizes affects competition, see Prendergast (1999) for a survey. However, empirical studies that explicitly analyze the incentive effects of affirmative action in tournament-like situations are scarce.[10] A recent exception is the empirical analysis of the mentioned road construction contracts by Krasnokutskaya and Seim (2007) and Marion (2007), where it is shown that bid preferences for small businesses, besides balancing the asymmetry of entrants, induce higher procurement costs because the entry decisions of low cost large firms are distorted.

Schotter and Weigelt (1992) study the incentive effects of AA in experimental tournaments in a laboratory set-up where effort exertion is modeled as an individual decision problem based on monetary costs. Subjects' exogenous disadvantage is induced through assigning different cost parameters for which individuals are later compensated for by affirmative action. This procedure allows varying the size of the asymmetry and tailoring the compensations in order to level the playing field. Similarly to the theoretical predictions we discuss below, they obtain that AA can either boost or worsen performance depending on the sizes of the cost disadvantage and the compensation implemented. In our study the incentive effects of AA policies are analyzed in a real effort tournament where the asymmetry between subjects existed ex-ante and was not induced by the experimenter. Since we did not have an exact ex-ante measure of the size of the asymmetry we relied on results from pilot experiments to roughly calculate two different types (and sizes) of compensations which, on average, level the playing field. Several recent

[9] Calsamiglia (2009) shows that an appropriately designed AA policy should equalize rewards to effort whenever the setup affects one of many aspects determining individual final welfare.

[10] See Holzer and Neumark (2000) for a survey on affirmative action policies in the US and their consequences.

experimental studies employ similar strategies based on natural occurring differences in characteristics among social groups, such as Hoff and Pandey (2006), where social caste differences are exploited, as well as Gneezy et al. (2003), and Niederle and Vesterlund (2007), where gender differences are used to analyze the performance of women versus men and their respective propensity to compete in mixed gender tournaments. Niederle et al. (2008) combines this insight on gender differences with participation decisions by individuals in an affirmative action framework based on gender quotas. Participation is not an issue in our framework because students had to stay in the classroom, although they could decide to do some other activity quietly, which some chose to do after a while. Our design also benefits from several of the features attributed to field experiments by Harrison and List (2004). For example, since subjects were school children, they were unaware that their choices were the object of study as the situation was presented to them as an extra-curricula activity in which they could obtain some prize. Such activities are not uncommon in the schools we selected for the experiment. Using children as subject pool has additional advantages, see Harbaugh et al. (2001): children react very spontaneously in competitive situations, they are not spoiled by questioning the underlying motivation of the experimenter and it is also relatively easy to provide incentives to them. It has also been shown that children react rationally and in line with economic theory--see Krause and Harbaugh (1999). Also, studying how children react to affirmative action is crucial since many of the social asymmetries may be ideally resolved at these early ages, before they are exacerbated.

The experimental results of our study suggest that the implementation of AA policies does not necessarily lead to performance losses by affected individuals. First, we obtain that knowledge about the existence of an asymmetry in ability in fact increases competitive pressure and results in (weak) increases in performance. Most importantly, when such asymmetry is corrected through AA policies, performance by both advantaged (experienced) and disadvantaged (non-experienced) individuals increases even further. We also show that increases in performance differ depending on the level of ability of the individual, and less importantly on gender. We also find that AA affects positively the confidence in winning that the non-experienced have and does not decrease that of the experienced. Finally, we show that the average effort of tournament winners does not significantly decrease when AA is implemented.

The rest of the paper is organized as follows. Experimental design and procedures are explained in section 2. Section 3 describes a theoretical model which illustrates the effects that the types and sizes of affirmative action policies may have on performance. Section 4 shows the results. Section 5 concludes. The Appendix contains experimental instructions.

2. Experimental Design and Procedures

We conduct pair-wise tournaments among 337 school children, aged 10-13, from two similar private schools in Barcelona, which have a systematic difference in experience with a specific real effort task, i.e., solving simple sudokus.[11] 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 in the "non-experienced" school (NE) do not.[12]

Each student from E was randomly and anonymously matched with a student from NE in their same school year (4^{th} or 6^{th} grade). Each pair competed in a tournament which lasted 30 minutes. Subjects had to correctly solve as many simple 4x4 sudokus as possible in order to beat their matched rival. To do so, a whole grid had to be filled in with numbers from 1 to 4 such that the same number can only appear once in each column, row, and box of the grid. 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.[13]

All subjects were handed the same answer sheet containing 96 sudokus randomly generated with the same level of difficulty by a computer program.[14] Figure 1 below shows one of the sudokus used in the experiment (a) and its solution (b).

	4		2
		3	
1			

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

[11] Although our analysis does not focus on the inter-school comparison, both school share many characteristics. They are both mixed-gender private schools located in the same upper-class neighborhood in Barcelona, they are not religious and they are both bilingual.

[12] An ex-post experimental questionnaire showed that some students from both schools were familiar with sudokus due to private experience. In any case, results from pilot experiments, as well as in this one, showed that subjects from NE were in fact disadvantaged in the competition (see section 4.1). The task was defined as "filling-up a grid" and the word "Sudoku" was never mentioned.

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

[14] 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 provided sudokus.

(a) Unsolved Sudoku (b) Correctly Solved Sudoku

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

Each pair of subjects was competing for a 7 Euro voucher from a bookshop located in Barcelona.[15] For each pair, the student who had correctly solved more sudokus during a 30 minute period won the voucher. In 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 consequence, on the output generated by subjects selected as tournament winners. We thus randomly assigned similar numbers of subjects from each school to each of six treatments. In treatment NK (Not Know), no subject was informed of the previous experience of subjects from the other school with solving sudokus, while in treatment K (Know) students in 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 in the NE school were not taught how to solve sudokus. In the remaining four treatments all subjects were informed about the difference in experience across schools and about the particular affirmative action policy applied on NE subjects. In treatments LH (Lump Sum High) and LL (Lump Sum Low), 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), NE subjects were given a number of solved sudokus proportional to the number of sudokus they correctly solved, one for every one correct in the case of PH and one for every two correct in PL. Table 1 summarizes our treatment design. Comparisons across treatments NK and K allow us to study the effects of information when no affirmative action policies are implemented. We will generically refer to a treatment pooling data from all treatments where affirmative action is implemented (LL, LH, PL, PH) as the AA treatment. Comparisons across treatments K and AA allow us to study the incentive effect of applying affirmative action policies once the asymmetry in experience is known. The size of the affirmative action policies were calibrated using results from pilot experiments. The objective was to design compensations that would roughly equalize the chances of winning once subjects react to the applied affirmative action policies. As shown below, on average the implemented policies roughly induced a "leveled playing field" ex-post (49% of subjects from the NE school won their respective tournaments).

