Social Stigma and Subsequent Competitive Behavior*

Natalia I. Valdez Gonzalez[†], Marco A. Palma[‡], Alexander L. Brown[§]

February 6, 2023

Abstract

Social signaling influences economic behavior. For instance, individuals may exhaust resources to competitively signal higher levels of social status. Conversely, individuals may avoid signaling their status to minimize the stigma associated with low status. We conduct a laboratory experiment to explore how benefit eligibility stigma drives subsequent decisions to enter a tournament competition. Similar to Friedrichsen et al. (2018), we induce a stigma associated with a benefit for the low status group. We then introduce a treatment in which the stigma is reduced by expanding the benefit eligibility to a middle-status group in a "plausible deniability" treatment. While we do not observe evidence of a stigma positively affecting benefit take-up for the low-status group, we do observe a difference in preferences for competitiveness in a subsequent and unrelated task; namely, when individuals in the middle group qualify for the benefit their rate of competition is reduced by 33% compared to the treatment in which they do not qualify. A potential interpretation of our results would suggest expansion of eligibility of certain government assistance programs may produce unintended consequences for the newly eligible.

*We thank the members of the Human Behavior Lab who assisted with data collection. We would like to thank Rodrigo Velez, Ragan Petrie, Yu Zhang, and Brit Grosskopf for their helpful comments to develop this work. We thank attendees of the Southern Agricultural Economics Association 2020, Economic Science Association North America 2021, and Agricultural and Applied Economics Association 2022 for their feedback. This work was funded by the International Foundation for Research in Experimental Economics.

[†]Department of Agricultural Economics, Texas A&M University

[‡]Department of Agricultural Economics, Texas A&M University

[§]Department of Economics, Texas A&M University

1 Introduction

It is well known that people are driven to showcase a desirable status in different market contexts through consumption of prestigious goods and services (Bursztyn and Jensen, 2017; Clingingsmith and Sheremeta, 2018). On the other hand, people also experience stigma from low-status environments, such as qualifying for special assistance programs due to low-income or performance (Moffitt, 1983). The extent and effects of this latter stigma is a crucial question for policy as government and nonprofit assistance programs are widespread in most developed nations; these programs are designed to help individuals meet basic housing and food needs, special education, or remedial training (Andrade, 2002; Daponte et al., 1999; Currie et al., 2001; Friedrichsen et al., 2018).[1] While social welfare programs are conceived to alleviate specific problems and benefit recipients, the stigma they carry may work in the opposite direction. Individuals in need may not access relevant programs out of stigma aversion or alternatively program participation may reduce the confidence of recipients later in their lives.

This paper studies how eligibility for a special social benefit based on low status influences take-up and later preferences for competitiveness. Since it is difficult if not impossible to imagine a field setting where variables such as welfare participation, stigma, and labor competition could be exogenously varied to test relevant questions, the experimental laboratory provides an ideal environment to study these relationships. The experiment consists of two main stages. In the first stage, subjects are randomly assigned to groups of three and are evaluated under a general knowledge quiz. Based on their performance, they are assigned a low, middle, or high status with higher earnings for higher performance. The low-status individuals are able to claim a welfare benefit which they claim by coming to the front of the room, thereby inducing feelings of stigma. In a separate Plausible Deniability

[1]For example, in the United States, people in the lowest income bracket are eligible to receive benefits from social welfare programs to assist them in procuring food and healthcare. As of 2019, 59 million Americans were eligible for one of the safety net programs, which accounts for about one fifth of the population (Minton and Giannarelli, 2019).

Treatment, we exogenously expand the benefit eligibility criteria to the middle performing individuals, thereby reducing the stigma associated with take-up. In both treatments, over 85% of eligible individuals claimed the benefit, suggesting that the stigma had a slight effect on overall benefit take-up.[2]

The second parts of both treatments are identical, featuring an elicitation of preferences for competitiveness based off Niederle and Vesterlund (2007). Subjects are paid to add up five two-digit numbers. Across treatments, subjects show no differences in performance when they are paid a piece-rate for this task. The basic results show that the performance in the first stage influences decisions to compete in a subsequent task. There are no differences in the propensity to compete for the high-status group across treatments nor for the low-status group. However, among the middle-status group roughly one-third as many subjects wish to compete in the control relative to the Plausible Deniability Treatment. No differences in performance, risk preferences, or confidence about their relative performance can explain the differences in competitiveness among the middle status group.

Our results also include a time-series of emotional responses of our subjects throughout the experiment. Interestingly, we find that benefits eligibility appears to increase feelings of positive emotions during the first stage of the experiment, but there is no observed effect on the second stage. Thus it does not appear the decision to not compete is associated with any negative emotional feeling, though perhaps the absence of a positive one.

Extensive literature has investigated the effects of social status on economic behavior, especially for high status in markets (Ball et al., 2001; Clingingsmith and Sheremeta, 2018; Bursztyn and Jensen, 2017). A high status is typically perceived as something desirable, and individuals seek to showcase or signal a high status through conspicuous consumption or consumption of prestigious goods visible to others (Veblen, 1899; Clingingsmith and Sheremeta, 2018; Bursztyn and Jensen, 2017; Palma et al., 2017). Clingingsmith and Sheremeta (2018) provided subjects the opportunity to make chocolate truffle purchases, a luxury good, in a

[2]Benefit take-up is negatively correlated with self-reported measures of shyness (Cheek and Buss, 1981), so to certain individuals, the social stigma was salient.

