Games and Discrimination:

Lessons From the Weakest Link

Kate Antonovics[*], **Peter Arcidiacono**[†], **Randall Walsh**[‡§]

January , 2003

Abstract

Empirically determining whether wage differentials arise because of discrimination is extremely difficult, and distinguishing between different theories of discrimination is harder still. This paper exploits a number of unique features of a high-stakes television game show to determine which contestants discriminate and why. In the show, contestants take turns answering a series of trivia questions, and, at the end of each round of questions, one contestant is voted off by the other players in the round. Our results suggest no evidence of discriminatory voting patterns by males against females or by whites against blacks. However, somewhat surprisingly, we find that in the early rounds of the game women appear to discriminate against men. We test three competing theories for the voting behavior of women: preference-based discrimination, statistical discrimination and strategic discrimination. In doing so, we highlight the types of experimental designs that could be used to distinguish between these theories. Only preference-based discrimination is consistent with the voting patterns.

Key Words: Discrimination, Experiments, Games.

JEL: J7, C9, C7

[*]Department of Economics, University of California at San Diego. Email: kantonov@ucsd.edu.

[†]Department of Economics, Duke University Email: psarcidi@econ.duke.edu.

[‡]Department of Economics, University of Colorado at Boulder. Email: walshr@colorado.edu.

[§]We would like to thank Nora Gordon, Roger Gordon, Tom Nechyba and Joel Watson for their helpful comments and advice concerning this paper. We also thank Patrick Dickinson and Pamela Maine for research assistance.

1 Introduction

The earnings of both women and blacks have consistently lagged behind those of white men. However, empirically determining whether these differences arise because of discrimination is extremely difficult, and distinguishing between the various theories of discrimination is harder still. As a result, researchers may need to look outside of labor markets to develop a better understanding of whether and why individuals discriminate. This paper builds on the emerging experimental literature on discrimination and explores the use of a high stakes game environment to reveal patterns of discrimination. To this end, we use data from the television game show *The Weakest Link* to determine whether contestants discriminate on the basis of race and gender and, if so, which theory of discrimination best explains their behavior.

The Weakest Link is a game in which players act cooperatively to build a pot of money by correctly answering a series of "general knowledge" questions. At the end of each round of questioning, the players vote to remove one player ("The Weakest Link") from the game. At the end of the game, the two remaining players compete head to head to determine the sole winner of the accumulated pot of money. There are a number of unique features of the game that allow us to determine whether and why contestants discriminate on the basis of race or gender. Additionally, because data are collected from two slightly different versions of the show, a daily show and a weekly show, the veracity of our results can be tested against two independent samples.

The principle difficulty in establishing the presence of discrimination is that it is almost impossible to determine whether unequal outcomes arise because of discrimination or because of unobservable race and gender differences in productivity. However, a unique feature of *The Weakest Link* is that we are able to observe the same explicit measure of individual ability–the percentage of questions answered correctly[1]–that is observed by the contestants. Our ability to observe a large fraction of the information that the contestants have about one another limits the extent of the omitted variables bias problem that plagues discrimination

[1] As discussed below, the game also involves individual players' decisions about "banking" money currently in the pot vs. accepting the risk of losing this money to facilitate a potentially larger reward. These decisions provide an additional dimension of information regarding player performance.

research. However, even if there remain important unobserved characteristics, the structure of the game allows us to control for these omitted factors by using information on the voting patterns of the other contestants. In particular, if we are trying to examine the factors that determine whether player A votes against player B, we can control for the proportion of other contestants who vote against B. In this way, we can account for other factors that are not directly included in our data but that are relevant to the players' voting decisions.

The structure of the game also allows us to distinguish between various theories of racial discrimination. Within economics, the two leading theories of racial discrimination are preference-based discrimination and statistical discrimination. In models of preference-based discrimination, employers act as if there is some cost associated with hiring workers from a particular group (Becker, 1957). The result is that, in equilibrium, workers from these groups are paid lower wages than workers from other groups. In contrast, models of statistical discrimination explain discrimination as a rational response to uncertainty in labor markets. Typically in these models, workers differ by some observable characteristic (i.e. the color of their skin) and can be classified into groups based on that characteristic. Because firms cannot perfectly observe worker productivity, they base their decisions on prior beliefs about the productivity of workers in different groups. Thus, discrimination can arise either because of group differences in average productivity or in the quality of information that firms have about workers (See, for example, Phelps(1972), Arrow (1973), Aigner and Cain (1977), Lundberg and Startz (1983), and Coate and Loury (1993)).

In a strategic environment such as *The Weakest Link*, we also need to consider a third type of discrimination: what we term "strategic discrimination." The notion here is that in a strategic setting race and gender may serve as focal points for collusive behavior. For example, Holm (2000) studies a one-shot coordination game similar to the Battle of the Sexes. He finds that players appear to use gender as a focal point for coordinating actions. While the game that Holm studies is very different from the game in *The Weakest Link*, we still wish to rule out strategic discrimination as an explanation for disparate outcomes by race and gender. A major contribution of this paper is to highlight various methods of distinguishing between the three theories of racial discrimination discussed above.

To look for evidence of discriminatory behavior, we evaluate the voting behavior of indi-

vidual players using a series of conditional logit models. A major advantage of the conditional logit model relative to analysis of the total votes cast against a particular group is that it allows us to consider interactions between the characteristics of the individuals doing the voting and the characteristics of those receiving the votes. Thus, for example, one can examine whether men are more likely to vote against women *and* whether women are more likely to vote against men. The results reveal no evidence of discrimination against either women or blacks. However, we consistently find that women are substantially more likely to vote against men in the early rounds of the game, even after controlling for a broad set of performance measures. That is, based on these findings, it appears that women discriminate against men (and in favor of women) in their voting decisions.[2]

Which of the three theories of discrimination best explains the voting behavior of women? We are able to show that neither statistical discrimination nor strategic discrimination explains this pattern. Instead, the evidence seems broadly consistent with women simply preferring to play the game with other women.[3]

[2]As is discussed below we believe that the earliest rounds are the easiest to interpret because in these rounds players have a clear incentive to vote off the weakest players. In later rounds, strategic concerns associated with the head-to-head competition between the last two remaining players make voting incentives much less clear.

