Undergraduate Economic Review

Volume 16 | Issue 1 Article 14

2019

Effects Of Relative Position on Success in Dancing Competitions - Lessons for Discrimination in Recruitment Processes

Jan P. Ringling
University of Oxford, jpspam@outlook.de

Effects Of Relative Position on Success in Dancing Competitions - Lessons for Discrimination in Recruitment Processes

Abstract

This article argues that there exists a further reason for discrimination beside taste and statistics, based on cognitive bias in the human thought process. The order in which one appears to recruiters in the hiring process also influences the chances of being hired in a non-linear way. When the characteristics of particular groups of people correlate with their order in which they appear in the hiring process, they stand a higher or lower chance of being hired. Preliminary evidence based on the results of the United Kingdom's largest student dancing competition supports this hypothesis, but finds only a small effect.

Keywords

Labour Economics, Dancesport, Relative Heat, Discrimination

1. Introduction

The early economic literature on discrimination encompasses two distinct, albeit challenging to differentiate, types of discrimination.

Gary Becker (1971) introduces the idea of taste-based discrimination, which involves that employers associate a cost with hiring a new employee based on the employee's group association. Where this cost is positive, employers will be more reluctant, ceteris paribus, to hire an equally qualified employee at the same wage as an employee belonging to a group with a lower cost. Taste-based discrimination, in other words, is what one would refer to as prejudice unsubstantiated in facts.

Kenneth Arrow (1973) and Edmund Phelps (1972) pay more attention to statistical discrimination. Phelps describes statistical discrimination as occurring when there is discrimination against certain groups "if [an employer] believes them to be less qualified, reliable, long-term, etc. on the average […] and if the cost of gaining information about the individual applicant is excessive".[1] Statistical discrimination hence ensues in a world of imperfect or costly to obtain information but is ultimately grounded in the economic pursuit of profits, not in the arbitrariness of humans.

Darity & Mason (1998) provide an excellent summary of other sources of discrimination, some of which extend the existing theories to explain the observed persistent wage gaps, which should typically close under the neoclassical assumptions, and some of which charter new ground. For the latter, they present a theory about gatekeepers, where wage differences are persistent because an insider group is able to capture the institutions responsible for job assignment. In this case, by preventing outsiders from even obtaining the necessary qualifications for a job, wage differences remain persistent.

[1] Phelps, 1972, p. 659

The existence of persistent wage differences and hiring rate differentials between different groups, especially races and genders, is corroborated by a wide variety of studies (cf. e.g. Bertrand & Mullainathan, 2004; Schmidt, 2016; Arias, Yamada and Tejerina, 2004). All of these tend to employ some variation of the Blinder-Oaxaca-Decomposition which splits the total wage gap into a part that is explained by productivity, e.g. by incorporating evidence on educational achievements, and a part attributed to discrimination, where discrimination is usually taken to correspond to the different types presented in the previous list. However, this paper argues that these studies might miss another crucial component that might result in systemic discriminatory behaviour against certain groups without any malintent due to the combination of cognitive biases and how different groups fall into the traps of cognitive biases.

The reason for this lies in one of the weak points of the Blinder-Oaxaca-Decomposition. In this decomposition, the second component is a residual. If it does not include all relevant factors or the measuring is inaccurate, "the residual will reflect these omitted influences […] and will therefore either over- or underestimate the extent of discrimination" (Cotton, 1988). While many of the unmeasured variables such as family background might represent past discrimination, the difference to the third category "cognitive bias" represented in this paper is that the latter is entirely unrelated to discrimination.

An early example of such a cognitive bias is the Relative Age Effect. Bedard & Dhuey (2006) for example show that children whose birthday falls just after the cut-off date for school entry, making them the oldest children in their cohort, perform 2-9 percentiles better than those whose birthdays are just before the cut-off date still in 8th grade. This effect endures even later, with the study showing that the relatively oldest students of a cohort are 11.6% more likely to attend a 4-year Accredited College than their youngest counterparts.

If we are then to assume that birth dates of people belonging to different races or genders cluster differently for some reason, studies will attribute the effect of the Relative Age Effect to discriminatory efforts, e.g. on part of the teacher in the grading process, of the admissions departments in the university admissions process, or of the employer in the hiring process. The Relative Age Effect itself appears to be an unlikely candidate for a source of such bias, despite some ominous websites claiming to have discovered a pattern to the gender of a child based on maternal age and date of conception.[2]

However, even if the Relative Age Effect is unable to explain persistent differences in wage setting and hiring behaviour, this does not exclude other sources of potential bias. In particular, this paper has found a suitable candidate for a bias, which influences the chance of success in a binary setting (i.e. success or no success). This bias is based on relative timing of when a decision is made: if there is a pile of applicants to be judged, humans appear to be more lenient in judgement for people who find themselves in the middle of this pile. While the evidence this paper will present was found in a setting most conducive to finding such an effect and has yet to be transferred to the labour market, it shows that such biases do exist and they can play a role in determining the effect group affiliation may have on hiring decisions.

Please note that this paper will refer to the Relative Age Effect for illustrative purposes multiple times. The reason for this is, again, purely illustrative, and the reader should not assume that hiring processes are systematically biased against certain groups due to the Relative Age Effect until a study can undoubtedly confirm such an effect. Other sources of systematic bias also exist but make for less clarifying examples.

---

[2] E.g. https://www.goodtoknow.co.uk/family/family-news/meghan-markle-baby-gender-revealed-452222, who inaccurately predicted the gender of the first child of the Duke and Duchess of Sussex to be a girl

Discrimination in hiring processes or university admission is unfair for the person who is discriminated against, who potentially was cheated of his preferred place. Additionally, discrimination in hiring processes will also affect the wage this person receives and will hence affect the overall wage differential. Discrimination in hiring and hiring rate gaps in hiring processes translate into aggregate level wage differentials through two direct and one indirect effects.

First of all, if admission procedures to college or university fall prey to such biases, it already stacks the cards against groups that fall prey to cognitive biases, e.g. groups that are born late in the school year. The effect will successively accumulate and lead to these groups earning less. The effect of these biases will hence be indirect and affect earnings and the wage gap through their effect on education.

Secondly, if somebody is less likely to be hired, their period of unemployment will likely be longer. This is simple probability – the chance of receiving at least one offer in 5 applications is 96.87% if the probability of receiving an offer in any single round equals 50%, but it is 98.98% if the probability of receiving an offer in any single round corresponds to 60%[3]. In other words, a 10 percentage points improvement in the probability of being hired increases the chances of receiving an offer in 5 applications by 2.1 percentage points. Although Arulampam (2001) finds no effects of unemployment duration that are statistically significant, by his own admission, his study might be influenced by a shift of reliance of unemployed people to sickness benefits in the period under investigation.[4] A study by Gregory & Jukes (2001) does find that the duration of unemployment permanently reduces an individual's wage. Gregg & Tominey (2005) find in an Instrumental Variable-setup that every month of youth

[3] $1-(1-50\%)^5$ and $1-(1-60\%)^5$ respectively, to be precise

[4] Arulampalam, 2001, F601

unemployment reduces "a male's conditional wage at age 33" by 1.8%".[5] In brief, the longer somebody takes to be hired, the smaller their wage will be, and, if e.g. cognitive biases systematically affect somebody's chances to be hired, they will hence also affect their wage.