laboratory-controlled setting and found that they purchased more to signal a high status when their consumption behavior was visible to others, indicating the presence of conspicuous consumption. Bursztyn et al. (2017) conducted a field experiment that exogenously altered the qualifications for eligibility of different credit card tiers; they find evidence of pecuniary emulation in that individuals sought the highest status level of credit cards they could attain. Butera et al. (2022) suggest that individuals are willing to pay to showcase their high status when status is related to gym attendance; they fit the behavior from their study to a structural model and find that high performers experience significant utility gains while low performers experience significant utility losses. While there has not been much literature focusing on stigma from low status, previous work documents that individuals experience stigma from participating in government benefit programs (Andrade, 2002; Daponte et al., 1999; Currie et al., 2001) or educational programs (Bursztyn and Jensen, 2015; Bursztyn et al., 2019). Some literature documents benefit-eligibility stigma driving the decision to not take welfare benefits (Moffitt, 1983; Major and O'brien, 2005; Andrade, 2002). Specifically, eligible low-status individuals are less likely to accept the benefit when their decision is visible to others compared to private environments (Friedrichsen et al., 2018). Our research induces three different statuses, which gives us the opportunity to observe behavior for three distinct status groups. We observe benefit take-up behavior when it is visible to others, yet we focus on the preferences for competitiveness and how benefit-eligibility stigma affects the decision to enter a tournament competition.

Preferences for competition have been extensively studied in previous literature. In general, the literature focuses on gender, finding robust results that women choose to compete at lower rates than men (Niederle and Vesterlund, 2007). However, identity seems to be a driver for competitive behavior (Shih et al., 1999; Ibarra et al., 2010; Zhang et al., 2020). Recent expansions in the competition literature also cover how identity more broadly affects preferences for competition. Social norms associated with a specific identity may be drivers for how individuals choose to compete (Benjamin et al., 2010). When exploring how being associated with low economic status affects competition, Banker et al. (2020) find that poverty salience is associated with lower rates of competition. Given that in our experiment the stigma arises from having low cognitive ability and lower earnings and as a result, a low status, our environment aligns with previous literature investigating the role of socioeconomic status, financial scarcity, and performance in cognitive ability for different status levels (Mani et al., 2013; Hoff and Walsh, 2018). The research exploring socioeconomic status and markets has focused mainly on the high and the low status groups, but the middle status group may provide an interesting group to study benefit eligibility through plausible deniability since many policy discussions relate to defining the limits of eligibility. Plausible deniability occurs when an agent uses the context or environment to deny responsibility or intent for their own actions (Bolton et al., 2021; Gillies et al., 2019). We add to the literature on social status by exploring whether plausible deniability reduces the stigma for the low-status group. We also study whether plausible deniability affects the rates of competition of the middle-status group.

We find that the social benefit eligibility reduces the competitiveness of the middle-status performing group. The social benefit is tied to a low-performing perception that seems to affect subsequent competitive decisions for the middle-status group. The low-status group did not increase competitiveness after the benefit is expanded to the middle-status group in the Plausible Deniability Treatment. The tasks in each of the two stages of our study are different by design and the performance in the first task did not affect the performance in the competitive environment task. The performance of the three status groups is the same during the second stage piece-rate and tournament tasks. Therefore, choosing to avoid competition was costly for the middle-status group. Those who were the top performers in the middle-income group could have increased their earnings by 60% if they chose to compete. While it is inappropriate to directly generalize stylized laboratory results like these into policy settings, a possible contextualization of these results is that expansion of government assistance programs may have an unintended negative effect on the newly eligible population. Of course, this effect would need to be weighed against the actual benefits of the policy. We return to this interpretation in our concluding section.

The rest of the paper is structured as follows. Section 2 presents the experimental design and procedures. Section 3 presents hypotheses. Section 4 provides results. Section 5 provides discussion of the findings and implications for future work in this area.

2 Experimental Design and Procedures

This experiment featured two between-subject treatments, Control and Plausible Deniability, randomly assigned at the session level. The difference between the two groups was the eligibility criteria for the additional benefit in Stage 1. Specifically, in the Control group, only the low-performance group were eligible for an additional benefit, similar to the treatment featured in Friedrichsen et al. (2018). In the Plausible Deniability Treatment, low and middle performance individuals were eligible for the additional benefit. Thus low-performers could "plausibly deny" they were low status (i.e., insist that they were middle status) in this particular treatment, reducing their stigma/visibility as low performance individuals. We also explore how the eligibility for this additional benefit affects willingness to compete in a second stage.

To determine our sample size, we conducted a power analysis based on the earned, subsidy and random, subsidy treatments in Friedrichsen et al. (2018).[3] We used these treatments because they are the closest to our application to observe the effect of plausible deniability and take up behavior. Given the take up proportions from these treatments, we employ a Pearson's chi-squared test and find that we need approximately 133 subjects per treatment to achieve a power of 0.8 and alpha level of 0.05.

Each experimental session had three stages that transpired as follows.

[3] Friedrichsen et al. (2018) had 69 and 48 subjects in the subsidy and random subsidy treatments respectively.

Table 1: Payment table

Rank	Status	Payment
1	High	\$6
2	Middle	\$4
3	Low	\$2

2.1 Stage 1: Status assignment and benefit allocation

In the first part of the experiment, subjects were randomly assigned to a group of three players that remained anonymous. Then, they had ten minutes to complete a 15-question general knowledge quiz that covered several topics and required no specific training to answer. The questions were selected from a bank of general knowledge questions (Kassas and Palma, 2019). Subjects were informed that their performance on the quiz relative to the other players in their randomly assigned group would determine their status and payment schedule as described in Table 1.