[3]Simultaneous to our paper, Levitt (2003) also examines discrimination in *The Weakest Link*. Our approaches are very different. Levitt focuses on aggregate discrimination (total votes cast against a particular group) rather than whether discrimination against a given group depends on who is doing the voting. For instance, if blacks vote against whites and whites vote against blacks, then these patterns could offset one another and would not be detected using Levitt's approach. In addition, Levitt's results differ from our own because of fundamental differences in our assumptions about the strategic incentives in the game and in the way we treat the data. In particular, Levitt's methodolody for distinguishing between taste-based discrimination and statistical discrimination relies on two assumptions. First, voting incentives must switch as game progresses, so that contestants first want to vote off weak players and later want to vote off strong players. However, given the structure of the game, it is not immediately obvious that the voting incentives will truly reverse. Indeed, we find that although the incentives to vote off the weakest player diminish as the game progresses, at no point is the strongest player ever more likely to be voted off than the weakest player. Second, Levitt states that taste discrimination should be the same across all rounds of the game. However, the implicit cost of taste-based discrimination rises as the show progresses because one's probability of winning the game is higher in later rounds. Hence, discriminatory outcomes due to taste discrimination should diminish as the game progresses. In addition, Levitt's treatment of the data differs from our own. Levitt is able to increase the precision of his estimates by pooling data across the weekly and the daily shows and

The remainder of the paper is organized as follows. Section 2 discusses the relevant background literature and outlines the problems with identifying and distinguishing different types of discrimination. Section 3 describes the details of the game and the data that we have collected. Section 4 shows how the game show data can be used to distinguish between preference-based, statistical, and strategic discrimination. Section 5 shows that males and whites do not discriminate, but that females discriminate against males in the early rounds. Section 6 tests which of the theories of discrimination best explains female voting behavior and Section 7 concludes.

2 Previous Work on the Nature and Extent of Racial Discrimination

It is not surprising that there exists an enormous literature on racial and gender discrimination in the labor market. Here we divide these studies into two groups: those that test for evidence of discrimination and those that attempt to distinguish between different theories of discrimination.

A large number of the studies that look for evidence of discrimination use basic regression techniques to control for observable differences between the groups. However, since there may be unobservable factors that make blacks and women less productive than white men, the race and gender differentials uncovered in these studies are not necessarily due to discrimination. Second, it is not always clear which observable characteristics should be controlled for in these regressions. For example, it is not clear whether experience should be controlled for in wage regressions since blacks may have lower levels of experience than whites partly because of discrimination. Thus, while these regression techniques provide interesting descriptions of the data, they do not provide solid evidence of discrimination.

A second strand of the literature uses audit studies to look for evidence of racial discrimination (i.e. Cross et al. (1990), Turner et. al (1991) and Neumark et al. (1996)).

across rounds. However, given that there are substantial differences in the amount of prize money available, number of rounds per game and the likely sample of contestants in the weekly and daily shows, we chose not to pool the data in this manner. Levitt's advantage over our own work is that he has a larger sample of daily shows and he analyzes discrimination against the elderly and Hispanics.

The majority of these studies find substantial evidence of discrimination against women and blacks in both the labor and housing markets. However, as Heckman and Siegelman (1992) and Heckman (1998) point out, these studies may fail to overcome the omitted variables bias that plagues standard regression techniques. In particular, the implicit assumption of audit studies is that the auditors are identical apart from their race and gender. However, the researcher who selects the auditors may not know or may not be able to observe all of the characteristics that are relevant to employers. Thus, if there is any relationship between race and gender and these unobserved components, then the differences in outcomes across these groups may not truly represent discrimination. As a means of circumventing this problem, Bertrand and Mullanaithan (2002) conduct a field experiment in which they send out a series of resumes to employers in the Chicago area that are essentially identical apart from the name at the top of the page.[4] They find that the response rates on resumes with white sounding names are 50 percent higher than those with African American names. One potential limitation of their study is that the African American names they use (i.e. Lakisha and Jamal) may connote socioeconomic status or other factors in addition to race.

Goldin and Rouse (1997) present one of the cleanest tests for discrimination. They examine whether the adoption of "screens" that hide the identity of auditioning musicians from judges has been responsible for the increase in the proportion of females hired at orchestras. Overall their evidence is consistent with discrimination against female musicians, and indicates that the use of the screen has been responsible for a large fraction of the increase in the number of women hired. However, this study, like many other studies that look for evidence of discrimination, does not reveal why people discriminate.

Detecting discrimination is difficult. Things become even more complicated when trying to empirically distinguish between various theories of discrimination, and there are only a handful of studies that even attempt to do so. Among the theories that attempt to test for preference based discrimination, two of the most well-known are Ashenfelter and Hannan (1986) and Kahn and Sherer (1988). Ashenfelter and Hannan examine the relationship between market competition in the banking industry and the relative employment levels of

[4]Neumark et al. also partially address this critique by emphasizing the use of resumes in the pre-interview stage.

women. According to the theory, taste-based discriminatory behavior should be less likely to occur in highly competitive markets, and Ashenfelter and Hannan's findings are consistent with this prediction; they find that there is a strong positive relationship between measures of market competition and the employment levels of women. Kahn and Sherer look for evidence of racial discrimination against basketball players in the NBA. They note that even after controlling for a wide range of player performance statistics, black players still earn twenty percent less than white players. Kahn and Sherer show that this coincides with lower attendance for basketball games as the number of black players increase, arguing for customer discrimination. However, the lower attendance may be due to customers valuing diversity as there are very few non-black players in professional basketball.[5]

Even fewer studies test for evidence of statistical discrimination, and those that do offer conflicting evidence. For example both Oettinger (1996) and Altonji and Pierret (2001) build models of statistical discrimination in which employers slowly learn about worker quality. In order to test the predictions of their models, both papers examine data from the National Longitudinal Survey of Youth. However, because their models have opposite predictions, their interpretations of the data are also at odds with one another; Oettinger interprets the fact the black-white wage gap grows with experience as evidence of statistical discrimination, while Altonji and Pierret interpret this trend as evidence against statistical discrimination.

Also related to our paper is a small but growing literature on how men and women play games (See, for example, Mason, Phillips and Reddington (1991), Holm (2000) and Gneezy, Neiderle, and Rustichini (2003)). However, very few of these experimental papers explicitly test for discrimination. One exception is Fershtman and Gneezy (2001) who analyze discrimination among Ashkenazic and Eastern Jews. They find discrimination against Eastern Jews in 'trust' games by both Ashkenazic and Eastern Jews. Conducting a series of experiments, the Fershtman and Gneezy argue that this result is due to incorrect expectations regarding the 'trustworthiness' of Eastern Jews.

[5] Arcidiacono and Vigdor (2002) test the diversity hypothesis using television ratings for football games as function of the race of the quarterback where black quarterbacks are in the minority. They find that ratings are substantially higher when there is a black quarterback, suggesting support for the diversity hypothesis.