Thirdly, there is likely to be a selection process in which high-paying jobs tend to be barred from people with a lower probability of being hired. Ceteris paribus, most people and particularly every species of the homo oeconomicus will prefer a high-paying job to a low-paying job. Those who have a high probability of being hired in any job, too, will prefer the high-paying contracts and choose them should they be offered these, leaving the lower-paying jobs to the people with less chance of being hired in a new hiring round. People with a lower probability of being hired anywhere are hence left to choose from jobs that are all lower-paying than the jobs held by people with a high probability of being hired, creating a wage gap.

To conclude this section, even if a bias, cognitive or other, strictly only affects the probability of being hired and not the wage, it will still retain an indirect effect on the wage and thus create wage gaps between groups who systematically differ in their ability to trigger the bias.

This paper will proceed as follows. In section 2, this paper will formalise the ideas presented above in a model, which relates hiring decisions to cognitive biases, and these cognitive biases to group affiliation. Section 3 introduces a data set compiled from the results of the Inter-Varsity Dance Competition, the United Kingdom's largest Latin and Ballroom student dancing competition held annually in Blackpool's Winter Gardens. This data set then serves as basis for the analysis of Section 4, which shows that, based on several indicators of success, success in the competition is indeed partially explained by the relative sub-group the

[5] Gregg & Tominey, 2005, p. 505

competition's hosts have assigned a competing couple. Section 5 concludes by providing avenues for further research, in particular to observe whether the effect found in this conducive environment can also be found in the real world.

2. Model

The model this paper proposes explains the decision to hire somebody in two steps. The first step models the vulnerability of a person to trigger one of the described biases that are distinct from discrimination and productivity. The second step relates the manifestation of a trigger to the effect of the bias. For example, in the context of the Relative Age Effect, the first step would explain how late a child is born based upon individual characteristics such as gender and race; the second step explains what effect being born late has on the likelihood of going to university.

When these two steps are eventually combined, one receives the full equation relating hiring decisions to possible biases. It should be noted that these equations only consider the direct effect the bias has on hiring decisions, and not any indirect effects, e.g. by prolonging a spell of unemployment and hence further depreciating an individual's perceived skillset.

2.1 Explaining the triggering of a bias

This paper argues that the manifestation of systematic bias of a group can be explained simply with this equation:

$$B_G = f(\alpha_B + X_G\beta + \varepsilon) \quad (1),$$

where B_G is the corresponding manifestation of bias, β is a vector of coefficients corresponding to the effect an increase in a group's attribute has, X_G is a vector of the mean value of attributes of a group, α_B is the intercept, and ε is the error term. In this equation, neither α nor β has a subscript 'G', meaning that the difference in the manifestation of bias in different groups is solely due to different values of the attributes explaining biases, and not a form of discrimination based on group affiliation. The function *f(.)* relates the intercept and effect of the attributes to the bias *B* and can take several forms. For the Relative Age Effect, *f(.)* might imply a linear combination, with *B* being a continuous variable. However, theoretically *f(.)* could also take another form such as the sigmoid function, if B turns out to be a binary variable. In the context of this paper's analysis of positional advantage, *f(.)* is assumed to be a linear model in which the intercept and the error term equal 0, all entries of X_G equal 1, and the coefficients equal the relative position of the respective couple.

To illustrate the function in another context focussing on the Relative Age Effect, *B* could correspond to the time somebody is born in the relative year, and *X* would be the determinants. If the conception of children of highly successful parents only takes place during the Christmas Break because at other times the parents are too busy working, their children's arrival nine months later coincides with the cutoff-date of school entry in many countries. The Independent Variable of 'having highly successful parents' would differ between different socioeconomic groups, and systematically affect the bias B of relative age.

This paper leaves open the question of what determines how certain biases manifest themselves or fail to manifest themselves for certain groups, instead focussing on the next step.

2.2 Explaining the effect of bias on hiring decisions

The second step then relates the manifestation of the bias to the outcome of the hiring decision through the following equation:

$$P(H_G = 1 | \alpha, P_G, D_G, B_G) = g(\alpha + P_G \beta_{P,G} + D_G \beta_{D,G} + B_G \beta_{B,G} + \eta) \quad (2)$$

In this context, H_G is a binary variable corresponding to the mean of people being hired from a group, $\beta_{P,G}$ is a vector of coefficients detailing the effect an increase in a productivity parameter displayed by the vector P_G has, D_G captures any discriminatory practices against a certain group, and $\beta_{B,G}$ is a vector of coefficients detailing the effect of certain biases, whose values are entailed in the vector B_G. α is the intercept, which will vary according to the job, with applications to professorships likely having a lower α, symbolising the smaller chance of being hired. η is the error term. The function $g(.)$ can take a variety of forms, but in this case, a probit- or logit-form would perhaps suit the situation best due to the binary nature of somebody being hired.

Besides the effect biases can have on hiring, which is the focus of this paper, this equation also captures the effects of productivity and discriminatory variables commonly found in the economic literature. Thus, it serves as an extension to the standard economic models describing the grounds on which hiring decisions are based. It captures any indirect effects biases have on the hiring decisions not through a separate term; another regression would be required to estimate the effect of a bias on the other variables like productivity. However, the presented equation still manages to capture the direct effects of bias.

To illustrate the meaning of this equation, consider the case of a wealthy family's kid. The excellent education and specific discriminatory influences will be assimilated in the respective vectors, and their familiar effects will enter the decision whether or not to hire the individual. Take again the previous illustration of the Relative Age Effect, in which children of wealthy parents are born just after the school cut-off date, making them the relatively oldest children in their school. Analyses that do not take into account the Relative Age Effect – if it were to play the role described – would be prone to take cognitive biases that systematically affect a group as evidence for discrimination for example in grading, particularly those that ascribe any wage differentials unexplained by productivity factors to discrimination.

In this paper, the equation (2) reduces to the form

$$P(H_G = 1 | \alpha, B_G) = g(\alpha + B_G \beta_{B,G} + \eta) \ (3),$$

with productivity and discrimination incorporated in the error term η. The reason for this is discussed in detail in the analysis of the dataset but is twofold. Firstly, as this paper will show later, the bias I evaluate is independent of any productivity – in the context of dancing effectively the dancing skills – and from discrimination. It even fulfils the strict condition of orthogonality. Secondly, data on certain variables are also not freely available or not even collected. However, the lack of data on these variables is a minor concern, as there is valid reason to conjecture orthogonality.

2.3 Putting the two equations together

Using simple substitution, compiling equations (1) and (2) into one equation yields

$$P(H_G = 1 | \alpha, \alpha_{B,} P_G, D_G, X_G) = g(\alpha + P_G \beta_{P,G} + D_G \beta_{D,G} + f(\alpha_B + X_G \beta + \varepsilon)\beta_{B,G} + \eta)$$

(4).

If one assumes *f(.)* to be a simple linear combination, the equation reduces to

$$\begin{aligned} P(H_G = 1 | \alpha, \alpha_B, P_G, D_G, X_G) &= g(\alpha + P_G \beta_{P,G} + D_G \beta_{D,G} + (\alpha_B + X_G \beta + \varepsilon)\beta_{B,G} + \eta) \\ &= g(\alpha + P_G \beta_{P,G} + D_G \beta_{D,G} + \alpha_B \beta_{B,G} + (X_G \beta)\beta_{B,G} + \varepsilon\beta_{B,G} + \eta) \\ &= g(\alpha_G + P_G \beta_{P,G} + D_G \beta_{D,G} + (X_G \beta)\beta_{B,G} + u) \end{aligned}$$

(5)

Verbalised, this means that the chance of a group to be hired for a job is a function of their perceived productivity, discrimination, biases that systematically play against the group, and other, uncorrelated factors.