After completing the quiz, but before revealing performance, subjects were informed of the additional \$1 benefit and the eligibility conditions of their respective treatment. The additional benefit represents a 50% increase in earnings from Stage 1 for those who earned a low status and 25% increase in earnings from Stage 1 for those who earned a middle status. This level of benefit also allowed us to maintain monotonicity in earnings across status groups. The strategy method was employed to ask all subjects whether they would accept a \$1 benefit if they were eligible. Subjects' beliefs about the proportion of eligible people in the room that would claim the benefit were also elicited. Only after a subject had completed these tasks would they learned their status based on their relative performance. For those who were eligible and requested the benefit, the session monitor called out their experiment ID publicly. Benefit claimers had to come up to the front of the room to receive a sheet of paper with an additional \$1 benefit to be added to their final compensation, following a similar procedure in Friedrichsen et al. (2018).

2.2 Stage 2: Competition

In the second stage of the experiment, the subjects engaged in a competitive tournament entry task following Niederle and Vesterlund (2007). First, they were asked to complete as many five two-digit summations as they could within five minutes under a piece-rate payment scheme. They were informed that they would receive \$0.50 per correct answer. They were provided with scratch paper but were not allowed to use calculators. After completing the first exercise, subjects were assigned to do the same exercise again but had the option to choose between a paid piece-rate, the same way as the first exercise, or to participate in a tournament against the other two people within their group, that is, the same group from Stage 1. Specifically, they would receive \$1.50 per correctly calculated answer if they were the top performer, and \$0 otherwise. They were not informed of their performance in this task. However, after the subjects completed these tasks, they were asked how they believed to have performed in the second task relative to the other two people within the group (top performer, middle performer, or lowest performer).

2.3 Stage 3: Additional tasks

The subjects then revealed their risk preferences using a multiple price list (Holt and Laury, 2002), completed a social preferences task (Bartling et al., 2009), completed a shyness scale (Cheek and Buss, 1981), and then answered some basic demographic questions. We collected these measures as control measures for competition and benefit claim rates. After completing these additional tasks subjects found out their earnings and privately received their payments in sealed envelopes.

2.4 Experimental Procedures

Thirty-one experimental sessions were conducted in the Fall of 2019. A total of 276 subjects participated in sessions of 6 to 12 people.[4] The sample was drawn from the student body of a large university. All subjects were invited through a university's bulk e-mail service. The subjects signed consent forms, were seated at a computer station, and were assigned a unique identification number. The session monitor gave some basic instructions and the experiment started. Subjects earned a $10 show up payment plus $20, on average, in additional earnings. The experiment lasted approximately an hour.

3 Hypotheses

Akerlof and Kranton (2000) introduced identity and self-image into a utility framework where one's own actions and the actions of others are directly parameterized in the utility function. Benjamin et al. (2010) add that inducing a particular social category has a marginal effect for increasing the strength of affiliation with an identity category. The utility function is then characterized by considering the strength of affiliation an individual has with a certain social category. The strength of affiliation to a social category drives an individual's decision to either engage with an activity as it aligns to a specific social category or engage in an activity that is opposite of in what a particular social category chooses to engage.

In our study, we induce three statuses based on the performance of a quiz during stage 1. We assume that the individuals will attain some utility from the earnings in the task. They also gain utility from the offered benefit minus whatever social stigma is associated with that benefit. We hypothesize that this social stigma is greater when there is no plausible deniability, that is, only individuals in the low-status group may receive the benefit.

Hypothesis 1. *Take-up of benefits is identical between treatments.*

[4]Power calculations based on Friedrichsen et al. (2018), estimates N=266 from their random and quiz treatments for a 0.8 power and alpha of 0.05.

Hypothesis 1a. *Take-up of benefits is greater for low-status individuals in the plausible deniability treatment where they may plausibly claim they are middle status.*

In the second stage of the experiment subjects will either gain utility from choosing the piece-rate payment scheme or from entering competition and winning it. However, individuals also gain or lose utility from entering the competition based on the prescribed notions of their assigned status and its implications for competitive decisions. More specifically, the unrelated assignment for the high-status groups gives license to engage in competitive behavior while it penalizes it for the low-status individuals. Between treatments, the middle status group's behavior remains an interesting question. Previous literature suggests that individuals seek to mimic high-status individuals to generate a sense of belonging to a higher status (pecuniary emulation); however, when those who already attained a status see that the status is becoming more widely available, they will seek to distinguish themselves to a higher status (invidious comparison) (Bursztyn and Jensen, 2017). Generally, we expect that the middle-status group will compete at a level between the rate of the high-status group and the low-status group. Our theoretical framework predicts that individuals are susceptible to the status assignment, and they will choose to compete according to the induced status during stage 1. Based on this simple framework, we present Hypothesis 2.

Hypothesis 2. *There is no difference in the competitive behavior in a subsequent task between the three status groups.*

Hypothesis 2a. *Competitive behavior is positively correlated with status. High groups enter competition most often, then middle groups, then low.*

By introducing plausible deniability, we expect that the low-status group may use this setting as a justification and respond by increasing their rate of competition relative to the control. Plausible deniability affords this group the opportunity to reduce the impact of the benefit claim eligibility to their status identity. We expect an opposite behavior for the middle status group. In the control, only the low-status group is eligible for the benefit transfer. However, in the plausible deniability treatment when the middle-status group is eligible for the benefit, it makes them experience the stigma associated with the benefit. We expect that the competitive rates of the middle-status group will decrease with the benefit eligibility in the Plausible Deniability Treatment. The high-status individuals are unaffected by the benefit eligibility in the control and treatment conditions and hence we expect that their rates of competition will remain the same across these conditions.