[15] Subjects were explicitly told that the voucher was redeemable for "books, collector's cards, toys, music or comics". Experiments took place around the time the final Harry Potter book was published in Spain.

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

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 a way that groups were balanced according to these pre-specified characteristics. Table 2 below shows descriptive statistics of subjects assigned to each treatment in each school.[16] Small variations across treatments were mainly due to absent students and latecomers.

Table 2: Description of the Subject Pool												
	Experienced						Non-experienced					
	NK	K	LH	LL	PH	PL	NK	K	LH	LL	PH	PL
% Female	41	43	48	47	50	50	46	46	48	39	48	47
% 6th Year	48	43	45	46	46	47	50	46	59	47	48	59
Average Math Grade	3.00	3.13	3.29	3.20	2.86	3.09	3.42	3.50	3.54	3.52	3.52	3.47
Number of subjects	29	30	31	30	28	32	24	24	27	23	27	32

Subjects were unaware of their participation in an experiment. With the help of the schools' faculty, subjects were explained that this was an extracurricular activity, not dissimilar to previous ones done during the same school year. Participation was quasi-mandatory, which helped avoiding selection biases and simplified matters for the school. Following the taxonomy suggested by Harrison and List (2004), our experimental setting may thus be described as a framed field experiment.

Experiments were carried out in two separate but close dates in the two schools. In each school experimental sessions took place at different times of the day for 4th and 6th graders due to practical reasons.[17] Once subjects arrived at their school, they 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 subjects calmed and equally uniformed of the experiment. The same

[16] Average school grades are slightly higher in the NE school than in the E school. Notice, however, that this is not necessarily an indicator of NE students being better at math since the grading methodology may differ across schools.

[17] We are unaware of cross-contamination between schools or between subjects from different school-years in the same school. The timing of the experiments was carefully designed such that these problems were avoided.

experimentalist and a research assistant arrived to each of the classrooms in twenty-minute intervals and sessions started. Teachers were not present during the experimental sessions in order to minimize their influence.[18]

The experimental session lasted one hour. First, the experimentalist read some general instructions about how to solve sudokus (see "Pre-instructions" in the Appendix). Then, subjects had a five-minute practice round to solve sudokus. 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 clear that each student was competing against an anonymous student from another comparable school and that students in the other school were systematically experienced (or not) in solving sudokus (for treatment NK this information was omitted). This difference in ex-ante experience was explicitly mentioned and was used to justify the implementation of the affirmative action bias in favor of the non-experienced group in treatments LL, LH, PL, and PH (see the Appendix for one of these treatments' instructions). The tournament rules were explained giving numerical examples (specific for each treatment) for the potential outcomes of the tournament, i.e., loosing, winning, and tying. Also, aggregate information with respect to the number of correctly solved sudokus (i.e. mean, minimum and maximum) by a comparable subject pool was provided. This information, identical for all subjects, was based on the results of our pilot experiments. The experimenter also showed a seven Euro voucher to increase the credibility of the prize offered to tournament winners. After that, subjects had thirty minutes to solve sudokus in two separate handouts. After the first fifteen minutes had passed, subjects were instructed to start working on the second handout, such that we could have a measure of whether there were intra-session learning effects and whether they were over-ruled by fatigue.[19] Subjects were explicitly told that they could stop solving sudokus and start any other activity, like drawing, as long as they kept quiet and did not bother others. In fact, some subjects did precisely this.

After the thirty minutes passed, handouts were taken away and a questionnaire about previous experience with sudokus, self-confidence measures and the perceived justice behind affirmative actions policies was distributed. Once the questionnaires had been filled out, subjects continued with their regular classes. Experimentalists then randomly matched participants from both schools, determined the winners and brought the vouchers to the schools, in order for them to be distributed by school faculty.

[18] This was the reason why 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 to repeat exactly the same cues across sessions.

[19] We did not find important differences in performance between the two parts, indicating that the effects of learning and fatigue possibly cancel out. Experienced subjects did one more sudoku in the second part (significant at the 1% level). Non-experienced subjects did the same number of sudokus in both parts.

3. Theoretical Model

We here abstract from psychological aspects and present a stylized non-cooperative tournament model to illustrate how in the context of pair-wise tournaments the asymmetry in capacity to compete leads to poor performance and how the different AA policies may affect performance by both advantaged and disadvantaged individuals. The results crucially depend on the relative size of the disadvantage and the compensation implemented. The model differs from the experimental design in that in our experiment neither the tournament designers nor the competitors were aware of the size of the asymmetry and thus, compensations, although calibrated based on pilot experiments with similar subjects, were chosen at two different levels.