3 Data and Rules of the Game

We use data collected from recordings of the nationally televised game show *The Weakest Link*. There are two versions of the show, an hour long weekly show and a half-hour daily show, with both versions following the same general structure. After excluding celebrity episodes where the contestants play for charity, our data consist of 28 weekly shows and 75 daily shows.[6] Each show is divided into a series of timed rounds, with the number of rounds corresponding to the number of players: eight rounds in the weekly show and six rounds in the daily show. Within each round, players are sequentially asked to answer general trivia questions where correct answers translate to an increase in the prize money. The first correct answer is worth $1,000 in the weekly show and $250 in the daily show. After a correct answer, a player can may choose to 'bank' the money for the team. If the player banks, the next correct answer is again worth $1,000 in the weekly show and $250 in the daily show. Should the player decide not to 'bank', the amount of money added to the pot following a correct answer increases. However, failure to answer a question correctly leads to the loss of any unbanked money for that round. A successive chain of eight (six) correct answers with no intermittent 'banks' leads to an $125,000 ($12,500) *increase* in the pot. Money banked from each round is accumulated into a team bank.

After each round, each player votes independently as to which player he would like to remove from the show, and the player who receives the most votes must leave the game. In the event of a tie, the 'strongest link' chooses which player to remove from the subset of players who received the most votes. The strongest link is the player who answers the highest percentage of his or her questions correctly.[7] Once the field of players is reduced to two (this occurs in round 7 of the weekly show and round 5 of the daily show), these two players first accumulate prize money in the same fashion as in the earlier rounds. Then, the two players directly compete against each other with the winner taking all the money in the team bank.[8]

[6] Some data is missing for particular rounds due to broadcast interruptions.

[7] Should there be a tie for the 'strongest link,' amount of money banked is used as the tie-breaker.

[8] In the final round, five (three) questions are asked of each contestant in the weekly (daily) show with the winner being determined by who answers the most questions correctly. All other contestants leave with nothing.

4 Discrimination and Strategic Behavior in the Weakest Link

The *Weakest Link* provides a unique environment in which to distinguish between various theories of discrimination. Playing the game well not only involves answering questions correctly, but also making astute inferences about the other player's ability to play the game. Since we observe players' decisions about who to vote off, we can determine whether race and gender are relevant factors in voting decisions even after controlling for each player's performance in the game.

Analysis of contestant behavior is complicated by the game's evolving strategic environment. In the early stages of the game, when the contestants act cooperatively to build the pot of prize money, there is a clear incentive to vote off weak players. However, voting incentives may shift as the game progresses because of the structure of the final round. As discussed above, in the final round, the two remaining players first build the pot of prize money as they do in the earlier rounds of the game. However, they then face one another in a head-to-head competition to determine the winner of the entire game. Thus, we would expect that concern about the head-to-head competition would create an incentive to vote off strong contestants in the rounds leading up to the final round. However, it is unclear whether the incentive to vote off strong players in later rounds will *outweigh* the incentive to vote off the weak players since even in the final round the two remaining players cooperatively build the prize money.[9] Due to this ambiguity in voting incentives, we base the majority of our analysis on the early rounds of the game, where incentives are more clear-cut. However, it is important to keep these shifting incentives in mind when considering how the different theories of discrimination affect voting behavior.

The following section presents a brief discussion of the expected voting behavior under the assumptions of a) statistical discrimination, b) preference-based discrimination, and c) strategic discrimination.

[9] Indeed, as mentioned earlier, we find that although the probability of voting off the weakest link diminishes as the game progresses, players are always more likely to vote off the weakest link than the strongest link.

4.1 Statistical Discrimination

In models of statistical discrimination uncertainty is the root cause of discrimination. In this situation, individuals may use race or gender as a proxy for unobserved ability. However, uncertainty alone is not sufficient to generate unequal outcomes. One of two additional factors must be present; either there must exist group differences in average ability or in the accuracy of the information available to others.

Because contestants are uncertain of one another's ability to play the game, uncertainty is likely to play an important role in *The Weakest Link*. As a result, contestants may rationally use race or gender as a proxy for underlying ability when attempting to assess the skills of their fellow contestants. In this setting, discriminatory voting patterns can arise either because of real group-level differences in the contestant's ability or because contestants have better information about the abilities of some players relative to others. These two sources of statistical discrimination (differences in ability vs. differences in the quality of information) have different predictions about voting patterns as the game progresses. This provides us with a method for uncovering the mechanism underlying that discrimination. That is, we can tell whether players are statistically discriminating because of perceived differences in players' average abilities or because of asymmetries in the accuracy of information that players receive about one another.

In order to see how these two "sources" of statistical discrimination differentially affect the voting patterns of players, assume that there are two groups: group A and group B. In addition, suppose that on average contestants from group A play the game better than contestants from group B, but that all players are equally capable of assessing one another's ability, regardless of group membership. Then we would expect that contestants from all groups would be more likely to vote for members of group B in the early rounds of the game when contestants clearly have an incentive to vote off players who they believe to be relatively weak. One simple method for examining whether this type of statistical discrimination could account for observed voting patterns is to look for group-level differences in the two most relevant factors that determine player ability: the percentage of questions correctly and decisions about whether and when to bank.

Now, suppose instead that contestants from group A and group B are equally skilled at

playing the game, but that contestants from group A have less accurate information about the ability levels of contestants from group B than they do about their fellow members from group A, then (assuming that players are risk neutral) members of group A will be more likely to vote off contestants from group A than they will contestants from group B. The reason is that, because members of group A have more accurate information about contestants from their own group than they do about contestants from group B, contestants from group A will place a higher weight on mean ability when assessing the ability level of contestants from group B than they do for contestants from group A. Thus, a contestant from group B with a relatively low signal will have a higher perceived ability than a contestant from group A who has that same signal. Thus, perhaps unexpectedly, in the early rounds of the game when the incentives are such that contestants will want to vote off weak players, if group ability distributions are the same and informational asymmetries lie at the root of statistical discrimination, then we would expect that contestants would be more likely to cast votes *against* members of their own group.

A final possibility is that the contestant's prior beliefs about the skills of blacks and women are simply inaccurate. It is possible, for example, that members of group A begin the game with the erroneous belief that members of group B are inferior at playing the game. In this case, we would expect to see that members of group B are more likely to be voted off in the early rounds of the game, and that this pattern would disappear as the incorrect prior is updated. Two types of analysis allow us to examine the possibility that players have such erroneous expectations. First, we examine data from episodes of the daily show for which all participants have the same occupation. Presumably members of the same occupation have better information about how particular groups will perform. If bad information is the reason that particular groups are voted off, one would expect to see a weaker pattern of discriminatory voting in shows where everyone has the same occupation. Second, if incorrect prior beliefs are driven by a characteristic that is not common to all males, then the test for explicit collusion discussed below can also be used to rule out incorrect prior beliefs as a source of discriminatory voting behavior.