Hypothesis 3. *There is no change in competitiveness for each group when comparing the control with the plausible deniability treatment.*

Hypothesis 3a. *Low status individuals will increase their competitive behavior relative to the control.*

Hypothesis 3b. *Middle status individuals will decrease their competitive behavior relative to the control.*

4 Results

Table 2 provides summary statistics of the main variables of our experiment both overall and separated by treatment. A key dependent variable is an indication to take benefits. In Stage 1 of the experiment—after subjects had completed their general knowledge quiz, but before they learned their score or standing in the group—subjects were asked, using the strategy method, whether they would take an additional $1 if they were eligible. In the control treatment the $1 is only available to the lowest (of three) performers in the group, but in the Plausible Deniability Treatment it is available to the lowest two. At first glance, it appears the Plausible Deniability Treatment is effective. More subjects indicate they would take the benefit (93.5%) under the treatment than under the control (86.3%). Both session

Table 2: Summary statistics of main explanatory variables overall and separated by treatment.

	Overall	Control	Plausible Deniability
will take benefit	0.899	0.870	0.928
(1=yes)	(0.302)	(0.338)	(0.260)
enters tournament	0.500	0.529	0.471
(1=yes)	(.0.501)	(0.501)	(0.501)
benefit eligible	0.500	0.333	0.667
(1=yes)	(0.501)	(0.473)	(0.473)
female	0.569	0.594	0.543
(1=yes)	(0.496)	(0.493)	(0.500)
stage 1 score	7.572	7.572	7.572
(questions correct)	(2.014)	(2.007)	(2.029)
stage 2 score	8.514	8.993	8.036
(questions correct)	(3.533)	(3.308)	(3.695)
tournament belief	2.402	2.420	2.384
(3=top performer)	(0.560)	(0.577)	(0.545)
expectation of group take-up	0.794	0.792	0.796
(0-1 range)	(0.265)	(0.263)	(0.267)
shyness score	2.838	2.903	2.773
(1-5 range)	(0.782)	(0.809)	(0.752)
behindness aversion	0.370	0.377	0.362
(0-1 range)	(0.484)	(0.486)	(0.482)
aheadness aversion	0.409	0.420	0.399
(0-1 range)	(0.493)	(0.495)	(0.491)
risk aversion	5.138	5.138	5.138
(safe choices on Holt-Laury)	(2.062)	(2.001)	(2.128)
year of study	2.467	2.428	2.507
(1-5 range)	(1.500)	(1.489)	(1.515)
observations	276	138	138
subjects	276	138	138
groups	92	46	46
sessions	31	14	17

level t-tests and rank sum tests (N=31), indicate some degree of statistical significance for this result ($p < 0.05$, $p < 0.1$, respectively).

However, the plausible deniability hypothesis (Hypothesis 1A) predicted a specific channel on how this take-up would occur. Low-performing subjects, now feeling they could hide their low status, would be more inclined to take benefits. As Figure 1 indicates, this is not the case. If anything, the lowest-performing subjects (based on decile of quiz performance) are *less likely* to take-up benefits under plausible deniability. The increase in benefits appears to flip around the third decile, the exact point where subjects (in expectation) would no longer be the lowest of three performers.

Table 3 provides three regressions that examine these relationships under more detail. Specification (1) which only includes the treatment variable, confirms our earlier result, indicating the plausible deniability treatment increases overall take-up by 6 percentage points ($p < 0.10$). Specification (2) provides similar results, controlling for the overall effect of quiz performance using percentile rank of subjects' score. Specification (3) includes a "kitchen-sink" of variables, including surveyed levels of risk-aversion (Holt and Laury, 2002), aheadness and behnindness aversion (Bartling et al., 2009), gender identification as female, and year of study at the university (1-5). None of these variables are predictive of take-up. However, scores from a scale of "shyness" (Cheek and Buss, 1981) are negatively correlated with take-up, the direction one would expect. Additionally, one's expectation about the take-up behavior of others (i.e., a subject's prediction of how many of the low performing subjects would take the benefit) is also predictive of take-up decision, consistent with norm-compliance behavior.

Appendix Table A.1 shows similar regressions to Table 3 with subject status from the quiz interacted with the treatment variable. The regressions show that the increase in take up due to the Plausible Deniability Treatment is driven by subjects that will be classified as middle-status. Interestingly, these subjects would receive a \$1 benefit under Plausible Deniability but not in the Control. Though subjects cannot know their quiz performance

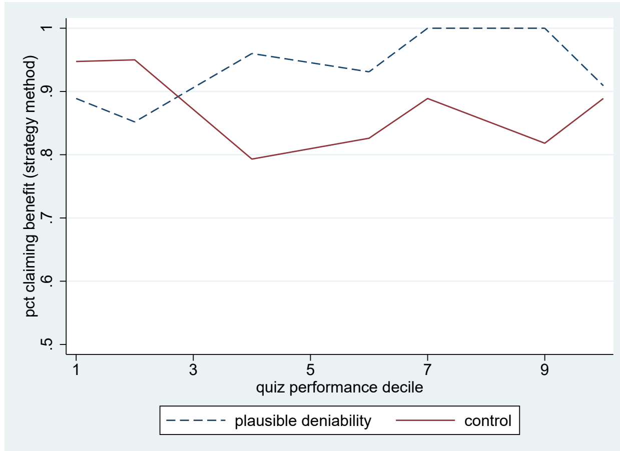

Figure 1: Percent of subjects opting to receive additional \$1 benefit if eligible strategy method in Control and Plausible Deniability Treatments by quiz performance decile. In the Control, only the lowest performer of three subjects will receive the benefit. The lowest two performers will receive the benefit in Plausible Deniability Treatment.

score nor standing in the group precisely when their take-up decision is made, it appears they are more likely to respond to the possibility as it becomes increasingly likely.[5]

Appendix Table A.2 repeats the analysis of Table 3, but restricted to low performance subjects. Across all three specifications, it appears the effect of plausible deniability on low performance subjects is statistically (as well as numerically) zero. There appears to be no plausible deniability effect as predicted by Hypothesis 1A. Thus, while there is a statistical basis to reject null Hypothesis 1, absent any plausible alternative, Hypothesis 1 remains the most plausible option.