As a result, the effect of factors contributing to the manifestation of biases on the hiring decision in this model depends both on β – how strongly the factors actually contribute to the manifestation – and $\beta_{B,G}$ – the direct effect of bias on the hiring decision. Since the sub-groups are assigned directly and solely by the hosts of the competition in a way that is indistinguishable from pure randomness, equation (5) reduces even further to:

$$P(H_G = 1 | \alpha_G, X_G) = g(\alpha_G + (X_G \beta)\beta_{B,G} + u) \quad (6)$$

In a Blinder-Oaxaca decomposition based on the underlying equation (5), in which P is taken as ability (or Marginal Product), both D_G and X_G would be caught with the last term (Bauer & Sinning, 2008), so that the part ascribed to discrimination would be inflated.

3. Description of the data and variables

The subsequent analysis is based on data derived from the Inter-Varsity Dance Competition of the year 2019. This competition is hosted annually by the Inter-Varsity Dance Association, "the representative and governing body for student dancesport in the United Kingdom" (IVDA Website, 2019), with the 2019 edition being the 57th instalment. According to the competition guide and the uploaded data, more than 1000 competitors competed "in a record 21 different open events" (IVDA Committee, 2019).

The results of this competition were all published on the online platform Scrutelle,[6] as is common in Dancing. Scrutelle also hosts results from other dancing competitions, for example the Blackpool Dance Festival, perhaps the largest and famous of all dancing competitions, which incidentally is hosted at the same place as the Inter-Varsity Dance Competition: The Blackpool Winter Gardens.

The Inter-Varsity Dance Competition hosts 21 different events. Events are held for each of the two traditional dancing styles Ballroom and Latin, for the team match, for the "Offbeat Competition", the Salsa formation competition and for Rock'n'Roll and Salsa. The events for Latin and Ballroom are, again, split according to the level of ability into four different events each: Beginner, Novice, Intermediate, and Advanced. Since a beginner event consists of only one dance,

[6] http://www.scrutelle.co.uk/cgi-bin/db1.pl?form=1&apple=estelle/2019/190223_ivda

each style entails two beginner events, one for each of the two basic dances. Furthermore, there are Ex-student competitions on Novice level and an open level involving all five dances of the respective style Latin or Ballroom. Rock'n'Roll is split into Acrobatic Rock'n'Roll, incorporating figures that involve a lot more time in the air for the follower, and Non-Acrobatic Rock'n'Roll, which is less sensational but also less injurious.

The dataset incorporates the results from all regular rounds except the finals from all the Latin and Ballroom events, as well as the team match and the ex-student novice competition. Enumerated, these events are the Beginner, Novice, Intermediate, and Advanced events, and the Ex-student Novice event, for Latin and for Ballroom, as well as the team match. Hence, the largest and most traditional events have been included, for which students practice year-round, and in which students compete regularly and sincerely.

The data is split into two datasets. The first dataset describes the performance of couples in every round, while the second entails information about the performance of competitors for each event.

The first dataset counts 3048 observations. It also includes the variable of whether a couple makes it to the next round.[7] The decision that a couple will make it to the next round is strikingly similar to the decision a hiring committee makes. The decision of who makes it to the next round in dancing should focus solely on the quality of dancing, while it should only focus on the quality of the applicant in hiring. Just like in the hiring process, there might exist some discrimination based

[7] The first round of the Team Match constitutes a special case for Recall, because a couple that does not receive a Recall is not eliminated from the competition. Instead, those who are not recalled enter another tournament, competing against all those that were also not recalled in the first round. In this analysis I treat couples who are not recalled in the first round of the Team Match as if they are fully eliminated, to ensure comparability of the recalls between Team Match and other events.

on gender,[8] but including an estimate of discrimination in the following analysis is not necessary, since discrimination appears to be orthogonal to the variable of interest – the sub-group. Furthermore, a large number of couples involved and the necessity of quick decisions – perhaps more quickly than in the job hiring process, since a judicator in dancing has only ca. 90 seconds to choose a heat's couples – introduces another similarity of judicator selection in dancing to the hiring process especially in competitive industries.

Other observed variables in this dataset are the Competitor Number of a dance couple (assigned only for this competition), the Heat (or sub-group), the round in which an observation took place, and the number of points a couple received in this round and event. The Total Points is simply the sum of the number of recall marks of a couple by all judicators across all dances in a round.

Unfortunately, the author of this paper did not receive the requested information about heat affiliation from the IVDA Committee. Instead, the paper reconstructs the supposed heat affiliation of a person based on how many people can reasonably be expected to dance on the dance floor, and compares this with the number of people in a particular round. A description of the employed algorithm has been included in the Appendix. Recognising potential shortcomings with this method, the paper formulates additional models that corroborate the results. Furthermore, the standardisation applied to the heat should also reduce the effect any inaccuracy might have on the analysis.

The Heat (sub-group) an observation was in, and Total number of Points a couple received in a round have both been standardised. The standardisation of the variable Heat into Standardized Heat was accomplished by rescaling the variable

[8] One possible candidate for discrimination in recent years focuses on same-sex couples, usually a female dancing with another female, since males tend to be in short supply in dancesport. Not too long ago, the British Dance Council for example considered banning same-sex couples from its competitions altogether (Topping, 2014).

into a 0-1 interval, where the lowest observed Heat in any round receives a value of 0, and the highest observed heat receives a value of 1.[9] This standardisation ensures both comparability between events and rounds with a different number of heats and should eliminate potential statistical biases in the analysis introduced by the reconstruction of Heat. Although there might be measurement error, the effect of this measurement error should simply bias the coefficients in the analyses incorporating Standardized Heat towards 0 (Stock & Watson, 2015). The coefficient estimates presented here hence represent a lower boundary, which is already found to be statistically significant.

The described standardisation is the most appropriate for the following analysis since the hypothesis asks for the effect of the relative position (the relative Heat). Therefore, it is crucial that the first heat of every round and the last heat of every round receive the same values throughout, and that the heats clustered in the middle also receive similar values. The presented standardisation fulfils this, while other common ways of standardisation, e.g. the standardisation to a Standard Normal Distribution would not fulfil these criteria.

The standardisation of the Total Number of Points in a round was straightforward and achieved by calculating the average Total Points across the number of dances in a round.[10] The number of judicators did not vary between rounds and thus did not need to be standardised.

---

[9] The formula used for this is as follows:

$$x_{\text{Stand}} = \frac{x - x_{\min}}{x_{\max} - x_{\min}} \;,$$

where x is the respective observation, x_{min} is the smallest observed heat number (usually corresponding to 1), and x_{max} is the largest observed heat number. Events whose minimum and maximum heat number are equal (i.e. which consist of only 1 heat and hence are undefined) are also excluded from observation in the regression analyses because they entail no middle heat, even though the middle heats are the target of the analyses.

[10] The sum of points of a beginner couple is hence divided by 1, since each Beginners events only involves one dance, the sum of points of a novice couple is divided by 2, since Novice events involve 2 dances which are scored together, etc.