Our model is based on a simple two player rank-order tournament as presented in Schotter and Weigelt (1992), which is a simplified version of Lazear and Rosen (1981). The setup is as follows: Two heterogeneous players $i=1,2$ compete for a given prize with value V by exerting non-negative effort e_1, e_2. Heterogeneity affects the cost function and is due to some sort of disadvantage on the second player. In particular we have that $c_1(e_1) = e_1$, while $c_2(e_2) = b*e_2$ where $b>1$.[20]

Exerted effort by an agent is not observable but generates observable output $y_i=e_i+\varepsilon_i$ for $i=1,2$ where ε_i is a random variable which is assumed to be uniformly distributed between $-a$ and a. The decisive variable for the outcome of the tournament is output, i.e., only the winner with the higher output level will win the tournament and obtain the prize. Hence, the expected payoff for agents *1* and *2* in a tournament with *AA* can be stated in the following way:

$$\pi_1 = \Pr[y_1(e_1) > y_2(e_2) + AA]V - \frac{e_1^{\,2}}{2} \text{ and}$$

$$\pi_2 = \Pr[y_2(e_2) + AA > y1(e_1)]V - b\frac{e_2^{\,2}}{2},$$

where affirmative action policy is captured by *AA* and can take on one of the following two forms:

1. Lump-sum *AA*, with $AA = L$ where L is a positive constant,
2. Proportional *AA*, with $AA = P*e_2$ where P is a positive constant.[21]

The probability of winning for agent *1* can be reformulated as $Pr(\varepsilon_2-\varepsilon_1<e_1-e_2-AA)$. Note that the composed random variable $\varepsilon=(\varepsilon_2-\varepsilon_1)$ has a triangular distribution with the following distribution function:

[20] The asymmetry can alternatively affect the productivity of effort (how effort converts into output) without altering qualitative results.
[21] An alternative (and more consistent) specification of proportional AA would be to scale up output y_2 by factor $(1+P)$, i.e., $f(e_2)=(1+P)(e_2+\varepsilon_2)$. In this case the random error would be affected as well by AA which complicates the analysis but leads to the same qualitative results.

$$F_1(x)=\begin{cases}\frac{1}{2}+\frac{x}{2a}+\frac{x^2}{8a^2}, \text{ for x} \in [-2a,0] \\ \frac{1}{2}+\frac{x}{2a}-\frac{x^2}{8a^2}, \text{ for x} \in [0,2a]\end{cases}.$$

The respective distribution function for agent *2* is: $F_2(x)=1-F_1(x)$.

In the interior equilibrium the effort levels *(e_1*, e_2*)* will satisfy first order conditions for both individuals:

$$\frac{\partial \pi_1}{\partial e_1}=\frac{\partial F_1(e_1^*-e_2^*-AA)}{\partial e_1}V-e_1^*=0 \quad and$$
$$\frac{\partial \pi_2}{\partial e_2}=\frac{\partial F_2(e_1^*-e_2^*-AA)}{\partial e_2}V-b*e_2^*=0.$$

The equilibrium effort levels will thus depend on the format of the Affirmative Action policy. We now describe the equilibrium for each of the different designs.

Lump sum AA

When affirmative action is lump sum we have that:

$$\frac{\partial F_1(e_1^*-e_2^*-L)}{\partial e_1}=\begin{cases}\frac{1}{2a}+\frac{e_1^*-e_2^*-L}{4a^2}, \text{ for } e_1^*-e_2^*-L<0 \\ \frac{1}{2a}-\frac{e_1^*-e_2^*-L}{4a^2}, \text{ for } e_1^*-e_2^*-L\geq 0\end{cases},$$

and vice versa for the second agent. Using these expressions and solving for e_1* and e_2* in the first order conditions we find that:

$$e_1^*=be_2^*,\ \ e_2^*=\frac{(2a-L)V}{4a^2b-(b-1)V}, \text{ for } L>\frac{(b-1)V}{2ab}$$
$$e_1^*=be_2^*,\ \ e_2^*=\frac{(2a+L)V}{4a^2b-(b-1)V}, \text{ for } L\leq\frac{(b-1)V}{2ab}.$$

We now analyze the effect of AA on the effort levels exerted in equilibrium. Comparative statics are carried out by taking derivatives of equilibrium effort levels with respect to AA, that is, with respect to *L*. Equilibrium effort levels are first increasing in the size of the bonus, until a maximum is reached when $L=\frac{(b-1)V}{2ab}$, and then decreasing.[22] This implies that implementing AA provides incentives to both agents as long as the bonus is not too high, in which case the advantage is reversed and incentives reduced. That is, if the bonus helps

[22] The derivative of e_2* with respect to *L* is $\frac{de_2^*}{dL}=\frac{-V}{4a^2b-(b-1)V}\leq 0, \text{ for } L>\frac{(b-1)V}{2ab}$ and $\frac{de_2^*}{dL}=\frac{V}{4a^2b-(b-1)V}\geq 0, \text{ for } L\leq\frac{(b-1)V}{2ab}$.

reduce the asymmetry then it increases incentives, but if it is too large it can reverse the advantage and reduce incentives again.[23]

Proportional AA

Under proportional affirmative action the marginal probability of winning can be expressed as:

$$\frac{\partial F_1(e_1^* - (1+P)e_2^*)}{\partial e_1} = \begin{cases} \frac{1}{2a} + \frac{e_1^* - (1+P)e_2^*}{4a^2}, \text{ for } e_1^* - (1+P)e_2^* < 0 \\ \frac{1}{2a} - \frac{e_1^* - (1+P)e_2^*}{4a^2}, \text{ for } e_1^* - (1+P)e_2^* \geq 0 \end{cases}.$$

For agent *2* they are now slightly different:

$$\frac{\partial F_2(e_1^* - (1+P)e_2^*)}{\partial e_2} = \begin{cases} \frac{1+P}{2a} + \frac{(1+P)[e_1^* - (1+P)e_2^*]}{4a^2}, \text{ for } e_1^* - (1+P)e_2^* < 0 \\ \frac{1+P}{2a} - \frac{(1+P)[e_1^* - (1+P)e_2^*]}{4a^2}, \text{ for } e_1^* - (1+P)e_2^* \geq 0 \end{cases}.$$

Using these expressions and solving the previously presented first order conditions we find e_1* and e_2*:

$$e_1^* = \frac{b}{1+P}e_2^*, \quad e_2^* = \frac{2a(1+P)V}{4a^2b + ((1+P)^2 - b)V}, \text{ for P} > \sqrt{b} - 1$$

$$e_1^* = \frac{b}{1+P}e_2^*, \quad e_2^* = \frac{2a(1+P)V}{4a^2b + (b + (1+P)^2)V}, \text{ for P} \leq \sqrt{b} - 1 .$$

Again, doing comparative statics of equilibrium effort levels on the size of the AA policy, *P*, we find that for low levels of proportional bonus, $\text{P} \leq \sqrt{b} - 1$, the equilibrium effort levels are increasing in the size of the bonus. But if the bonus is high, $\text{P} > \sqrt{b} - 1$, increasing it further will decrease e_1* and have ambiguous effects on e_2*. [24] As before there exists a threshold value of $\text{P} = \sqrt{b} - 1$, that maximizes total effort and corresponds to the *P* that exactly levels the competition.