4.2 Strategic Discrimination

In a strategic setting such as *The Weakest Link*, collusion may also play a role in discriminatory outcomes. Two types of collusion are possible here: explicit collusion (where agreements are made between groups of contestants on their voting behavior) and implicit collusion (where no agreements are discussed but individuals play strategies that yield focal points).

Contestants meet one another prior to the game and presumably there are opportunities to discuss with one another various (quite possibly collusive) strategies for playing the game. Whether and how players are able to commit to this strategy of collusion remains an open question, however, assuming that players are able to successfully act cooperatively, coalitions may be likely to form along racial or gender lines. To test for explicit collusion, we examine whether the votes of the individual players are correlated with the votes of the same group members after controlling for the votes of all the contestants. That is, there may be unobservable characteristics which make particular individuals more likely to be voted off. These will be captured by the total votes cast against a person by all the other players. If, however, a particular group votes together then the voting patterns of the other group members should affect the voting patterns of individual group members.

Discriminatory outcomes may also result from implicit collusion where individuals naturally play particular equilibria. For example, discriminating against members of group B may be a best response for members of group A given that the other members of group A are playing a discriminatory equilibrium. These focal point equilibria should be reenforcing. Hence, if group A succeeds in removing a member of group B in round 1, then this should make it more likely that they will vote off a member of group B in round 2. We can therefore test for explicit collusion by examining how who is voted off affects voting behavior in the next round.

4.3 Preference-Based Discrimination

In the prototypical model of preference-based discrimination, members of the majority group simply dislike working with members of the minority group and act as if there is some nonmonetary cost associated with hiring them. As a result, in equilibrium, workers from minority groups may earn lower wages on average.

In the Weakest Link, it is possible that contestants from one group may dislike playing the game with contestants from some other group. Thus, if we believe that members of group A dislike members of group B, we would expect that members of group A would be consistently more likely to vote off members of group B in every round. As the game progresses, however, the probability that any single player will win the game increases. As a result, the implicit cost of preference-based discrimination also increases. For this reason alone, even if the contestants have discriminatory preferences, their propensity to vote off blacks may fade as the game progresses. Thus, if the implicit cost of discrimination becomes too high, it may become difficult to separate statistical discrimination from preference-based discrimination. Therefore, if preference based discrimination exists, we would expect it to diminish with each round of the show.

5 Which Contestants Discriminate?

5.1 Descriptive Evidence

In order to understand the broad patterns in the data, we start by analyzing voting patterns by gender and race. Table 1 summarizes the voting behavior by round for both the daily and weekly shows. The table evaluates the voting behavior of three demographic groups: men, women, and whites[10] to determine if members of these demographic groups discriminate against players who are not group members (for example do men discriminate against women in their voting patterns). The task of discerning discrimination is complicated because even if both men and women voted randomly, men would cast more votes against women and women would cast more votes against men simply because a contestant will never vote for him or her self. To account for this problem, and the fact that the distribution of demographic types varies across episodes and rounds, we describe the voting behavior in terms of the mean "group bias statistic". For individual i voting in round r, the "group bias statistic" is given by equation 1.

$$\textbf{Group Bias Statistic} \;=\; \frac{1_{ir}}{\left[\frac{G_{ir}-1}{N_r-1}\right]}. \tag{1}$$

[10]The voting behavior of blacks is not considered due to small sample size.

Where 1_{ir} is an indicator variable that takes on a value of one if individual i votes for a contestant from his/her group in round r. G_{ir} is the number of contestant i's type in round r of his/her episode and N_r is the total number of contestants in round r of contestant i's episode.[11] A mean value of one for the group bias statistic implies no discriminatory outcomes, a value less than one implies discrimination against the other group and a value greater than one implies discrimination against one's own group.

The descriptive evidence show some surprising results. First, there is virtually no evidence of discrimination by whites against blacks; all of the values of our discrimination statistic are indistinguishable from one. Second, there is no discrimination against women by men in the early rounds, though in round 3 of both shows men are more likely to vote for women. The most surprising result, however, is that in both the weekly and the daily show women are substantially more likely to vote for men in round 1. This pattern of women voting off men continues in rounds 2 and 3 of the daily show, though at a diminishing rate.

Obviously, these patterns may reflect race and gender differences in the average ability to play the game. In particular, if men are not as successful at playing the game well, then this may explain why they are more likely than women to receive votes in the early stages of the game. The analysis of the following section controls for these differences. Also, note that the sample sizes for blacks are extremely small in the later rounds of the game. Thus, it may be difficult to draw firm conclusions about racial discrimination given the small number of observations available in the later stages of the game.

5.2 Evidence from Conditional Logits

In order to control for any possible race or gender differences in the contestants' abilities to play the game, we estimate conditional logits by show and by round of the probability contestants cast votes against other players as a function of those players' characteristics. We model the utility of player i voting for contestant j in round r as:

$$U_{ijr} = X_{ijr}\beta_r + \epsilon_{ijr}, \tag{2}$$

[11] Note that this statistic is only calculated for rounds in which contestant i has the option of voting for both group members and non-group members.

Table 1: Own-Group Bias in Voting Behavior

		Females			Males			Whites		
			Group Bias Statistic†			Group Bias Statistic†			Group Bias Statistic†	
Show	Round	Sample	Mean	Std. Dev.	Sample	Mean	Std. Dev.	Sample	Mean	Std. Dev.
	1	222	0.739*	1.141	222	0.976	1.220	331	0.949	0.634
Daily	2	189	0.783*	1.212	186	1.336	0.336	235	0.982	0.703
	3	150	0.910	1.177	123	0.756*	1.205	119	0.966	0.856
	4	66	0.909	1.003	62	0.935	1.006	68	1.000	1.007
	1	111	0.783*	1.086	113	1.000	1.185	133	1.004	0.479
	2	98	0.907	1.119	98	0.978	1.344	103	1.019	0.526
Weekly	3	84	0.933	1.160	83	0.689*	1.172	82	1.037	0.500
	4	64	1.115	1.307	72	0.806	1.162	62	0.871	0.689
	5	44	0.784	0.252	53	1.301	1.349	32	0.703	0.851
	6	26	1.077	0.287	22	1.091	1.019	12	0.833	1.030

† The "group bias statistic" is equal to $1_{ir}/\left[\frac{G_{ir}-1}{N_r-1}\right]$. Where 1_{ir} is an indicator variable that takes on a value of one if individual i votes for a contestant from his/her group in round r. G_{ir} is the number of contestant i's type in round r of his/her episode and N_r is the total number of contestants in round r of contestant i's episode. A mean value of one implies no discriminatory outcomes, a value less than one implies discrimination against the other group and a value greater than one implies discrimination against one's own group.