Table 1: Dataset Single Round p. Competitor

Statistic	N	Mean	St. Dev.	Min	Pctl(25)	Pctl(75)	Max
Competitor Number	3,047	277.40	167.69	1.00	127.50	423.00	572.00
Heat	3,048	3.91	2.70	1	2	6	12
Standardized Heat	2,866	0.49	0.35	0.00	0.14	0.80	1.00
Recall?	3,048	0.63	0.48	0	0	1	1
Round	3,048	2.12	1.23	1	1	3	6
Total Points	3,047	6.32	4.69	0.00	4.00	7.00	32.00
Number of Heats in the round	3,048	6.90	3.23	1	4	9	12
Average Points	3,047	4.05	1.89	0.00	3.00	5.58	8.00

The second dataset describes the competitors of every event themselves and has 1,242 observations. It features the Competitor Number of every competitor, their final placing based on the rounds they made and the total points they received in their last round, and the Last Round they made before the couple had to drop out of the competition (an indicator for how successful they were). The Competitor Number and the Last Round have also been standardised for each event since different events have a different number of competitors and hence also differ in how many rounds they make until the finals. The standardisation is the same as that applied to the Heat in the other dataset and is based on the same criteria.

Table 2: Dataset Competitors

Statistic	N	Mean	St. Dev.	Min	Pctl(25)	Pctl(75)	Max
Competitor Number	1,242	269.52	172.08	1	107	420.8	572
Standardized Competitor Number	1,242	0.47	0.31	0.00	0.19	0.73	1.00
Placing	1,242	61.50	43.49	1	24	94	159
Last Round	1,242	2.57	1.45	1	1	3	7
Standardized Last Round	1,242	0.31	0.28	0	0	0.5	1

While both the first and second dataset is built from the same competition, naturally the dataset showing observations about the competitors is smaller than the dataset showing observations about a competitor in each round. After all, every competitor participates in at least one, and usually more than one round (see that the average round a couple made before they dropped out is 2.57).

4. Analysis and Discussion

4.1 Hypothesis

In this section, the paper presents the evidence that the decision to let a couple advance to the next round is biased by the relative Heat a couple is placed in. The bias is hypothesized to be non-linear so that the model corresponds to

$H_G = g(\alpha_G + \beta_1 X_G + \beta_2 X_G^2 + u)$, with $\beta_1 X_G + \beta_2 X_G^2 = \beta_{B,G} X_G$ when compared to (6),

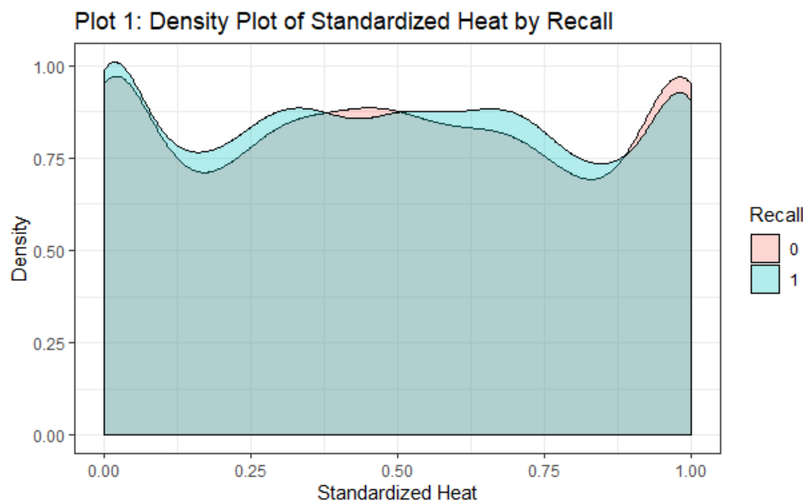
With *g(.)* being the integral of the standard normal distribution. The Hypothesis states that couples featuring in the middle Heats have a better chance of progressing to the next round than couples that are presented to the judicators either very early or very late. This is because of cognitive bias: Judicators will hold back with marks in the first few Heats because they are aware that there are plenty more couples coming, and they only have a limited number of marks to dispense. Towards the end of the heats, judicators have already distributed a significant number of marks and might not even be able to award every deserving couple with a mark. The couples dancing in the middle Heats should then have the best chances of receiving a surplus of marks enabling them to move forward to the next round.

To corroborate these findings, the paper presents two additional models. The first of these additional models provides a link between the number of Rounds a couple made in an event and its Competitor Number. The intuition behind this is that couples, who have a Competitor Number that is relatively in the middle compared to the other Competitor Numbers in an event, will tend to end up in middle heats. Since the hypothesis states that couples in middle heats will perform slightly superiorly, couples with a middle Competitor Number should make more rounds than couples with peripheral Competitor Numbers.

The last model presented in this paper links the Total number of Points a couple received to the Heat in which the couple is placed. Since the hypothesis relies on the cognitive biases of individual judges, the link should also manifest itself when aggregating the ratings of the judges.

4.2 Descriptive Statistics

A first look at Plot 1, a density plot of Heats, with the density shown by whether a couple is recalled or not, gives a first indication of support for the hypothesis.

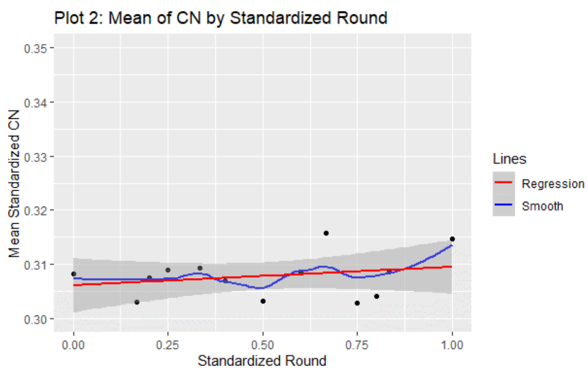


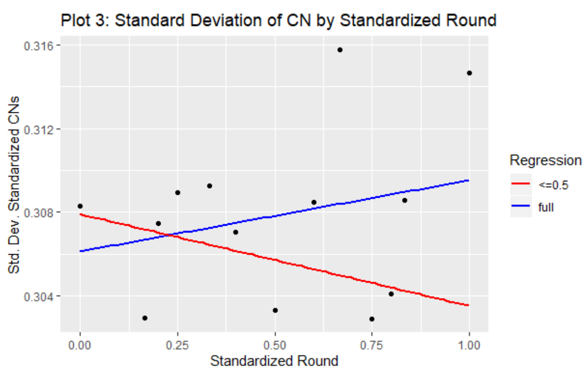
The plot shows that recalled couples tend to cluster more in the middle, while not recalled couples cluster towards the borders. The plot excludes any observations of rounds in which there were 2 or fewer Heats since in these cases there is no middle Heat.

Looking at differences between means is not particularly helpful in this context. The reason for this lies in the hypothesised non-linear relation between Recall and Standardized Heat, in which the mean as a measure of linear relation becomes quite useless. However, since the hypothesis states that recalled couples will cluster around a particular value, which is close to 0.5, the standard deviation as a measure of dispersion is apt to examine the relationship. If the standard deviation of recalled couples is lower than the standard deviation of all couples, this will support the hypothesis, because it indicates a clustering of Recalls around some point.

I found that indeed the standard deviation in Standardized Heat of couples receiving Recalls is marginally lower (33.63%) than the standard deviation of all couples (33.96%) so that the standard deviation of all couples is ca. 1% higher than that of couples receiving Recalls.[11]

A similar result can be found for the relationship between Competitor Number and Rounds.