Summarizing then, we find that Affirmative Action should increase subjects' performance, no matter whether the compensation is lump-sum or proportional to performance, as long as the bonus reduces the asymmetry, that is, if it is not too large to

---

[23] Schotter and Weigelt (1992) report that lump sum Affirmative Action theoretically reduces effort, but this is because they only looked at the case where the compensation is too large, i.e., $L > \frac{(b-1)V}{2ab}$.

[24] For $\text{P} > \sqrt{b} - 1$, $\frac{\partial e_1^*}{\partial P} = -\frac{4ab(1+P)V^2}{(4a^2b - (b + (1+P)^2)V)^2} < 0$ and $\frac{\partial e_2^*}{\partial P} = \frac{1}{b}(1+P)\frac{\partial e_1^*}{\partial P} + \frac{1}{b}e_1^*$, which is of ambiguous sign. For $\text{P} \leq \sqrt{b} - 1$, $\frac{\partial e_1^*}{\partial P} = -\frac{4ab(1+P)V^2}{(4a^2b + (b - (1+P)^2)V)^2} > 0$, and $\frac{\partial e_2^*}{\partial P} = \frac{1}{b}(1+P)\frac{\partial e_1^*}{\partial P} + \frac{1}{b}e_1^* > 0$.

induce reversed discrimination. Notice however that since we did not know the exact size of the asymmetry in experience, the model only serves as an illustration of the possible effects different types and sizes of affirmative action policies can have. We did in any case changes these parameters in order to enrich the illustration, although we cannot fully test the model.

4. Results

4.1 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). Notice that there is high heterogeneity in the 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, a key hypothesis justifying our experimental design. It also shows that being in a more advanced grade improves performance.

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))	**38.92** (16.10)	**24.66** (15.43)	**33.27** (16.44)	**20.38** (12.80)
K	**29.94** (12.45)	**17.69** (10.70)	**43** (17.97)	**29.09** (13.43)	**35.60** (16.21)	**22.92** (13.13)
AA	**29.18** (13.74)	**19.26** (9.48)	**45.47** (12.06)	**28.08** (12.12)	**36.59** (15.29)	**24.03** (11.80)
LH	28.47 (11.89)	23.36 (9.19)	44.29 (11.86)	29.50 (14.43)	35.67 (14.19)	27 (12.73)
LL	27.47 (12.21)	19.41 (11.79)	51.38 (11.19)	26 (9.01)	37.83 (16.71)	22.56 (10.85)
PH	29.60 (12.45)	17.92 (9.05)	44.69 (11.09)	26.53 (11.15)	37.54 (14.50)	22.07 (10.84)
PL	31.23 (18.22)	17.07 (7.59)	40.27 (12.38)	29.16 (12.83)	35.47 (16.17)	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 AA measures are not implemented (NK and K). Notice 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 lower 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 0.003 for K).

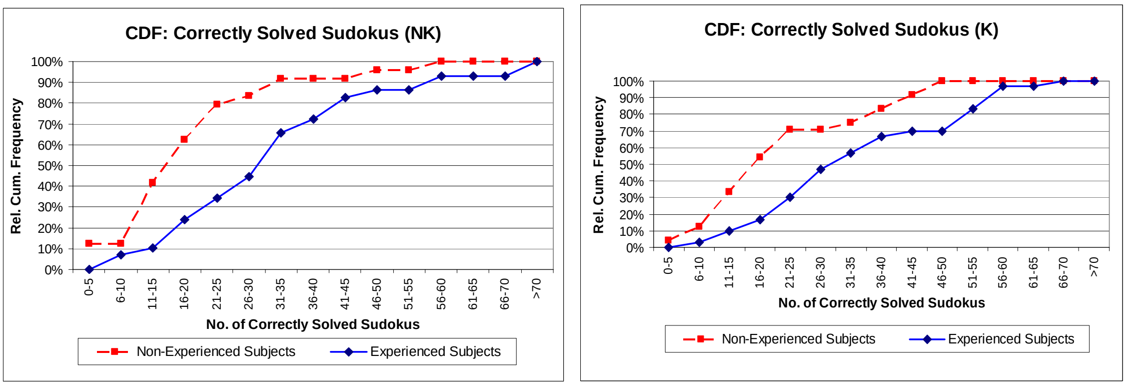

Figure 2: CDFs of the number of correct sudokus by E and NE in the NK and K treatments.

Intra-school comparisons across treatments are less clear cut. Table 3 reports the number of correctly solved sudokus for both, the unpooled data for the treatments where affirmative action is implemented (LH, LL, PH and PL), and the pooled data under the AA label. Figure 3 draws the CDFs of the number of correct sudokus for the NK, K and AA treatments in each of the two schools. Visually, the CDF of the K treatment "almost stochastically dominates" the CDF of the NK treatment in both graphs, suggesting that the provision of information on the existence of a disadvantaged group does not decrease performance, and in fact it may weakly enhance it. Similarly, the CDFs of the AA treatment also lie below the CDFs of the NK treatment and generally not above the CDF of the K treatment in both schools, suggesting that subjects facing AA policies do not decrease their performance, even with respect to the K treatment in which they are aware of the disadvantage. Mann-Whitney tests comparing the distribution of all treatments do not generally show significant differences at the standard levels, apart from the comparison of NK with AA for 6th year experienced subjects (average of 38.92 correct sudokus in NK and average of 45.47 in AA, p-value of 0.03).