* Statistically different from 1 at the 95% level.

where β is a vector of coefficients to be estimated and ϵ_{ijr} is the unobserved preference individual i has for voting for contestant j. Included in X_{ijr} are controls for the percentage of the questions the player answered correctly in that round, whether the player was statistically the weakest link (answered the fewest questions correctly) in that round and also whether the player was the strongest link in that round. In addition, we control for the gender and races of the other contestants and whether the individual voting was of the same gender or race. Due to small sample sizes, we restrict our analysis of race to the daily show.[12] We assume that the ϵ_{ijr}'s are distributed i.i.d. extreme value, implying that the probability of voting for contestant j in round r is given by:

$$P_{ijr} = \frac{\exp(X_{ijr}\beta_r)}{\sum_{k=1}^{N_r} \exp(X_{ijr}\beta_r)} \tag{3}$$

where N_r is the number of contestants in round r.

Results for each round are shown in the appendix, with the results for round 1 of both the daily and the weekly show given in Table 2. We have the most confidence in the round 1 results for three reasons. First, sample sizes are much larger here than in the later rounds. Second, strategies may change as the game progresses and may depend upon the history of play. Finally, the pool of contestants in all but the initial round is endogenously determined by the voting behavior.

The first set of rows examines the voting behavior of women. Consistent with the descriptive evidence, women appear to be much more likely to cast votes against men than against other women. As shown in the appendix, this effect disappears immediately after round 1 in the weekly show while slowly diminishing in the daily show. This effect is quite large. Consider a daily show with six white contestants, three men, and three women, all with identical performance. The probability of a given female voting for a particular male contestant is 23.3% while the probability that she votes for a particular female contestant is substantially lower at 15.1%.

Also consistent with the descriptive evidence, the next two sets of rows show that there is no indication of discrimination by men against women or by whites against blacks. This holds true for all rounds of the game except for round 3 where in both the weekly and the daily show men are more likely to vote for women, though the effect is not statistically significant.

[12] None of the qualitative results change if race is included in the weekly show.

However, this discriminatory behavior by men may be due to the discrimination against men and the corresponding selection of men who make it to round 3.

As one would expect, for all demographic groups in the early stages of the game the higher the percentage of the questions that the player answers correctly the less likely other contestants are are to cast votes against that player. However, as the game progresses, the percent correct becomes less and less important. This confirms the basic logic that players have an incentive to vote off weak players in the early rounds of the game but in the later rounds, this incentive is partially offset by the incentive to vote off stronger players.

6 Why Women Vote for Men

This section attempts to distinguish between the three possible hypotheses for why women are more likely to vote for men: statistical discrimination, strategic discrimination, and preference-based discrimination.

6.1 The Case Against Statistical Discrimination

Recall that statistical discrimination with correct priors can take two forms. In the first, the mean performance level is different across groups. For statistical discrimination of this type to explain females discriminating against males, the performance by males must be on average worse than the performance by females. Table 3 documents the average percent correct for males and females by round and by show type. There is virtually no difference in performance levels for males and females in either show in the early rounds, while in the later rounds males actually perform better than their female counterparts.[13] In addition, if men are truly worse at playing the game, then both women *and men* should be more likely to vote against men in the early rounds of the game. We find no evidence of this. Hence, there is no evidence that statistical discrimination based upon mean group performance is driving females to vote against males.

The second type of statistical discrimination, where the signals on ability are more infor-

[13]That males preform better in later rounds may be a result of relatively low ability males being voted off earlier than relatively low ability females due to the discrimination.

Table 2: Conditional Logit Estimates of Round 1 Voting Behavior[†]

	Females Only			
Characteristics of Other Contestants	Daily Show		Weekly Show	
Male	0.433*	(0.164)	0.538*	(0.224)
Percent Correct	-2.667*	(0.470)	-2.913*	(0.506)
Weakest Link	0.305	(0.242)	0.429	(0.283)
Strongest Link	-0.016	(0.342)	0.275	(0.510)
Black	-0.145	(0.292)		
Same Race	-0.189	(0.291)		
Observations	222		111	
	Males Only			
Characteristics of Other Contestants	Daily Show		Weekly Show	
Female	-0.099	(0.150)	0.084	(0.210)
Percent Correct	-2.018*	(0.410)	-2.754*	(0.488)
Weakest Link	0.085	(0.234)	0.354	(0.263)
Strongest Link	-0.596	(0.352)	-0.728	(0.746)
Black	0.199	(0.308)		
Same Race	-0.201	(0.309)		
Observations	222		113	
	Whites Only			
Characteristics of Other Contestants	Daily Show		Weekly Show	
Female	-0.170	(0.124)		
Same Gender	-0.271	(0.123)		
Percent Correct	-2.508*	(0.352)		
Weakest Link	0.057	(0.192)		
Strongest Link	-0.360	(0.284)		
Black	0.000	(0.144)		
Observations	349			

[†] Conditional logit estimates of the probability of an individual voting for a particular contestant. Standard errors in parenthesis.

* Statistically different from zero at the 95% confidence interval.

Table 3: Performance by Race and Gender

		Females			Males			Blacks			Whites		
Show	Round	Sample	Percent Correct	Standard Deviation	Sample	Percent Correct	Standard Deviation	Sample	Percent Correct	Standard Deviation	Sample	Percent Correct	Standard Deviation
	1	225	0.668	0.306	225	0.667	0.340	97	0.639	0.346	353	0.675	0.317
Daily	2	186	0.630	0.341	184	0.632	0.336	79	0.657	0.325	291	0.624	0.342
	3	159	0.600	0.308	141	0.614	0.337	63	0.583	0.366	237	0.613	0.309
	4	115	0.580	0.353	110	0.659	0.292	48	0.647	0.334	177	0.611	0.325
	1	111	0.608	0.367	113	0.631	0.328	27	0.629	0.391	196	0.618	0.342
	2	98	0.613	0.311	98	0.611	0.344	23	0.563	0.343	173	0.618	0.325
Weekly	3	85	0.579	0.303	83	0.631	0.302	20	0.698	0.319	148	0.592	0.299
	4	67	0.569	0.274	73	0.655	0.308	18	0.580	0.323	122	0.619	0.292
	5	49	0.511	0.252	63	0.639*	0.246	12	0.636	0.277	100	0.577	0.254
	6	40	0.539	0.287	44	0.548	0.236	8	0.413	0.289	76	0.558	0.254

* Statistically different from female performance in the round at the 95% confidence level.

mative for one group than another, is ruled out by the fact that women are discriminating against men rather than against women. Recall that information-based statistical discrimination means within-group performance is more informative that out-of-group performance. Hence, poor performance is weighted more heavily when it is by a member of one's own group. Should this type of discrimination occur, women would be more likely to vote for women than for men. We find the exact opposite result.