These graphs plot the mean and standard deviation of all couples who danced in a particular round, whether they made it to the next round or not. While the mean of Competitor Number stays relatively constant until the end, the standard

[11] As the plot, this value was calculated with Observations in which Observations with a number of Heats of 2 or lower being excluded.

deviation of Competitor Numbers increases over the observed rounds, i.e. the standard deviation of Competitor Numbers in the finals is higher than the standard deviation in the middle.

While this seemingly speaks against the hypothesis, it is important to note that several additional effects are manifesting themselves towards the final rounds. Firstly, the number of heats continuously reduces with progressing rounds, making it easier for judicators to make expert judgements across multiple Heats. Hence, the importance of being in a middle Heat should reduce with the total number of Heats. Secondly, judicators have had some rounds to become familiar with couples and their quality of dancing. They might remember the boy with the sparkling bowtie, whose Waltz was so superb, or the girl with a feather in her hair, whose Tango caused hair to stand on end. They also remember just about in which Heat the couple might appear – towards the beginning, the middle, or the end. This is a further effect that might mitigate the importance of Heat position.

Indeed, if the sample on which the regression is based is restricted to Standardized Round <= 0.5, the slope reverses and becomes negative, indicating that the standard deviation of Competitor Numbers decreases during the first few rounds, centring on the mean.

4.3 Regression Analysis

The American psychologist Dr Wayne Dyer reportedly once said "When you dance, your purpose is not to get to a certain place on the floor. It's to enjoy each step along the way" (Seibel, 2015). As a dancer myself, I might add to this quote "But enjoyment is twice as sweet when you make the next round".

Making the next round, which I have compared to the decision of being hired, is the most important variable of all those under observations. While it is undoubtedly vital to reach a high number of Total Points, the points are non-transferable, so once one reaches enough points for the next round, any more points have a marginal benefit of 0. Conversely, the total number of points does not matter much if one cannot progress to the next round. The main result of the regression hence focusses on the chances of making the next round, modelled by a Probit-regression.

Before turning to regression analysis however, it is crucial to see whether the distribution of dancers across the Competition Numbers and Heats is random, or whether there are any patterns. If all the good dancers happen to end up in the middle Heats, the results might otherwise prove spurious.

Unfortunately, it proves hard to observe the ability of a dancer. Ability might be inferred by how far a dancer came, or how many points the dancer received from judicators, but these indicators are obviously endogenous to the analysis.

Comparing the results from the Inter-Varsity Dance Competition to other, previously held, competitions also proves challenging for three reasons: Firstly, there is no necessity for couples to dance with the same partner at different competitions. As the old proverb “It takes two to Tango” says, the success in dancing is dependent on both dancers in a couple and is more like an interaction between the dancers than a simple average. Secondly, since the Inter-Varsity Dance Competition is the most important competition, many dancers miss previous competitions and only attend this one, so that there is no information available about a significant share of dancers. Sparse information could hence make detection of a tiny effect untenable. Thirdly, and most significantly, all evaluations in dancesport are relative to the other couples attending a competition. The same couple might receive a score of 0 or full score, depending on whether they compete with the British champions, or with adolescent beginners. The analysis would then require turning the results of previous competition into an absolute rating, alike to the Elo score used in chess, which goes beyond the scope of this article.

Instead, this article evaluates the spread of dancers based on their university affiliation. This has two reasons. If universities are spread randomly, it appears likely that their dancers are spread randomly as well. Furthermore, university affiliation can be taken as a proxy for the achievement of a dancer. Since 1996, the winners of the A-Team in the Team Match[12] were either Oxford, Cambridge, or Imperial with exceptions in 1998 (University of London), 2003 (Oxford, Cambridge & Cardiff ranked jointly first), and 2005 (Bristol). The winner of the overall champions[13] since 1996 has been Oxford, Cambridge, and Imperial as well, with one exception in 1998 (University of London). Coaching opportunities vary between universities, with Oxford featuring a dedicated coaching team with a head coach and several auxiliary coaches according to their website (OUDC Website, 2019), which many other universities cannot feature.

The paper conducts two tests to test for the independence of University and Position.

Firstly, the homogeneity of variance is tested using the Fligner-Killeen Test. A standard test like a likelihood ratio test should not be applied in the case under study, because it relies on the assumption of normality (Conover, Johnson & Johnson, 1981). However, Competitor Numbers and Heat assignment are necessarily uniformly distributed, thus violating a crucial assumption. The Fligner-Killeen Test is a "non-parametric test of the null hypothesis that samples have been drawn from populations with a common distribution, with the alternative being that the distributions have the same mean (or median) but different scales (and thus different variances)" (Upton & Cook, 2014). Testing the assumption that Competitor Number and Heat in the first round of an event are randomly distributed across universities yields a non-significant p-value[14], indicating that we can assume homogeneity of variance across universities.

[12] In the Team Match, every University sends up to 6 teams consisting of 4 couples, each of which competes against every other team including the ones from its own university
[13] The best performing university throughout the whole competition
[14] Distribution of Competitor Number across Universities: $p = 0.3145$, Distribution of First-Round Heats across Universities: $p = 0.216$

Secondly, differences in means are tested using the Kruskal-Wallis Rank Sum Test, another test which does not rely on the assumption of normality. It is an extension of the Mann-Whitney test to more than two groups (Upton & Cook, 2014; Hollander, Wolf & Chicken, 2014). This test, too, yields non-significant p-values.[15]

In the light of this evidence, it is reasonable to assume that dancers are distributed randomly across the Competitor Numbers and Heats so that any systematic difference between middle and outer heats can be assigned to the relative position, and not to differences in quality.

To model the non-linear effect of Heat on Recall, I employ a probit model in a quadratic specification, as previously presented. The models all report McFadden's Adjusted Pseudo-R^2, which adjusts McFadden's Pseudo-R^2 with a punishment term when including too many Independent Variables (citation needed). The table also reports Log-Likelihood and the Akaike Information Criterium as further measures of fit. The results are displayed in Table 3.

The table reports four models which differ in their sample size and Independent Variables.

Model (1) includes observations from all rounds with at least 3 different heats and incorporates Dummy variables for all universities. The coefficients on both terms based on Standardized Heat are statistically significant, although the coefficient on the linear term is statistically significant only at the 10% level. Thus, it appears that the hypothesis is substantiated.

For all subsequent models, the sample has been restricted further to exclude the Team Match. This should make the detection of positional effects easier. The reason is that while all other events are split according to ability, in the Team Match

[15] Distribution of Competitor Number across Universities: $p = 0.6488$, Distribution of First-Round Heats across Universities: $p = 0.5578$

Beginners with less than half a year of experience compete with Advanced couples, with more than five years of experience.

Model (2) is based on the restricted sample and also incorporates university dummies, but is only linear. The linear coefficient on Heat is not statistically significant; its Standard Error is higher than its estimated value.

Model (3) is a quadratic specification but excludes university dummies. Without accounting for ability through this proxy, the level of significance has dropped, with the coefficient on the linear term being no longer statistically significant at conventional levels, and the coefficient on the quadratic term being significant just at the 10% level. The difference between the sizes of the coefficients in Model (1) and (3) is smaller than the standard error in either model, so might be just due to chance. However, it is still remarkable that an effect in the hypothesised direction, albeit very small, can be found without including any further controls.

Finally, model (4) is the quadratic specification on the restricted sample and includes dummy controls for university. The coefficients on both Heat terms are statistically significant at the 5% level, and both feature the predicted signs.