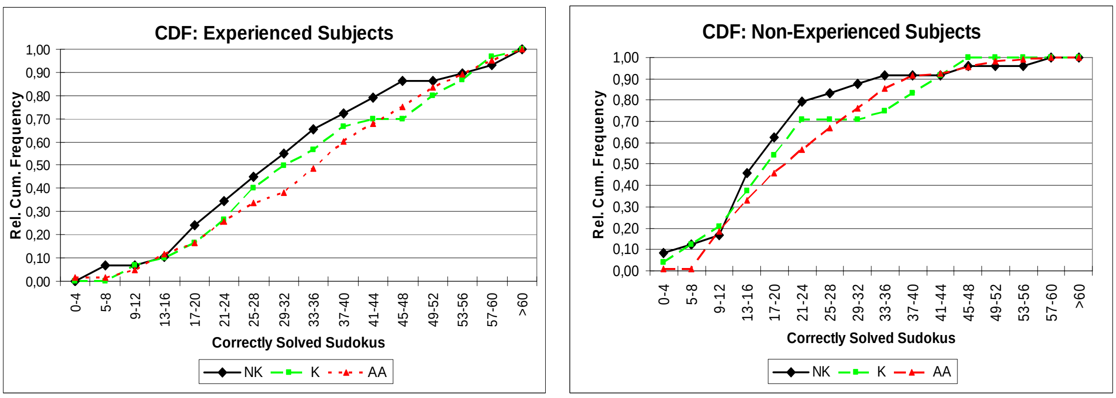

Figure 3: CDFs of the number of correct sudokus by school in the NK, K and AA treatments.

Mann-Whitney tests comparing the distribution of correctly solved sudokus in all four treatments where affirmative action is implemented (LH, LL, PH and PL) as well as a similar graphical analysis using CDFs for all four treatments do not show important differences (p-values confirm that differences are not statistically significant at the usual levels). Notice that with respect to our theoretical model this implies that none of the compensations we used, no matter how high they seemed in our pilot experiments, were high enough to reverse the asymmetry between the two schools so as to reduce performance by subjects. Additionally, notice that the theoretical model predicts similar effects to either lump-sum or proportional compensations. Therefore, in the following we perform our analysis with both pooled and upooled data for the treatments where affirmative action in implemented.

Low levels of significance can be attributed to the high variance among the subjects in a given school. But part of this variance can be attributed to individual heterogeneity that we control for in the following section.

4.2 The Effects of Information and Affirmative Action on Performance

In this section we study the effects of providing information on the ex-ante difference in experience between subjects and how performance in the tournament is affected by the implementation of AA policies, using suitable controls for individual heterogenity. We run linear regressions with the number of correctly solved sudokus as dependent variable (Table 4). Our baseline treatment is K, where subjects are aware of the existent disadvantage but no AA policy is implemented. Explanatory variables are the AA treatment dummies (both unpooled and pooled in two separate regressions) and a dummy variable for treatment NK in order to control for the effect of providing information on the asymmetry in experience. From the theoretical model presented above, asymmetries in ability to compete should affect performance, independently of their potential source. Therefore, an important variable we need to control for is unobserved "ability". We use as a proxy the number of correct sudokus solved by subjects in a "Pretest", i.e., the five minute practice rounds subjects perform before competition was introduced. Additional regressors include several other individual characteristics such as "Gender", year of schooling ("Year"), and the math grade obtained by students in the previous term ("Grade"). Since we anticipated that AA may affect subjects differently depending on their level of ability, we included an interaction term ("AA*Pretest"). 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). The regressions presented in table 4 show the results for both E and NE schools when controlling for the different AA treatments, pooled in the case of OLS (1) and (3) and unpooled in OLS(2) and (4).[25]

[25] Since "Pretest" and "Grade" may be interpreted as different but correlated measures of individual ability, we run separate regressions using just one of each variable as regressor. Qualitative results are maintained and statistical levels of significance are slightly improved.

Table 4: Correct Sudokus, Information and Affirmative Action				
	Experienced		Non-Experienced	
	OLS (1) Dep. Var: # Correct Sudokus	OLS (2) Dep. Var: # Correct Sudokus	OLS (3) Dep. Var: # Correct Sudokus	OLS (4) Dep. Var: # Correct Sudokus
Constant	-12.96 (4.43)***	-13.12 (4.42)***	5.16 (3.29)	5.22 (3.34)
NK	2.73 (2.49)	2.75 (2.49)	0.40 (2.15)	0.40 (2.19)
AA	8.31 (4.80)*	-	-1.59 (2.68)	-
AA*Pretest	-1.64 (0.96)*	-	1-45 (0.56)**	-
LH	-	11.10 (5.96)*	-	-3.83 (4.10)
LL	-	-1.66 (8.19)	-	-2.01 (4.02)
PH	-	13.38 (6.02)**	-	-0.38 (3.55)
PL	-	1.39 (7.20)	-	-0.66 (3.51)
LH*Pretest	-	-2.29 (1.24)*	-	2.12 (0.88)**
LL*Pretest	-	0.31 (1.62)	-	1.58 (1.03)
PH*Pretest	-	-2.17 (1.23)*	-	1.01 (0.87)
PL*Pretest	-	-0.58 (1.45)	-	1.12 (0.81)
Pretest (0=Min, 6=Max in E) (0=Min, 12=Max in E)	6.96 (0.81)***	6.95 (0.81)***	3.34 (0.42)***	3.34 (0.43)***
Grade (1=Worst,5=Best)	3.33 (0.72)***	3.43 (0.72)***	0.45 (0.70)	0.45 (0.71)
Year ($0=4^{th},1=6^{th}$)	12.03 (1.63)***	11.77 (1.65)***	4.14 (1.24)***	4.28 (1.28)***
Gender (0=Male,1=Female)	1.99 (1.40)	2.05 (1.41)	1.04 (1.19)	0.76 (1.25)
Adj. R^2	0.65	0.66	0.65	0.64