Result 1. *The Plausible Deniability Treatment increases overall rates of take up, but not in the way hypothesized. Subjects with better quiz performance are more likely to increase take-up under plausible deniability treatment. There is no evidence of increased take-up among the low-status group.*

To test the other hypotheses in this paper, it is crucial that performance in stage 1

[5]This result would appear to be at odds with most theoretical models of monotonicity. If mean subject expectations of benefit eligibility are p_c in the control and p_t under plausible deniability, a subject should claim a the same rate as long as $p_c, p_t > 0$. However, if $p_t > p_c$ a basic model of bounded rationality could accommodate the difference in take-up. Essentially, subjects are optimizing more often when non-zero costs of sub-optimal play are more likely to be realized.

Table 3: Regression analysis of indicated willingness to take benefit (strategy method) on treatment, stage 1 quiz performance, elicited preferences, and demographics.

	(1) will take benefit	(2) will take benefit	(3) will take benefit
plausible deniability	0.058*	0.058**	0.046*
	(0.030)	(0.029)	(0.027)
percentile of quiz score		0.000	0.001
		(0.001)	(0.001)
shyness scale			-0.085***
			(0.023)
expectation of group take-up			0.484***
			(0.085)
risk aversion			-0.007
			(0.007)
aheadness aversion			-0.014
			(0.035)
behindness aversion			-0.028
			(0.039)
female			0.009
			(0.033)
year of study			-0.005
			(0.010)
constant	0.870***	0.861***	0.768***
	(0.025)	(0.032)	(0.108)
observations	276	276	276
number of groups	92	92	92
number of sessions	31	31	31
r^2	0.010	0.010	0.253

*** p<0.01, ** p<0.05, * p<0.1.

Notes: All three regressions use random effects terms for subject group, and use cluster-robust standard errors at the session level.

on a general knowledge quiz is not predictive of performance in stage 2's summation task. Otherwise subjects' response to status on task 1 may reflect a rational expectation of stage 2 performance. Luckily, performance on either task does not appear to be correlated. Table 4 provides a two-way counts table of subject tercile in the stage 2 summation (math) task and status in the stage 1 (general knowledge quiz) task. A chi-square test does not reject the null hypothesis that distribution is random ($p \approx 0.582$).

Figures 2(a) and (b) show general knowledge quiz scores (stage 1) with performance on summation tasks in the paid round and tournament round (stage 2), respectively. Neither

Table 4: Two-way contingency table of general knowledge quiz status (stage 1) and summation math task tercile (stage 2). Binning does not appear to differ from random ($\chi^2 = 2.858$, $p \approx 0.582$).

stage 1 assigned status	stage 2 group performance 1st tercile	2nd tercile	3rd tercile	total
low	42	25	25	92
medium	32	31	29	92
high	34	32	26	92
total	108	88	80	276

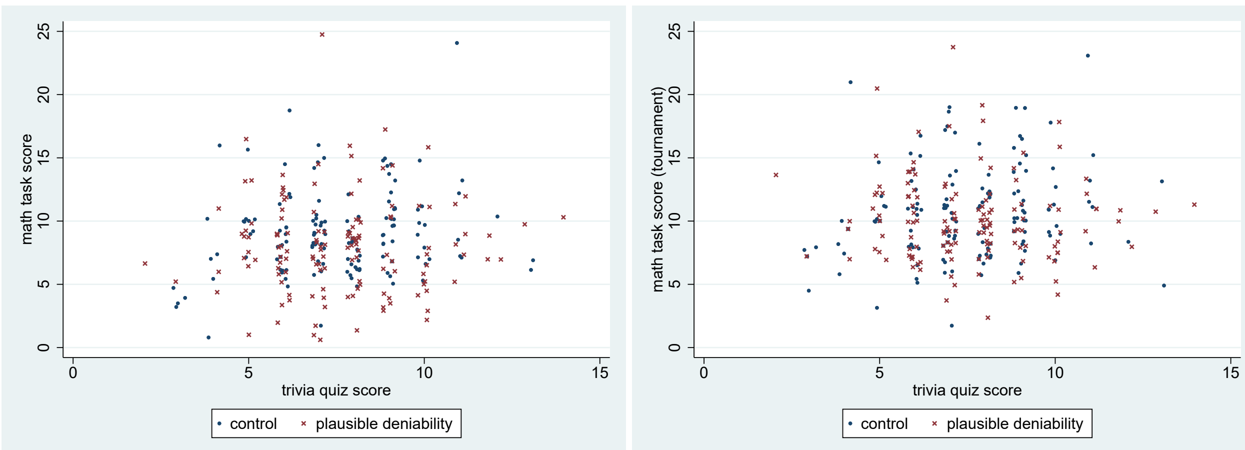

Figure 2: Scatterplot of trivia quiz score on math task score (left) ($\rho \approx 0.079$, $p \approx 0.191$) and the second math task score used for tournaments (right) ($\rho \approx 0.046$, $p \approx 0.447$).

figure appears to show any correlation between quiz scores and summation performance. Table 5 provides four regression specifications, two with dependent variable being rank in group on the math task (i.e., 1st=3, 2nd=2, 3rd=1), two with the dependent variable of being in the top rank in the math task (i.e., 1st=1, 2nd=0, 3rd=0). Using status in the general knowledge quiz as well as actual number of answers correct in the stage 1 general knowledge quiz, we see no significant correlations with stage 2 performance.

Result 2. *There is no evidence that stage 1 performance or status is predictive of rank in group in stage 2.*

The main research question in this paper is whether a stigma induced by benefits for low performance in one task can carryover to reducing competitiveness on a second, unrelated,

Table 5: Regression of stage 2 rank (specifications (1), (3)) and top rank (specifications (2), (4)) on assigned stage 1 status (specifications (1), (2)) and number of quiz questions correct, stage 1, (specifications (1), (2)). There is little evidence of correlation between stage 1 and 2 variables.