The differences in model fit are small, with the quadratic specifications with university dummies of Models (1) and (4) performing the best, and the worst-performing model does not include any controls for ability. This is reassuring since it implies that there is a broad scope for talent and ability beside the proven positional advantage. The negligible difference in coefficient size between the quadratic models including university dummies and the model not including university dummies provides further backing to the assumption that ability, at last when measured by the university, is spread randomly throughout all Heats.

Table 3: Probability of Recall

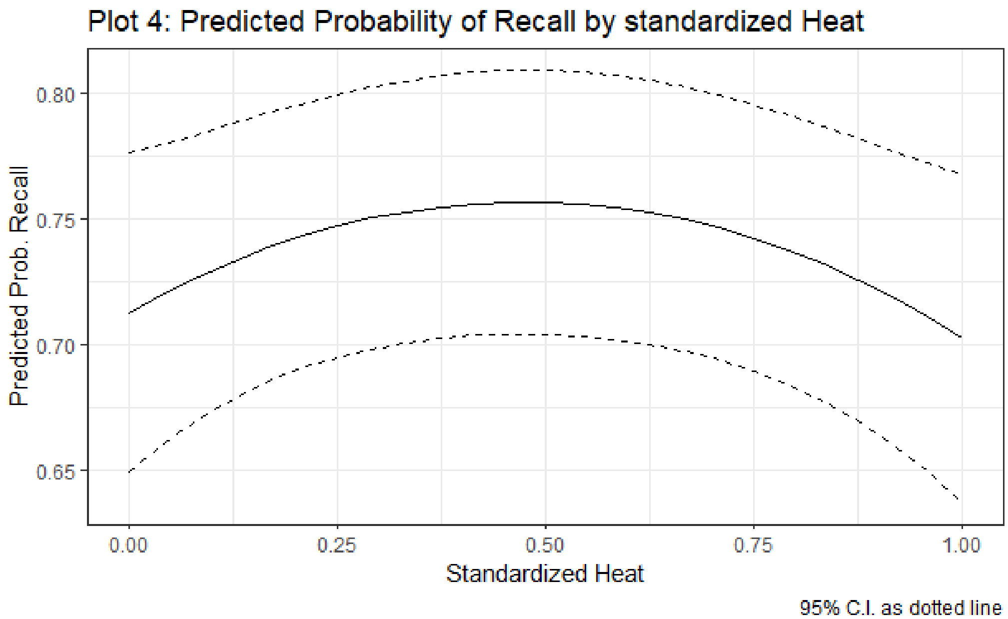
	Recall			
	probit			
	(1)	(2)	(3)	(4)
Constant	-0.265	-0.165	0.351***	-0.238
	(0.241)	(0.242)	(0.056)	(0.244)
Standard Heat	0.505*	-0.027	0.393	0.567**
	(0.268)	(0.077)	(0.265)	(0.273)
(Standard Heat)^2	-0.546**		-0.423*	-0.597**
	(0.259)		(0.256)	(0.264)
University Dummies	Yes	Yes	No	Yes
McFadden's Pseudo-R^2	0.01	0.007	-0.001	0.008
N	2,658	2,561	2,561	2,561
Log Likelihood	-1,668.595	-1,611.845	-1,652.695	-1,609.265
Akaike Inf. Crit.	3,417.189	3,285.689	3,311.390	3,282.530

Notes: ***Significant at the 1 percent level.
**Significant at the 5 percent level.
*Significant at the 10 percent level.

A likelihood-ratio test comparing models (2) and (4) to each other and to a probit-model regressing Recall on a Constant yield that Model (4) provides a statistically significantly better fit.[16] The results support the theory that the effect of Standardized Heat is non-linear.

Finally, Plot 4 displays the predicted probability of being recalled for each value of Standardized Heat, displaying the non-linear relationship. Depending on whether one is in a central or peripheral heat, the predicted probability varies ca. 6%, i.e. by switching from a peripheral to a central heat one can increase one's chances of making the next round by almost 8.6%!

[16] The improvement of fit between the Model (4) and the constant is statistically significant at the 0.1% level; the improvement of fit between models (2) and (4) is statistically significant at the 5% level.



The next regression is a linear regression with the number of rounds a couple made regressed on its Competitor Number. The logic here, as outlined above, is that a middle Competitor Number ensures a couple is placed in a middle heat most of the time so that the couple's number of rounds should be higher than a couple with a very high or very low Competitor Number.

The results of these do not provide definite proof for the hypothesis, but neither do they contradict them. None of the coefficients on Competitor Number are statistically significant. Reported standard errors are robust. Both of the regressions were conducted on the full set of Competitors, including the Team Match. Excluding observations gathered from the Team Match mildly increases the level of significance, but the standard errors remain higher than their coefficients (not displayed here). It should be noted, though, that in Model (2), which includes dummies for the university, the signs correspond to expectation.

The question arises why there is no statistically significant result to be found here, contrary to the previous set of regressions. There are three different, but not mutually exclusive, explanations. Firstly, the sample size is smaller than in the previous regression. The sample size might hence simply be too small to detect a very minor and statistically significant effect. Secondly, the sample includes rounds and events in which the number of couples would be too small for the explained bias to manifest itself. The coefficient estimates, however, display the average effect over all rounds and are hence biased downward. It was easily to filter these observations out of the sample in the first set of regressions, but they are still included here. Thirdly, Competitor Number is only a proxy for the Heat a couple is placed in every round, so it can only inaccurately measure the real coefficient. If an Independent Variable can only be observed with some measurement error, its coefficient is biased towards 0, as is the case here. All reported standard errors are robust.

Table 4: Number of Rounds

	Standard Rounds	
	OLS	
	(1)	(2)
Constant	0.324***	0.147***
	(0.070)	(0.019)
Standard CN	-0.013	0.011
	(0.096)	(0.096)
(Standard CN)²	-0.039	-0.060
	(0.097)	(0.098)
University Dummies	No	Yes
N	1,242	1,242
R^2	0.003	0.104
Adjusted R^2	0.001	0.075
Residual Std. Error	0.284 (df = 1239)	0.274 (df = 1202)
F Statistic	1.868 (df = 2; 1239)	3.581*** (df = 39; 1202)
Notes:		***Significant at the 1 percent level.
		**Significant at the 5 percent level.
		*Significant at the 10 percent level.

I tested two restrictions in Model (2) with an F-Test. Firstly, I tested the null hypothesis that the coefficient on the quadratic term is equal to 0. I was unable to reject the null hypothesis.[17] Secondly, I tested the null hypothesis that the coefficients on both the linear and the quadratic term are equal to 0. As with the first null, I was unable to reject the null hypothesis,[18] albeit at a much lower p-value.

The final set of models regress the Total Number of Points a couple receives in any round on the Heat the couple was in. The sample-set for the first two models, once again, only includes observations from rounds with at least 3 Heats, so that

[17] $p = 0.534$
[18] $p = 0.159$

there is always at least one interior Heat. The third model also excludes the Team Match for the reason mentioned earlier.

The first model, which does not include dummies for the university, shows no statistically significant coefficients, although the signs point in the correct direction.

The second model includes university dummies and yields statistically significant coefficients on Heat. While the coefficient on the quadratic term is significant at the 5% level, the coefficient on the linear term is only significant at the 10% level. The difference between the values of the coefficient of Model (1) and (2) is less than one standard error in either model.