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

Our proxy for unobserved ability, measured by results in the practice rounds (”Pretest”), being in sixth grade instead of fourth grade (“Year”), and math grades (“Grades”) have positive and significant effects,[26] while the effect of “Gender” is not significant. Our first result relates to the effect of information on the disadvantage of NE subjects on performance. Notice that the coefficients for the NK dummy variable are not significant in any of the regressions which, together with Figure 3 allow us to conclude:

Result 1: *Knowledge of the existence of an asymmetry in experience does not decrease performance by experienced nor non-experienced subjects.*

We now focus on the effect of affirmative action on performance, starting with the experienced subjects (E). OLS (1) shows that when pooling all affirmative action treatments, the coefficient for AA has a positive and significant impact (although only at the 10% level), i.e., experienced subjects of the basis group (with low ability) solve statistically 8.315 sudokus more if they compete with subjects being favored by an affirmative action policy. However, this

[26] “Grade” is not statistically significant for NE subjects.

increase in performance AA declines the higher the ability of the experienced subject (measured by "Pretest"), since the interaction term ("AA*Pretest") is negative and significant (at the 10% level).[27] For subjects with the highest ability the incentive enhancing effects of AA are nullified or even slightly negative.[28] This result suggests that it is precisely those experienced subjects who may realize they may now not win their respect tournament, once non-experienced subjects are being helped, the ones who react more to competitive pressure. On the other hand, experienced subjects with high ability may not have any reason to react, since they may win their respective tournament anyways The unpooled analysis of the AA treatments in OLS (2) shows generally the same signs (apart from the highly non-significant LL treatment) although significance is low, which may be due to the lower sample size. In any case, the more intensive AA treatments, i.e. LH and PH, are the main contributors for the described incentive effects. We thus conclude:

Result 2: *Affirmative Action policies (weakly) enhance the performance of experienced subjects. The lower the individual ability of the experienced subject the higher is the effect.*

We now focus on the non-experienced subjects (NE). Notice that as before, the inclusion of the interaction term ("AA*Pretest") implies that a subject of the base group has low ability. For the pooled data (OLS (3)) the coefficient for AA for low ability subjects is negative although not statistically significant from zero. Hence, low ability subjects do not seem to react to AA, i.e., AA by itself does not reduce the performance of non-experienced subjects with low ability. This result is substantively different for subjects with higher abilities as signalized by the positive and significant coefficient at the 5% level for the interaction term ("AA*Pretest"): subjects with higher ability solve significantly more sudokus under AA than without, i.e., one correct sudoku more in the practice round implies 1.45 more correct sudokus in the tournament if AA is implemented.[29] We thus conclude:

---

[27] The statistical effect of AA on subjects with higher ability can therefore be calculated as "AA"+"AA*Pretest".

[28] We did not anticipate experienced subjects to solve more than 6 sudokus in their 5 minute practice rounds. Thus, experienced subjects were only provided with 6 sudokus as part of their trials. 40% of the subjects solved all 6 of them correctly. Therefore our measure of ability for experienced subjects is truncated at 6, since it includes individuals who would possibly have solved 6 or more sudokus. We thus expect that our estimated parameters are smaller and less significant (due to higher variance) in comparison to properly specified trials with a larger number of sudokus. For the subsequently run experimental sessions with non-experienced subjects, we extended the number of trial sudokus to 12. Truncating those data artificially in the same manner as for experienced subjects, we could verify our conjecture that results became slightly less significant and weaker in absolute size without altering the qualitative results.

[29] Note that non-experienced subjects could solve up to 12 sudokus in the pretest which implies that the performance enhancing effect of AA, as expressed by the interaction term "AA*Pretest", is of importance. For instance, a subject with 11 sudokus in the pretest (the highest observed value in the experiment) would statistically solve 14.4 (=-1.59+11*1.45) more sudokus in the tournament with AA than without it.

Result 3: *Affirmative Action policies do not decrease the performance of non-experienced subjects. In fact, it increases their performance the higher the ability of the non-experienced subject.*

The results obtained for the non-experienced subjects are in contrast to those for experienced subjects. While low ability subjects without experience are not affected by AA, high ability subjects without experience tend to react very positively to AA. For experienced subjects this relation is reversed. This difference in behavior is intuitive because non-experienced subjects with high ability and experienced subjects with low ability are the ones most affected by the implementation of AA. They are at the margin of competition and are most directly affected by the leveled playing field that is imposed through AA. The results for the separated AA treatments in OLS (4) are similar in direction but mainly not significant.[30] For non-experienced subjects this conclusion is also verified by analyzing the results of a similar linear regression where the dependent variable is the number of wrongly solved sudokus. Here, low ability students have significantly lower number of wrongly solved sudokus under AA. This effect declines for students with higher abilities.[31]

Given the existence of an important literature analyzing how male and female individuals react differently to competition (see Gneezy et al. (2003) and Niederle and Vesterlund (2007)), we run similar regressions controlling additionally for gender and its interaction with affirmative action and the ability proxy. Table 5 presents regressions OLS (5) and OLS (6) which are extended versions of OLS (1) and OLS (3) for experienced and non-experienced subjects respectively. They are supplemented by the interaction terms "AA*Gender", "Pretest*Gender" and "AA*Gender*Pretest".

Table 5: Correct Sudokus, Affirmative Action and Gender

	Experienced	Non-Experienced
	OLS (5) Dep. Var: # Correct Sudokus	OLS (6) Dep. Var: # Correct Sudokus
Const	10.497 (5.342)*	11.084 (3.556)***
AA	4.419 (6.513)	-5.846 (3.235)*
AA*Pretest	-0.783 (1.333)	2.888 (0.727)***
AA*Gender	8.265 (8.856)	8.318 (4.491)*
AA*Gender*Pretest	1.841 (1.916)	-2.779 (1.070)***
Pretest*Gender	1.315 (1.567)	3.262 (0.796)***
Pretest	6.425 (1.036)***	1.762 (0.554)***
Gender	-3.769	-9.484

30 Again, the reason for the low significance of the AA-treatment dummies in OLS (4) in comparison to the pooled treatment AA in OLS (3) may be due to the small sample size for each singular treatment.