Next, we turn to the case of incorrect prior beliefs regarding group performance. It is possible that in the early rounds of the game, women have not yet learned that men perform as well as women in the game. Our data set contains 13 episodes of the daily show in which all of the contestants have the same occupation. Should erroneous prior beliefs exist, it seems reasonable to expect that workers in the same occupation would have more informative priors on the abilities of their opposite-sex contestants than do contestants of differing occupations. In Table 4 we interact a dummy variable for the shows in which all of the contestants have the same occupation with the the male indicator. Under the incorrect priors hypothesis, one would expect a negative coefficient on this interaction term. However, the coefficient on this term is positive, although insignificant–providing weak evidence that women in the same occupation are more likely to vote off men then are women in different occupations. Taken by itself, this result does not conclusively rule out the possibility of incorrect prior beliefs. However, given additional evidence against this hypothesis provided below in the discussion of explicit collusion, we feel comfortable focussing our attention on other potential explanations for the observed voting patterns.

6.2 The Case Against Strategic Discrimination

We now test whether women are acting cooperatively with one another. There are two basic types of collusive behavior in which women might engage: implicit collusion and explicit collusion. Under explicit collusion, the presumption is that the women are following some agreed upon pattern of voting. Under implicit collusion, the presumption is that women have a tacit agreement to vote for men. To test for explicit collusion, we include in the female round 1 conditional logits the total votes cast for each contestant by the other contestants as well

Table 4: Do Females Have Wrong Priors on Male Ability? Evidence From Same Occupation Shows †

Characteristics of Other Contestants	Daily Show- Females Only	
Male	0.398*	(0.178)
Male × Same Occupation Show	0.238	(0.472)
Percent Correct	-2.649*	(0.471)
Weakest Link	0.309	(0.243)
Strongest Link	-0.026	(0.342)
Black	-0.139	(0.292)
Same Race	-0.192	(0.292)
Observations	222	

† Conditional logit estimates of the probability of a female voting for a particular contestant. Standard errors in parenthesis.

* Statistically different from zero at the 95% confidence level.

Table 5: Are Female Votes Correlated? Testing for Explicit Collusion[†]

	Females Only			
Characteristics of Other Contestants	Daily Show		Weekly Show	
Male	0.470*	(0.200)	0.480*	(0.242)
Percent Correct	-1.920*	(0.504)	-2.312*	(0.551)
Weakest Link	0.156	(0.262)	0.243	(0.300)
Strongest Link	0.028	(0.354)	0.198	(0.512)
Black	-0.161	(0.321)		
Same Race	-0.237	(0.321)		
Total Votes Cast By:				
Other Contestants	0.541*	(0.102)	0.223	(0.120)
Other Female Contestants	-0.054	(0.179)	-0.028	(0.206)
Observations	222		111	

[†]Conditional logit estimates of the probability of a female voting for a particular contestant. Standard errors in parenthesis.

* Statistically different from zero at the 95% confidence level.

as the total votes cast by the other female contestants.[14] The first variable captures the fact that there are unobservable characteristics that may lead individuals to vote for particular contestants. In order for there to be explicit collusion among females, the coefficient on the latter variable, total votes cast by other women, must be positive. Results for this specification are shown in Table 5. Note that the total votes cast by other women has no more predictive power than the total votes cast by men as the coefficient on the total votes cast by other women is small and insignificant. Hence, there is no evidence of explicit collusion.

[14]We also performed this analysis for the male sub-sample and for round 2 of the daily show. Gender again had no effect on the voting behavior of men while the results for round 2 are very similar to the results from round 1; female votes are not correlated except through total votes.

That the coefficient on females votes is small and insignificant allows us to reject many other explanations for the discriminatory behavior. In particular, it suggests that women are not discriminating on some unobserved characteristic correlated with being male as this too should have led to a positive coefficient on votes cast by women. If particular men are arrogant or make poor (in the eyes of women) banking decisions, then the votes of the other women should reflect this. Instead, the results suggest that whatever characteristics are unappealing to women are equally unappealing to men. Note that this also helps to rule out discrimination based upon bad expectations. In particular, the only bad expectations that can exist involve women expecting all men to perform equally poorly–that is, there can be no correlation with any observable (to the voter, not to the researcher) characteristics.

We next test whether or not women are implicitly colluding to vote off men. The first evidence that this is not the case comes from the diminishing coefficient on same sex as we move to the later rounds. If implicit collusion was working, then there is no reason to stop colluding against men in the later rounds. The second indication that implicit collusion is not driving the results is that implicit collusion does not appear to be reenforcing. That is, if implicit collusion were occurring successfully (a man was voted off), then implicit collusion should be more likely to happen in the next round than if the implicit collusion did not work. Table 6 reports conditional logits for rounds 2 and 3 of both the daily and weekly shows with an indicator variable for whether a man (two men) were voted off prior to round 2 (3). Although not significant, three of the four interactions are negative, implying that, if anything, removing males makes it less likely that women will vote off men in future rounds. Hence, we find no evidence of strategic discrimination through either implicit or explicit collusion.

6.3 The Case for Preference-Based Discrimination

The only remaining explanation is that women simply prefer playing with women. Consistent with this explanation, the coefficient on male diminishes in later rounds as the price of discriminating based upon preferences increases. Further, the taste discrimination falls faster in the weekly show, disappearing after one round. This is consistent with the theory as the cost of discriminating is higher in the weekly show where the total prize money is substantially

Table 6: Do Females Use Males as Focal Points? Testing for Implicit Collusion†

Characteristics of Other Contestants	Females Only- Round Two			
	Daily Show		Weekly Show	
Male	0.538	(0.332)	-0.423	(0.413)
Male × Male Voted Off in Round 1	-0.314	(0.389)	-0.759	(0.505)
Percent Correct	-1.568*	(0.441)	-3.387*	(0.623)
Weakest Link	0.240	(0.247)	0.057	(0.308)
Strongest Link	0.053	(0.283)	-0.381	(0.774)
Black	-0.048	(0.298)		
Same Race	-0.079	(0.300)		
Observations	179		98	
	Females Only- Round Three			
	Daily Show		Weekly Show	
Male	0.242	(0.240)	0.302	(0.289)
Male × Male Voted Off in Rounds 1 & 2	-0.163	(0.394)	0.443	(0.624)
Percent Correct	-1.750*	(0.487)	-2.678*	(0.687)
Weakest Link	-0.028	(0.136)	0.342	(0.333)
Strongest Link	-0.140	(0.290)	-0.359	(0.588)
Black	0.665	(0.556)		
Same Race	-1.081	(0.553)		
Observations	156		85	

†Conditional logit estimates of the probability of a female voting for a particular contestant. Standard errors in parenthesis.