	(1) stage 2 group rank[a] (tercile)	(2) top rank (tercile)	(3) stage 2 group rank (tercile)	(4) top rank (tercile)
low status	-0.152	-0.043		
	(0.143)	(0.075)		
high status	-0.054	-0.033		
	(0.119)	(0.078)		
stage 1 score			0.019	-0.001
			(0.024)	(0.014)
constant	1.967***	0.315***	1.756***	0.295***
	(0.079)	(0.050)	(0.179)	(0.109)
observations	276	276	276	276
number of groups	92	92	92	92
number of sessions	31	31	31	31
r^2	0.006	0.002	0.002	0.000

*** p<0.01, ** p<0.05, * p<0.1.

Notes: All three regressions use random effects terms for subject group, and use cluster-robust standard errors at the session level.

[a] To preserve sign, higher performance is associated with a higher number in rank. That is, within a group, 3 is top rank and 1 is bottom rank.

task. At first glance, the answer appears to be definitively positive. Of the subjects that receive benefits in the first stage of the treatment (i.e., all "low" performance and "middle" under the Plausible Deniability treatment) only 34.8% (48 of 138) of subjects compete in the second round tournament compared to 65.2% (90 of 138) in the other groups. The difference is statistically significant under both session-level rank sum and t-tests ($N = 31$, $p < 0.01$, both tests). Of course, this comparison does not differentiate between whether the stigma of "low status" vs. the stigma of benefits is ultimately responsible for the reduced competitiveness. To isolate the effects of eligibility for benefits on competitive behavior, we must examine the middle group across the two treatments. Among those in the middle group, 27 of 46 (58.7%) in the control and 18 of 46 (39.1%) in the Plausible Deniability Treatment compete in the subsequent task. The result is marginally significant under a *t*-test ($p < 0.1$) but not under a non-parametric rank sum ($p \approx 0.129$), both run at the session level ($N = 31$).

To control for the explanatory power of other variables, Table 6 provides regressions of the binary decision to compete in the tournament in stage 2 on status from the stage 1 task. As expected, earning low status on the first task reduces competitiveness on the second task by 19–26 percentage points ($p < 0.01$) relative to a middle-status subject in the control treatment (who is not eligible for benefits). Low-status subjects, regardless of treatment, are always eligibility for additional benefits in stage 1. Interestingly, this effect is somewhat mitigated, but not eliminated, in regressions that control for beliefs on tournament performance. This relation suggests the status effect works in more ways than just through subjects' expectations about future performance. In contrast, high status does not significantly differ from middle status under the control and its sign is nearly reduced to zero when controlling for other variables. Neither high nor low status significantly differs across treatments. However, middle-status subjects are 17–20 percentage points less likely to enter the tournament in the subsequent stage 2 task under the Plausible Deniability Treatment ($p < 0.1$). Among the control variables, notably female gender, as expected, reduces competitiveness ($p < 0.01$). For each successively higher rank expected in the tournament (of 3), subjects are 23-26 percentage points more likely to enter the tournament ($p < 0.01$). Unsurprisingly, better performance on the same task under a piece-rate structure also increases competitiveness. For each additional year of student study, subjects are 4 percentage points more likely to enter the competition ($p < 0.1$).

Taken together, a useful simplification of these result may be that middle-status subjects under the control compete like high-status subjects. In contrast, middle-status subjects under the Plausible Deniability Treatment behave like low-status subjects. Of course this categorization is identical to the grouping of eligibility of benefits after stage 1. In general, the results are highly supportive of Hypothesis 2A over Hypothesis 2 as tournament entry is positively correlated with stage 1 status. Further as low status subjects to not act differently under the treatments, but middle-status subjects do, we reject Hypothesis 3 in favor of Hypothesis 3B.

To further examine the way stage 1 benefits may have affected subjects tournament entry decision in stage 2, we take a closer look at tournament beliefs. Before making the decision to enter the stage 2 tournament each subject gave their belief in their performance in the 3-person group should they enter the tournament. Table 7 shows the distribution of overall beliefs and split across treatments. Status appears to have an effect on subsequent beliefs. Few subjects believe they will finish last in the tournament. Most high-status subjects believe they will finish first. In contrast, most low-status subjects believe they will finish 2nd. Middle-status subjects fall somewhere in between. Appendix tables A.3 and A.4 show this table separated by treatment. The tables are mostly similar though there appears to be a small effect of less confidence of middle-status subjects in the Plausible Deniability Treatment.

Table 8 shows regressions of expected rank in tournament (i.e., 1st=3, 2nd=2, 3rd=1) and expectation of top rank on status and the treatment variable. Consistent with other results, subjects assigned low status in stage 1 expect to finish 0.2 ranks lower than subjects assigned

Table 6: Regression analysis of tournament entry (stage 2) on treatment, stage 1 and 2 performance, elicited preferences, and demographics.