31 The linear regression results for ex-ante experienced subjects are less clear and mainly insignificant. Regressions where the number of wrongly solved sudokus is the dependent variable are available upon request.

(0=Male,1=Female)	(7.068)	(3.568)***
NK	2.722 (2.502)	-0.820 (2.069)
Year	12.032 (1.643)***	3.640 (1.192)***
Grade	3.301 (0.722)***	0.283 (0.672)
Adj. R^2	0.650	0.680

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

Results for experienced subjects are not statistically significant. However, for non-experienced subjects we find significant gender effects in line with the existent literature, which may partially explain the low significance of results in regressions (3) and (4). Women with low ability do worse in competition than men with equally low ability since the coefficient for "Gender" in OLS (6) is large, negative, and significant at the 1% level. However, this effect is reversed for women with higher ability. This can be seen from the fact that the coefficient "Pretest*Gender" is significant at the 1% level and positive. We can also observe that the effect of AA is different for men and women. The coefficient for "AA" is significant at the 10% level and negative, implying that low ability non-experienced male subjects with low ability decrease their performance in the presence of affirmative action. However, since "AA*Pretest" is significant at the 1% level and positive, this effect on male is reversed as ability increases. Women's performance is enhanced under AA since the coefficient for "AA*Gender" is significant at the 10% level and positive, and it does so independently of the ability level since the significant coefficients at the 1% level on "AA*Pretest" and "AA*Pretest*Gender" almost cancel out.[32] Therefore we conclude:

Result 4: *Non-experienced women perform worse than men in competitions as long as the asymmetry in experience is not corrected through affirmative action. Once the disadvantage has been corrected, women increase their performance more than men do.*

We finally look at how expectations on winning their respective tournament are affected by affirmative action. After the tournament was finished, subjects answered a voluntary questionnaire which included questions about their individual perception of the probability of winning their respective tournament.[33] As there was no information about the identity and characteristics of the respective opponent (with the exception on ex-ante experience in the AA treatments and treatment K) we use these answers as a measure of confidence in winning. OLS (7) and (8) in table 6 regress our measure of confidence in winning on "Pretest" and the treatment dummy for affirmative action ("AA"). Understandably, we find that high ability subjects have higher confidence in wining their respective tournament as "Pretest" has a

[32] Similar regressions for separated AA treatments confirmed these results in the sense that all AA treatments had similar effects. The results of those regressions are available on request.
[33] In the relevant question 6 (see appendix) students could rank their expectation to win the tournament against their respective rival on an ordinal scale from 1 ("Sure") to 5 ("For Sure Not").

positive and significant coefficient at the 1% level. More interestingly, we find that while for experienced subjects the presence of AA does not significantly affect reported self-confidence, it significantly increases at the 5% level the confidence of non-experienced subjects. These results are consistent with experienced subjects not feeling frustrated by the introduction of affirmative action while at the same time. it correctly increases the expectations of non-experienced subjects of winning their respective tournament.

Table 6: Expected Winning Probability, Affirmative Action and Ability

	Experienced	Non-Experienced
	OLS (7) Dep. Var.: Win Prob.	OLS (8) Dep. Var.: Win Prob.
Constant	1.40 (0.214)***	1.43 (0.177)***
AA	-0.121 (0.143)	0.344 (0.155)**
Pretest	0.182 (0.042)***	0.098 (0.033)***
Adj. R^2	0.086	0.069

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

Finally, regressions OLS (9) and OLS (10) show that while self-confidence is not gender specific for non-experienced subjects, experienced women report a significantly lower confidence in winning than male subjects of the same ability at the 5% level.

Table 7: Expected Winning Probability, Affirmative Action and Gender

	Experienced	Non-Experienced
	OLS (9) Dep. Var.: Win Prob.	OLS (10) Dep. Var.: Win Prob.
Constant	1.00 (0.306) 0.000 ***	1.373 (0.230) 0.000 ***
AA	-0.238 (0.193) 0.217	0.501 (0.213) 0.020 **
Gender	0.694 (0.427) 0.106	0.160 (0.360) 0.657
AA*Gender	0.245 (0.286) 0.393	-0.345 (0.313) 0.272
Pretest	0.289 (0.062) 0.000 ***	0.106 (0.045) 0.021 **
Pretest* Gender	-0.198 (0.084) 0.020 **	-0.024 (0.068) 0.724
Adj. R^2	0.100	0.066

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

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

In the previous subsection we have shown that affirmative action has weak performance enhancing effects on both experienced and non-experienced subjects. However, a question remains regarding the efficiency of implementing such policies from an ex-post point of view. As we have seen, AA increases the probability of winning of subjects who have an inherent disadvantage. Thus, those disadvantaged subjects who win their respective tournament may do so not because they perform better than their respective advantaged competitor but because now they are given a compensation. As a result, the performance of tournament winners may be lower than the one which would have been obtained had affirmative action not been implemented. However, given that in our tournament affirmative action had performance enhancing effects, it is still possible that when comparing performance with and without affirmative action there may not be any reduction and thus, no efficiency loss. We first show that our affirmative action measures "leveled the playing field" at the aggregate level. Table 8 reports the percentage of tournaments winners from the NE school as a result of performing all possible pair-wise matches between competitors from both schools in each treatment.[34] Notice that the percentage of NE tournament winners in the 4th grade (58.29%) is higher than in the 6th grade (45.81%), suggesting that the sizes of the compensations may have been a bit too high (low) given the performance of 4th (6th) graders. Comparing the NK and K treatments with the pooled AA treatments for the whole sample we observe that NE subjects pass from representing roughly just 24% of the tournament winners when there is no affirmative action implemented (NK and K treatments) to roughly 52% when it is implemented. This difference is significant at all standard significance levels.