Finally, the third model excludes observations from the Team Match but retains dummy variables for university. The level of statistical significance on both terms for Heat increases to the 5% level. The difference in coefficients between Model (3) and either Model (1) or (2) is less than the standard error on the coefficients in any model. The increased size of the coefficients on Heat when excluding Team Match observations lends support to the prior speculation that the heterogeneity of ability in the Team Match supersedes the positional effect.

Table 5: Total Points

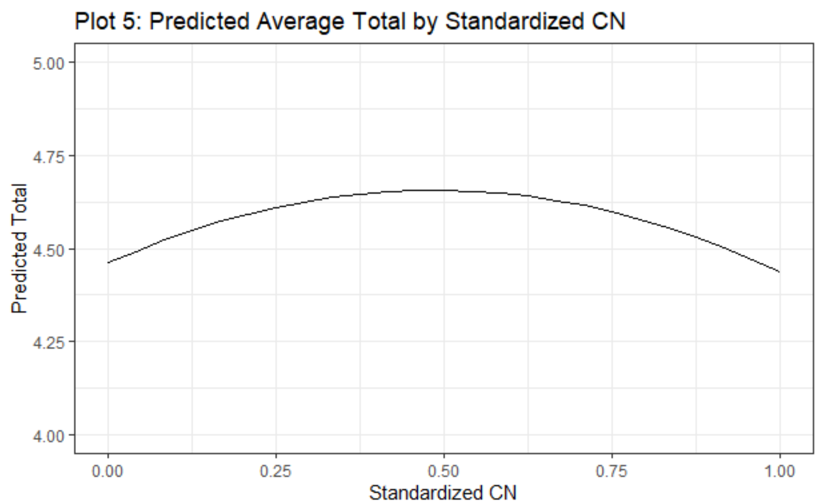
	Averaged total points		
	OLS		
	(1)	(2)	(3)
Constant	4.068***	2.851***	2.876***
	(0.081)	(0.352)	(0.356)
Standard Heat	0.451	0.676*	0.805**
	(0.384)	(0.382)	(0.387)
(Standard Heat)²	-0.515	-0.748**	-0.830**
	(0.370)	(0.368)	(0.373)
University Dummies	No	Yes	Yes
N	2,658	2,658	2,561
R^2	0.001	0.065	0.058
Adjusted R^2	0.0001	0.051	0.046
Residual Std. Error	1.900 (df = 2655)	1.851 (df = 2618)	1.843 (df = 2529)
F Statistic	1.125 (df = 2; 2655)	4.633*** (df = 39; 2618)	5.016*** (df = 31; 2529)

Notes: ***Significant at the 1 percent level.
**Significant at the 5 percent level.
*Significant at the 10 percent level.

It is clear that including University dummies as a proxy to ability provides a much better fit. This is again reassuring since it indicates that much more variation in the Total Number of Points can be explained by ability than by positional advantage. The third model provides a slightly worse fit than the second model in terms of both R^2 and adj. R^2. A reason for this could lie in the role of the Team Match, in which University affiliation is a much greater factor of success, as described above, and hence University dummies become better predictors of success.

An F-Test conducted on both the linear and quadratic coefficient of Heat could reject the null Hypothesis of both coefficients being equal to 0 at the 10%

level.[19] An F-Test testing for the coefficient on the quadratic term being equal to 0 could reject the null hypothesis at the 5% level.[20]



Plot 5 display the predicted non-linear relationship between Heat and Average Total Points, tentatively confirming the initial hypothesis that judicators tend to award more marks in the middle rounds. The range of Predicted Total points is 0.22.

To conclude this section, there is a variety of evidence that supports the hypothesis that the relative position of a couple in the field of competitors has a slight influence on this couple's chance in the competition. In other words, a cognitive bias directly influences success when success is defined by outside sources, be it HR managers or judicators. The influence of the cognitive bias derived from relative position appears particularly strong in cases in which there are many applicants of comparable quality, in which it is hard to rank applicants. Conversely, the influence somewhat diminishes when there is a clear rank-order between applicants, the number of applicants is overall low, and the time available to examine candidates is high.

[19] $p = 8.3\%$
[20] $p = 2.63\%$

5. Directions for further research and concluding remarks

The analysis in this paper has shown that there is a certain role for cognitive biases to affect the success of people. However, for this cognitive bias to affect the aggregate, a group of people must be systematically affected by this bias. The question is then why and how are certain groups of people more affected by a bias than other groups.

I have found two possible candidates warranting closer examination.

The first candidate is UK university admissions. In the United Kingdom, most undergraduate applications are submitted to universities via a portal called "UCAS". In 2019, applications via UCAS could be submitted from 4 September (UCAS, 2019). Then there are two deadlines for applications: Oxford, Cambridge, and some courses like medicine, dentistry, and veterinary medicine require an application by 15 October, the standard application deadline for all other courses is 15 January 2020. Since there is only one single application possible, which is then sent to multiple universities, people applying to Oxbridge will submit their application earlier than other students. If the composition of people applying to Oxbridge is different, this might convert into a significant advantage for some groups solely due to positional advantage and no conscious or subconscious discrimination.

Indeed, an analysis by the head of the Oxford college Lady Margaret Hall Alan Rusbridger (2018) reports that from all British pupils with the requisite grades to attend Oxford, 14% of underprivileged pupils apply to Oxford, while 25% of all state-school pupils apply, and 37% of independent school pupils do. A contributing factor to the overrepresentation of independent school pupils in general might hence be because they happen to apply at a more convenient point of time.

This is not to say that this is the only obstacle in the way of underprivileged students on their way into Higher Education – far from it, the same analysis lists a whole number of other obstacles – but it does point to easily implementable policy recommendations which somewhat level the playing field. Such an example could be mandating universities to refrain from considering applications on a rolling basis, and instead only consider applications after the deadline and in a randomised order. This way, the positional advantage of public school students would be somewhat mitigated.

A second candidate could be hiring processes. Applications to internships in areas like Consulting and Investment Banking in the United Kingdom are due around November – barely a month after the university term has started. Students who are aware of these deadlines likely can submit their application with sufficient time, while students who are not familiar with these areas submit their application very late. The unfamiliar students are more likely to be from disadvantaged backgrounds, and thus face an additional obstacle in being hired into well-paying jobs. The policy recommendations here are similar to the case of admissions described above. An additional potential policy consists of offering year-round applications, which avoids a high cluster of application processing at any one point.

Future research into these matters should be two-fold. Firstly, research should look into the impact of other biases that might systematically affect the chances of a member of a specific group. Secondly, once found, these biases should be assessed on the likelihood that they impact certain groups more than others. Policies can then be proposed to mitigate the effect of those biases that systematically impact a group's chance of success.