	(1) entry into tournament	(2) entry into tournament	(3) entry into tournament
low status	-0.261***	-0.212***	-0.188***
	(0.091)	(0.067)	(0.066)
high status	0.087	0.043	0.013
	(0.115)	(0.093)	(0.105)
plausible deniability ×	0.022	0.009	0.026
low status	(0.096)	(0.096)	(0.103)
plausible deniability ×	0.000	0.003	0.009
high status	(0.105)	(0.096)	(0.096)
plausible deniability ×	-0.196	-0.179*	-0.174*
middle status	(0.119)	(0.101)	(0.100)
female		-0.142***	-0.129***
		(0.045)	(0.046)
stage 1 quiz			0.015
questions correct			(0.020)
stage 2 non-			0.015**
tournament performance			(0.007)
believed tournament		0.266***	0.232***
tercile		(0.044)	(0.053)
shyness			-0.014
			(0.038)
expectation of group			0.070
take-up			(0.098)
risk aversion			0.008
			(0.015)
aheadness aversion			0.009
			(0.059)
behindness aversion			0.016
			(0.069)
year of study			0.038*
			(0.022)
constant	0.587***	0.025	-0.310
	(0.094)	(0.152)	(0.215)
observations	276	276	276
number of groups	92	92	92
number of sessions	31	31	31
r^2	0.099	0.217	0.247

*** p<0.01, ** p<0.05, * p<0.1.

Notes: All three regressions use random effects terms for subject group, and use cluster-robust standard errors at the session level.

Table 7: Two-way contingency table of general knowledge quiz status (stage 1) and expected performance in 3-person tournament (stage 2). Binning significantly differs from random ($\chi^2 = 16.819$, $p < 0.01$).

	stage 2 tournament belief			
stage 1 assigned status	1st tercile	2nd tercile	3rd tercile	total
low	5	59	28	92
middle	3	51	38	92
high	2	35	55	92
total	10	145	121	92

Table 8: Regression analysis of belief in expected rank in tournament (specification (1)) and top rank expectation (specification (2)) on treatment interacted with status.

	(1) expected rank in tournament (tercile)	(2) expected top rank
low status	-0.196*	-0.174**
	(0.117)	(0.088)
high status	0.152	0.130
	(0.133)	(0.117)
plausible deniability ×	-0.022	-0.022
a high status	(0.100)	(0.097)
plausible deniability ×	0.022	-0.000
low status	(0.122)	(0.101)
plausible deniability ×	-0.109	-0.130
middle status	(0.131)	(0.120)
constant	2.435***	0.478***
	(0.105)	(0.089)
observations	276	276
number of groups	92	92
number of sessions	31	31
r^2	0.061	0.066

*** p<0.01, ** p<0.05, * p<0.1.

Notes: All three regressions use random effects terms for subject group, and use cluster-robust standard errors at the session level.

[a] To preserve sign, higher performance is associated with a higher number in rank. That is, within a group, 3 is top rank and 1 is bottom rank.

middle status in the control treatment ($p < 0.1$). They are also 17.5 percentage points less likely to expect to finish first ($p < 0.05$). Though the sign is positive with an economically significant magnitude, high-status subjects do not differ significantly in expectations from middle-status subjects in the control. Middle-status subjects under the Plausible Deniability Treatment have significantly lower expectations than the high-status subjects ($p < 0.01$, both specifications), but do not significantly differ from the low-status subjects ($p \approx 0.5, 0.7$). It appears expectations of tournament performance may be one driver of our main results, however, the results in specifications (2) and (3) in the regressions of Table 6 show they are not the sole drivers of our observed effects.

Result 3. *Eligibility for benefits—assigned based on stage 1 performance—affects subsequent, stage 2, competitive behavior. While benefits eligibility may be entangled with status assignment for the low-status group, among the middle-status group, both competitive expectations and decisions resemble the high-status group in the control and low status group in the Plausible Deniability Treatment.*

4.1 Emotions Data

We collected data on our subjects emotional responses using Affectiva through the iMotions software. Affectiva is an AI software based on a large data set and algorithms trained to recognize facial expressions through webcam recordings (Senechal et al., 2015). The emotions data captures seven different emotions on a scale from 0 to 100, which represents the likelihood that it is the emotion expressed in the facial reading. There is also a valence measure, which indicates an overall positive, negative, and neutral emotion. The number for valence ranges from -100 (negative) to 100 (positive), where 0 represents neutral. Data were collected at points stages during the experiment (see Table 9). Appendix Table A.5 provides summary statistics for each of the variables over the 11 stages of the experiment.

It is impractical to observe treatment effects individually across each one of our biometric variables. For one, the variables may be highly correlated and, absent any ex-ante hypothe-

Table 9: Description of each stage of emotions observation during the experiment.

period	description	number of subject observations
0	Introduction	248
1	Group Assignment	248
2	Eligibility	246
3	Results	242
4	Belief Elicitation	247
5	Token Claim	248
6	Compensation Screen	243
7	Performance Beliefs	232
8	Stage 2 Task	239
9	Shyness Scale	238
10	Earnings Disclosure	227

Table 10: Emotions structure from principal factor analysis.

variable	f1 ("positive emotion")	uniqueness
valence	0.712	0.493
joy	0.450	0.798
surprise	-0.129	0.983
contempt	-0.178	0.968
anger	-0.316	0.900
sadness	-0.407	0.835
disgust	-0.189	0.964
fear	-0.031	0.999
pupil	0.092	0.992

ses, it is difficult to explain a significant difference of one biometric variable and not other. Additionally, the comparison of over ten variables may yield false significance simply due to the number of multiple comparisons. Since we do not have hypotheses tied to each of these variables, we use factor analysis to generate a single factor to express the changes in the ten variables between experiments and across stages of the experiment (Table 10). The factor we identify is highly correlated with positive valence and joy, but negatively correlated with contempt, anger, sadness, and disgust. We view this factor as a "positive emotional" reaction.