* Statistically different from zero at the 95% confidence level.

larger.

We have attempted to further test this explanation by including controls for the amount of money banked (the size of the pot) at the voting stage. If preference-based discrimination exists, higher amounts of money banked should lead to less discrimination. Estimates of models of this type were mixed, with evidence that at the lowest quartile of money banked discrimination increases (consistent with the theory) but discrimination also increases at the highest quartile of money banked (inconsistent with the theory). One possible explanation for the latter is that, in order to bank a large sum of money, every contestant needs to perform well, and if every player is similarly talented at playing the game, then there is no extra cost associated with voting against men, regardless of the amount of money banked. Unfortunately, the small sample sizes make it difficult to further test this hypothesis.

The finding that women have discriminatory preferences against men initially may seem rather surprising. However, there are a number of possible explanations for this type of behavior. First, women may dislike certain aspects of how men play the game.[15] Women may also feel more compassionate towards women than towards men and may be more willing to forgive women for incorrect answers. An additional possibility is that women may not like playing with men because they fear that they will not compete as well against men as women in the later rounds of the game. Gneezy, Niederle and Rustichini (2003), for example, find evidence that, in an experimental setting, women perform more poorly when competing against men than against women.[16] Understanding more about the sources of women's discriminatory preferences in this game is an intriguing area for future research.

[15]Though for this to be true, it would have to be how males in general play the game and not correlated with any other feature (such as how certain males speak or how certain males look). If the discrimination result was driven by females disliking the way males with particular features played the game, the test we performed for explicit collusion would have shown that female votes were correlated beyond the correlation with male votes.

[16]If this were the case, we would expect discrimination to persist and even increase as the game progresses which we do not see in the data. The selection of contestants into the later rounds, however, may be why we do not see this in the conditional logit estimates.

7 Conclusion

Understanding the nature of discrimination and it's contribution to both racial and gender earnings inequality involves tackling two questions. First, we would like to know whether individuals discriminate. Second, we would like to know why individuals discriminate. In this paper we attempt to address both of these questions by examining the voting behavior of contestants in *The Weakest Link*.

Interestingly, we find no evidence of discrimination against either blacks or women. However, we do find that women discriminate against men in the early rounds of both the daily and weekly shows. We are able to rule out both statistical discrimination and strategic discrimination as explanations for women's behavior. The one theory consistent with the observed voting trends is preference-based discrimination. In other words, it appears that women simply prefer playing with women rather than with men.

It is difficult to determine to what extent the results of our paper shed light on the extent and nature of labor market discrimination. The behavior of a select group of contestants playing a game that is broadcast on network television may be very different from that of employers in the labor market. Nonetheless, our paper reveals provocative patterns in the behavior of women towards men in strategic environments. In addition, a major contribution of this paper is to shed light on the types of experimental designs that could generate the data needed to distinguish between statistical discrimination and preference-based discrimination. Toward this end, we believe that there are a number of basic insights from our paper. First, it should be possible to test for preference-based discrimination by designing experiments in which there is variation in the cost of discriminatory actions. Second, by designing experiments in which participants assess the ability of other participants based on limited information, it should be possible to test for statistical discrimination by changing the quantity and quality of that information.

In particular, although the complex structure of *The Weakest Link* precludes us exploiting the fact that voting incentives shift as the game progresses, this feature of the game suggests that one could potentially use changing incentives to distinguish between preference-based discrimination and statistical discrimination. By comparing how individuals choose partners in experimental settings where payoffs depend positively on the ability of one's partner

with games where payoffs depend negatively on the ability of one's partner (since statistical discrimination is presumably sensitive to this switch in incentives, whereas preference-based discrimination is not) it would be possible to distinguish between the two hypotheses. We are hopeful that future work will make use of this insight.

References

[1] AIGNER, D. and CAIN, G. (1977), "Statistical Theories of Discrimination in the Labor Market," *Industrial and Labor Relations Review*, **30**, 175-87.

[2] ALTONJI, J. and PIERRET C. (2001), "Employer Learning and Statistical Discrimination," *Quarterly Journal of Economics*, textbf116, 313-350.

[3] ARROW, K. (1973), "The Theory of Discrimination" in Orley Ashenfelter and Albert Rees (eds.), *Discrimination in Labor Markets* (Princeton, NJ: Princeton University Press).

[4] ASHENFELTER, O. and HANNAN, T. (1986), "Sex Discrimination and Market Competition: The Case of the Banking Industry." *Quarterly Journal of Economics*, 149-173.

[5] ARCIDIACONO, P. and VIGDOR, J. (2002), "Do People Value Racial Diversity? Evidence from Nielsen Ratings," working paper.

[6] BECKER, G. (1957), *The Economics of Discrimination*, Chicago, IL: The University of Chicago Press.

[7] BERTRAND, M. and MULLAINATHAN, S. (2002), "Are Emily and Brendan More Employable than Lakisha and Jamal? A Field Experiment on Labor Market Discrimination," University of Chicago, Graduate School of Business Working Paper.

[8] COATE, S. and LOURY, G. (1993), "Will Affirmative Action Eliminate Negative Stereotypes?" *The American Economic Review*, **83**, 1220-1240.

[9] CROSS, H., KENNY, G. and ZIMMERMAN, W. (1990), *Employer Hiring Practices: Differential Treatment of Anglo and Hispanic Job Seekers* (Washington, DC: Urban Institute Press).

[10] FERSHTMAN, C. and GNEEZY, U. (2001), "Discrimination in a Segmented Society: An Experimental Approach." *The Quarterly Journal of Economics*, **116**, 351-377.

[11] GNEEZY, U., NIEDERLE, M. and RUSTICHINI, A. (2003), "Performance in Competitive Environments: Gender Differences," forthcoming *Quarterly Journal of Economics.*

[12] GOLDIN, C. and ROUSE, C. (1997), "Orchestrating Impartiality: The Impact of "Blind" Auditions on Female Musicians," *NBER Working Paper No. 5903.*

[13] ECKMAN, J. (1998), "Detecting Discrimination", *The Journal of Economic Perspectives*, **12**, 101-116.