Appendix

1. Description of the process used to derive the heat affiliation of a couple

The exact Heat a couple was in is not reported by the hosts of the Inter-Varsity Dance competition. The main criteria for Heat assignment are equality and size. This means that couples are split as equal as possible between Heats and that there are not too many couples on the dance floor at the same time due to space constraints. I have hence reconstructed the heat in the following way:

I have first sought to reconstruct the number of Heats in a round before assigning a couple their respective Heat. To do this, I calculated the number of dancers who participate in the respective round of an event. Afterwards, I estimated the maximum number of dancers who viably can be on the floor at the same time under the space constraints. For the Empress Ballroom, this is 18 couples, for the Pavilion Theatre this is 12 couples. To receive a rough estimate of the required number of Heats, I then divided the number of dancers in the respective round by the viable number of dancers who can be on the floor at the same time (i.e. in the resp. Heat). Following this, I divided the number of dancers by the rough number of heats and rounded up the result to receive an estimate of how many dancers are on the floor at the same time. The purpose of rounding up was to ensure that the number of dancers on the dance floor at the same time never exceeds the maximum viable number of dancers on the floor at the same time

For example, if a round consists of 42 couples in the Pavilion Theatre, the approximate number of dancers per heat would be 11 dancers in 4 heats:

$$\text{Approx. No. Heats} = \frac{\text{No. Dancers}}{\text{Max. No. Dancers p. Heat}} = \frac{42}{12} = 3.5 \approx 4$$

$$\text{Approx. No. Dancers p. Heat} = \frac{\text{No. Dancers}}{\text{Approx. No. Heats}} = \frac{42}{4} = 10.5 \approx 11$$

Of course, 11 * 4 dancers equals 44 total dancers and not 42, so there must be some Heats that do not include the full number of dancers. To calculate the precise number of dancers in every single Heat, I then employed the following algorithm:

I set the variable *rem* (short for *remainder of dancers*) to be equal to *No. Dancers*. If the division of *No. Dancers* by *(Approx. No. Dancers p. Heat – 1)* was equal to 0, I set the number of couples in the Heat equal to *(Approx. No. Dancers p. Heat* – 1*)* and subtracted *(Approx. No. Dancers p. Heat – 1)* from *rem*. If *rem* was not divisible by *(Approx. No. Dancers p. Heat – 1)*, I set the number of couples in the Heat equal to *Approx. No. Dancers p. Heat* and subtracted *Approx. No. Dancers p. Heat* from *rem*. I then repeated this process with the new value of *rem* until *rem* is equal to 0, i.e. all dancers are distributed to a Heat. The process is displayed schematically in the following table for the above's example:

Heat	rem	rem % (Approx. No. Dancers p. Heat – 1)	Assigned No. Dancers	New rem
1	42	2 (!= 0)	11	31
2	31	1 (!= 0)	11	20
3	20	0 (== 0)	10	10
4	10	0 (== 0)	10	0

For the case that the number of dancers in the round was perfectly divisible by the maximum number of dancers, this also yielded an equal distribution across the heats, with each heat featuring the maximum number of dancers. To illustrate, consider the case of 54 dancers in the Empress Ballroom:

Heat	rem	rem % (Approx. No. Dancers p. Heat – 1)	Assigned No. Dancers	New rem
1	54	3 (!= 0)	18	36
2	36	2 (!= 0)	18	18
3	18	1 (== 0)	18	0

The assignment of couples to Heats then was simple: I went through all Competitor Numbers in the Round from smallest to highest and continued to assign them the respective Heat until that Heat was full, afterwards repeating this process with the next Heat.

Reference list

Arias, O., et al. (2004). "Education, family background and racial earnings inequality in Brazil." International Journal of Manpower 25(3/4): 355-374.

Arrow, K. (1971). The Theory of Discrimination. Working Papers, Princeton University.

Arulampalam, W. (2001). "Is Unemployment really Scarring? Effects of Unemployment Experiences on Wages." The Economic Journal 111(475): F585-F606.

Bauer, T. K. and M. Sinning (2008). "An extension of the Blinder-Oaxaca decomposition to nonlinear models." AStA Advances in Statistical Analysis 92(2): 197-206.

Becker, G. (1971). The economics of discrimination. Chicago, Illinois, Chicago University Press.

Bedard, K. and E. Dhuey (2006). "The Persistence of Early Childhood Maturity: International Evidence of Long-Run Age." The Quarterly Journal of Economics 121(4): 1437-1472.

Bertrand, M. and S. Mullainathan (2004). "Are Emily and Greg More Employable than Lakisha and Jamal? A Field Experiment on Labour Market Discrimination." The American Economic Review 94(4): 991-1013.

Conover, W. J., et al. (1981). "A Comparative Study of Tests for Homogeneity of Variances, with Applications to the Outer Continental Shelf Bidding Data." Technometrics 23(4): 351-361.

Cook, G. U. I. (2014). A Dictionary of Statistics. Oxford, Oxford University Press.

Cotton, J. (1988). "On the Decomposition of Wage Differentials." The Review of Economics and Statistics 70(2): 236-243.

Darity, W. A. and P. L. Mason (1998). "Evidence on Discrimination in Employment: Codes of Color, Codes of Gender." The Journal of Economic Perspectives 12(2).

Economic Theory Blog (2016). ROBUST STANDARD ERRORS IN R. economictheoryblog.com.

Fox, J., et al. (2019). Companion to Applied Regression. CRAN.

Gregg, P. and E. Tominey (2005). "The wage scar from male youth unemployment." Labour Economics 12: 487-509.

Gregory, M. and R. Jukes (2001). "Unemployment and Subsequent Earnings: Estimating Scarring among British Men 1984-94." The Economic Journal 111(475): F607-F625.

Hlavac, M. (2018). stargazer: Well-Formatted Regression and Summary Statistics Tables. . https://CRAN.R-project.org/package=stargazer.

Hollander, M., et al. (2014). Nonparametric statistical methods. Hoboken, New Jersey, Wiley.

Hothorn, T., et al. (2019). Testing Linear Regression Models. CRAN.

Inter-Varsity Dance Association (IVDA) (2019). "Homepage." Retrieved 04 September, 2019, from http://www.universitydancesport.com.

IVDA Committee (2019). The 57th Inter Varsity Dance Competition. I.-V. D. A. (IVDA).

Lang, D. T. and CRAN Team (2019). General Network (HTTP/FTP/...) Client Interface for R. CRAN.

Leeper, T. J. (2018). Marginal Effects for Model Objects. CRAN.

McKinney, W. and others (2010). Data structures for statistical computing in python. Proceedings of the 9th Python in Science Conference. 445: 51-56.

Oliphant, T. E. (2006). A guide to NumPy, Trelgol Publishing USA.

Oxford University Dancesport Club (2019). "OUDC - Coaches." Retrieved 06 September, 2019, from http://www.oudancesport.co.uk/coaches/.

Phelps, E. S. (1972). "The Statistical Theory of Racism and Sexism." The American Economic Review 62(4): 659-661.

Rusbridger, A. (2018). "If Oxford Shrugs." Retrieved 07 September, 2019, from https://www.prospectmagazine.co.uk/magazine/if-oxford-shrugs-alan-rusbridger-admissions-lmh.

Schmidt, J. (2016). Entgeltgleichheit – Die gesamtwirtschaftliche Perspektive – Welche Ursachen hat der Gender Pay Gap? IW Kurzberichte, Institut der deutschen Wirtschaft Köln.

Seibel, S. (2015). "Dr. Wayne Dyer Lives On." Retrieved 19 August, 2019, from https://www.huffpost.com/entry/dr-wayne-dyer-lives-on_b_8080952.

Signorell, A. (2019). Tools for Descriptive Statistics. CRAN.

Stock, J. H. and M. W. Watson (2015). Introduction to Econometrics. Harlow, Pearson Education Limited.

Topping, A. (2014, 10 July 2014). "British Dance Council bids to ban same-sex couples from the ballroom." The Guardian. Retrieved 05 September, 2019, from https://www.theguardian.com/society/2014/jul/10/british-dance-council-same-sex-couples-ballroom-ban.

UCAS (2019). "UCAS Undergraduate: when to apply." Retrieved 09 September, 2019.

Wickham, H., et al. (2019). Create Elegant Data Visualisations Using the Grammar of Graphics. CRAN.