Table 8: Percentage of Non-Experienced Tournament Winners by Treatment and School Year

	4th Year	6th Year	Overall
NK	**25**	**23.81**	**24.42**
K	**21.27**	**27.27**	**23.94**
AA	**58.29**	**45.81**	**51.81**
LH	83.42	57.14	72.32
LL	49.51	10.49	31.84
PH	61.43	55.03	58.40
PL	40.27	53.68	45.96

We now look at the average performance of all possible tournament winners by treatment and school year. Table 9 reports these averages with standard deviations in parenthesis. We find that both separating by school year and pooling all data ("Overall"), the average number of correct sudokus solved by tournament winners is not lower in the AA treatments than in the NK. However, there exist a small negative difference when comparing the AA treatments with respect to the K treatments, i.e., when looking at the effect of affirmative action among competitors who are aware of the existence of a disadvantage.

[34] Notice that the particular match used to reward subjects in our experiment was just one random realization of this process.

Table 9: Average Correct Sudokus by All Possible Tournament Winners in Each Treatment and School Year			
	4th Year	6th Year	Overall
NK	**30** (13.541)	**42.04** (15.52)	**35.81** (15.71)
K	**31.83** (11.62)	**46.91** (13.03)	**37.75** (14.23)
AA	**29.70** (13.35)	**43.36** (12.53)	**36.53** (14.63)
LH	28.20 (11.98)	42.91 (12.58)	36.22 (14.31)
LL	29.75 (12.79)	51.09 (11.12)	38.54 (16.04)
PH	27.94 (11.00)	41.30 (12.80)	33.90 (13.56)
PL	32.56 (16.27)	41.08 (11.51)	37.36 (14.41)

Notice that since we are performing all possible matches within each treatment, the sample size increases exponentially and thus statistical tests are now more powerful. Table 10 reports the percentual change in these averages between treatments as well as Mann-Whitney tests comparing these differences. A negative sign indicates that the first treatment compared has a lower average than the second compared treatment.

Table 10: Percentage Change of the Average Correct Sudokus by Tournament Winners			
	4th Year	6th Year	Overall
NK Vs. K	-5.76**	-10.38***	-5.14**
NK Vs. AA	1.02	-3.05**	-1.95*
K Vs. AA	7.19**	8.17***	3.36

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

Notice that when comparing the NK with the AA treatment, performance is in fact a significant 1.95% higher (at the 10% level) when affirmative action is in place. Performance only decreases by a non-significant 3.36% when comparing AA with K for the overall data (although a significant loss of 7.19% for 4th year and of 8.17% for 6th year) indicating that the average performance was not lower under affirmative action.[35] Thus we conclude:

Result 5: *While affirmative action policies induced a leveled playing field, they did not do so at the expense of a loss in performance by the respective tournament winners.*

5. Conclusion

This paper contributes to the debate on the effects of AA by providing evidence on the change in performance of the participants and in the resulting pool of selected winners. It provides experimental evidence that in a competitive environment with a clear and unjustified

[35] Notice that performance is a significantly lower 7% when comparing the performance by tournament winners under AA with respect to the hypothetical situation in which affirmative action was announced (such that the performance inducing effects are present) but then not used to select tournament winners. Notice that such policy, although more efficient from a performance point of view, may be difficult to justify and not credible in the long term.

asymmetry, levelling the playing field through AA may improve all individuals' performance. Moreover it may not worsen the quality of the selected winners.

Our results thus imply that there are circumstances under which affirmative action policies do not generate a loss in efficiency, neither from an ex-ante point of view, as measured by the tournament effort, nor from an ex-post point of view, as measured by the average performance by tournament winners. Our design does not allow us however to discuss the possible long term effects of using affirmative action measures to select a pool of winners composed by a higher percentage of disadvantaged subjects. Although we have seen that affirmative action corrects the asymmetry in the short term, the disadvantage between both types of subjects may remain in the long term. Thus, future research will look at how disadvantaged subjects behave in subsequent competitions once affirmative action is removed depending on whether the asymmetry disappears in the long term or not.

The contribution of this paper is to separate the incentive effects of affirmative action from the normative issue of whether affirmative action policies should be implemented on the basis of moral or fairness justifications. Once we fully understand the effects on performance affirmative action may have and the factors affecting these effects a more informed debate could take place.

6. References

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

Calsamiglia, C., (2009). "Decentralizing Equality of Opportunity". International Economic Review 50(1).

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

Franke, J., (2007). "Does Affirmative Action Reduce Effort Incentives? A Contest Game Analysis". UAB and IAE Working Paper Series 711.07. Barcelona.

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

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

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.

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

Harrison, G., and J. List, (2004). “Field Experiments”. Journal of Economic Literature 42(4), pages 1009-1055.

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

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

Krasnokutskaya, E., and K. Seim,(2007). “Determinants of the Participation Decision in Highway Procurement Auctions”. Mimeo.

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

Myerson, R., (1981): “Optimal auction design”, Mathematics of Operations Research 6, 58-73.

Krause, K., and W. Harbaugh, (1999). “Economic Experiments that you can perform at home on your children”. Working Paper.

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

Niederle, M., C. Segal and L. Vesterlund, (2008). “How Costly is Diversity? Affirmative Action in Competitive Environments”. Mimeo.

Niederle, M., and L. Vesterlund, (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, 1999, *37*(1), pp. 7-63.

Schotter, A., and K. Weigelt, (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.

Steele, C.M., S.J. Spencer, and J. Aronson, (2002). “Contending with Group Image: The Psychology of Stereotype and Social Identity Threat” in: M. Zanna (Ed.), Advances in Experimental Social Psychology Vol. 37. Academic Press.

9. Appendix

Experimental Instructions

Below you can find a translation of the experimental instructions for one of the treatments (Experienced students in treatment PH). Other treatments' instructions were identical, although changing the size and type of compensation. Complete instructions are 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 are supposed to do. You have to complete 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 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 those 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 have less practice than you we are going to give them 1 grid more for each grid that they have solved correctly.

For example:

- If your matched participant correctly solves **12** grids, they count as 12 +12 = 24 grids. Therefore 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 scorrectly 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 by chance 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?

   Sure Probably yes I don’t know Probably not For sure 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? ____