[14] ECKMAN, J. and SIEGELMAN, P. (1992), "The Urban Institute Audit Studies: Their Methods and Findings," in Michael Fix and Raymond Struyk, *Clear and Convincing Evidence: Measurement of Discrimination in America* (Lanham, MD: Urban Institute Press).

[15] HOLM, H. (2000), "Gender-Based Focal Points," *Games and Economic Behavior*, **32**, 292-314.

[16] KAHN, L. and SHERER, P. (1988), "Racial Differences in Professional Basketball Players' Compensation," *Journal of Labor Economics*, **6**, 40-61.

[17] LEVIT, S. (2003), "Testing Theories of Discrimination: Evidence from the Weakest Link," NBER Working Paper 9449.

[18] LUNDBERG, S. and STARTZ, R. (1983), "Private Discrimination and Social Intervention in a Competitive Labor Market", *The American Economic Review*, **73**, 340-347.

[19] MASON, C. PHILLIPS O. and REDDINGTON, D. (1991), "The Role of Gender in Non-Cooperative Games," *Journal of Economic Behavior and Organization*, **15**, 215-335.

[20] NEUMARK, D., BANK, R. and VAN NORT, K. (1995), "Sex Discrimination in Restaurant Hiring: An Audit Study," *NBER Working Paper No. 5024.*

[21] OETTINGER, G. (1996), "Statistical Discrimination and the Early Career Evolution of the Black-White Wage Gap," *Journal of Labor Economics*, **14**, 52-78.

[22] PHELPS, E. (1972), "The Statistical Theory of Racism and Sexism," *American Economic Review*, **62**, 659-61.

[23] TURNER, M., FIX, M. and STRUYK, R. (1991), *Opportunities Denied, Opportunities Diminished: Racial Discrimination in Hiring* (Washington, DC: Urban Institute Press).

Table A.1: Conditional Logits of Voting Behavior in the Daily Show[†]

Characteristics of Other Contestants		Round 1		Round 2		Round 3		Round 4	
Females Only	Male	0.433*	(0.164)	0.314	(0.172)	0.183	(0.190)	0.088	(0.271)
	Percent Correct	-2.667*	(0.470)	-1.573*	(0.440)	-1.742*	(0.490)	-0.832	(0.744)
	Weakest Link	0.305	(0.242)	0.239	(0.247)	-0.024	(0.135)	0.080	(0.354)
	Strongest Link	-0.016	(0.342)	0.049	(0.283)	-0.145	(0.290)	0.431	(0.318)
	Black	0.145	(0.292)	0.042	(0.298)	0.685	(0.555)	0.723	(0.575)
	Same Race	-0.189	(0.291)	-0.083	(0.299)	-1.102	(0.552)	-1.110	(0.579)
	Observations	222		179		156		113	
Males Only	Female	0.099	(0.150)	-0.047	(0.167)	0.412	(0.226)	0.077	(0.265)
	Percent Correct	-2.018*	(0.410)	-1.658*	(0.428)	-2.234*	(0.609)	-0.692	(0.756)
	Weakest Link	0.085	(0.234)	-0.001	(0.245)	0.037	(0.170)	0.123	(0.346)
	Strongest Link	-0.596	(0.352)	-0.452	(0.312)	-0.120	(0.353)	0.232	(0.346)
	Black	0.199	(0.308)	0.030	(0.311)	0.328	(0.434)	0.586	(0.473)
	Same Race	-0.201	(0.309)	0.034	(0.313)	0.204	(0.434)	-0.273	(0.480)
	Observations	222		179		138		110	
Whites Only	Male	0.170	(0.124)	0.171	(0.135)	-0.044	(0.161)	0.162	(0.212)
	Same Gender	-0.271*	(0.123)	-0.159	(0.134)	-0.295	(0.159)	0.089	(0.207)
	Percent Correct	-2.508*	(0.352)	-1.818*	(0.350)	-1.920*	(0.455)	-0.638	(0.578)
	Weakest Link	0.057	(0.192)	-0.045	(0.197)	-0.086	(0.155)	0.092	(0.284)
	Strongest Link	-0.360	(0.284)	0.024	(0.228)	-0.203	(0.246)	0.324	(0.256)
	Black	0.000	(0.144)	0.002	(0.161)	0.000	(0.189)	0.020	(0.253)
	Observations	349		280		232		175	

[†] Conditional logit estimates of the probability of an individual voting for a particular contestant. Standard errors in parenthesis.

* Statistically different from zero at the 95% confidence interval.

Table A.2: Conditional Logits of Voting Behavior in the Weekly Show[†]

Characteristics of Other Contestants		Round 1		Round 2		Round 3		Round 4		Round 5		Round 6	
	Male	0.538*	(0.224)	-0.067	(0.235)	0.395	(0.259)	-0.170	(0.282)	0.547	(0.366)	0.018	(0.424)
Females	Percent Correct	-2.913*	(0.506)	-3.472*	(0.623)	-2.726*	(0.686)	0.171	(0.770)	-0.177	(1.037)	0.333	(1.591)
Only	Weakest Link	0.429	(0.283)	-0.002	(0.301)	0.340	(0.334)	0.866	(0.383)	0.193	(0.433)	0.822	(0.614)
	Strongest Link	0.275	(0.510)	-0.377	(0.774)	-0.321	(0.585)	-0.788	(0.499)	-0.561	(0.470)	0.025	(0.562)
	Observations	111		98		85		66		49		39	
	Female	-0.084	(0.210)	0.220	(0.230)	0.420	(0.263)	0.250	(0.287)	-0.487	(0.297)	-0.519	(0.495)
Males	Percent Correct	-2.754*	(0.488)	-3.058*	(0.602)	-2.247*	(0.657)	-2.396*	(0.851)	0.136	(1.095)	1.587	(1.363)
Only	Weakest Link	0.354	(0.263)	-0.107	(0.313)	0.346	(0.317)	0.170	(0.328)	0.379	(0.393)	1.266*	(0.608)
	Strongest Link	-0.728	(0.746)	-34.060[‡]		-0.452	(0.648)	-0.451	(0.585)	-0.039	(0.412)	0.304	(0.540)
	Observations	113		98		83		71		63		42	

[†] Conditional logit estimates of the probability of an individual voting for a particular contestant. Standard errors in parenthesis.

[‡] No males voted for the strongest link in round 2 of the weekly show.

* Statistically different from zero at the 95% confidence interval.