As we look at the time series of this positive emotional factor across the stages of the

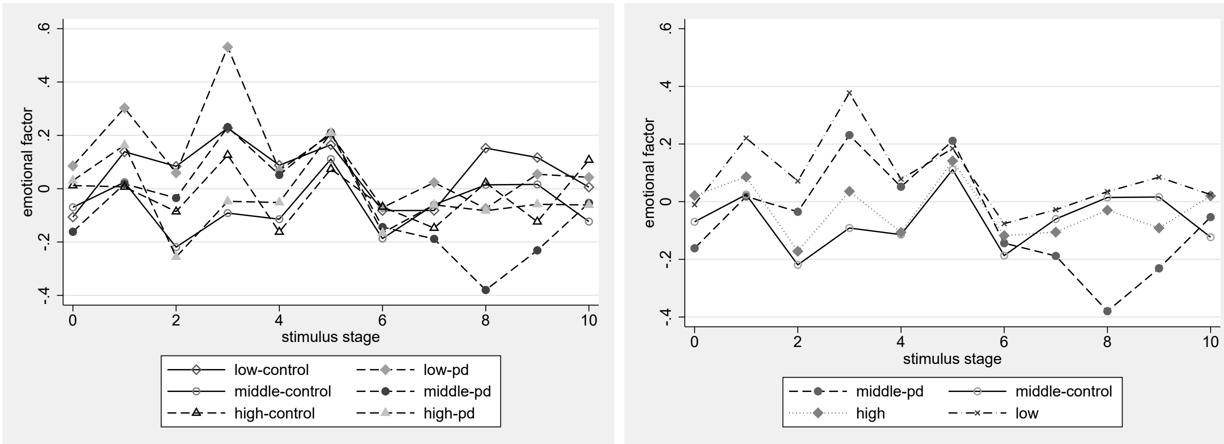

Figure 3: Time-series of mean positive emotional factor over the 11 stages of stimulus (see Table 9) by subject status and treatment. (a, left) Emotional factor by every status and treatment. (b, right) Emotional factor with high and low statuses merged across treatments.

experiment, we note a few things. First, compared to middle-status, emotional factor does not differ as much across treatments for low and high status subjects. Low-status subjects appear to have a higher factor than high status subjects during the first half of the experiment (i.e. before benefits claim, period 5) and then express emotions at similar levels. A similar separation can be observed for middle status subjects in the Plausible Deniability Treatment vs. the control. Overall, it appears that subjects that are eligible for benefits express more positive emotion during the first half of the experiment. Interestingly this emotion peaks around period 3 when subjects learn the results of the stage 1 competition and benefits elligibility.

Table 11 provides regressions to identify these differences across treatment. Subjects that are eligible to receive benefits have 0.15 point higher positive emotional factor over the first half of the experiment (before receiving benefits in stage 5), this disparity is effectively zero in the second half of the experiment. Indicating no observable emotional difference between the eligible subjects that are less likely to compete in the tournament (see Result 3) than the subjects without eligibility.

Result 4. *Subjects that are eligible to receive benefits based on their low-performance ex-*

Table 11: Regression analysis of positive emotional factor on eligibility for benefits.

	(1) positive emotional factor
eligible for	0.161**
benefits	(0.074)
post benefit	-0.008
take-up (stage 5)	(0.049)
eligible × post	-0.168**
benefit take-up	(0.079)
constant	-0.056
	(0.058)
observations	2,410
number of subjects	249
number of sessions	31
r^2	0.008

*** p<0.01, ** p<0.05, * p<0.1.
Notes: All three regressions use random effects terms for subject, and use cluster-robust standard errors at the session level.

hibit more positive emotions than those who do not qualify for benefits in the first stage of the experiment. There is no difference in emotions expressed over the second stage of the experiment.

5 Discussion

We implement a laboratory experiment where we can control for factors that are normally correlated in observational data, specifically the provision and stigma of social welfare benefits and subsequent labor entry decisions. We create an environment where the lowest-performing third of subjects can supplement their earnings with an additional benefit, however, that choice is directly observed. In one treatment, we allow this group to have "plausible deniability" by expanding the benefits eligibility to include the middle-status group. Surprisingly, we find no effect on benefit take-up of the lowest-performing group. However, in eliciting competitiveness in a subsequent and unrelated task we find competitiveness of the middle-status group decreases in the Plausible Deniability Treatment. Other related measures such as performance ability, confidence and risk do not explain the difference in the rates of competitiveness. This result may imply that being exposed to social identity status in social benefit eligible programs may have significant effects on the later preferences for competition of the subjects.

Eligibility for the social benefit is tied to a low-performing perception that seems to affect subsequent competitive decisions. Choosing to compete at a lower rate due to being in a low status is consistent with findings in the literature. Specifically, being eligible for an additional government benefit is associated with experiencing poverty or financial scarcity. Inducing poverty or financial constraints has been found to be detrimental for the cognitive performance of the poor, but not the rich (Mani et al., 2013; Hoff and Walsh, 2018). Furthermore, recent evidence points to lower propensity to choose a challenging task with higher payoffs or experience financial avoidance when facing financial scarcity (Banker et al., 2020; Hilbert et al., 2022). While we may see evidence of pecuniary emulation in the middle status group in the Control treatment, the introduction of plausible deniability reduces their rate of competition entry. The findings from our study add to this literature focusing on expanding eligibility to the middle group and discovering that being eligible for an additional benefit is tied to lower entry into competition even when controlling for ability on the same task across status levels. Further, our study also adds to this discussion by recording mood of subjects. Interestingly, if anything, benefits eligibility appears to increase the positive emotional state of subjects. There is no evidence that subjects' reduced propensity to enter competitions is correlated with negative emotions. This may imply that emotions like complacency rather than those like discouragement are responsible for any subsequent reduced competitiveness of persons that receive social benefits.

The findings from this controlled setting may be considered for a broader context. Much of the literature focuses on the behavior of those who are high status or are low status, but literature on the group in the middle is sparse. However, this group is the most susceptible when considering policy changes that can affect eligibility into government assistance programs. Recent economic shocks such as recessions and pandemics affect eligibility rules, which generally lowers the barriers to partake in government benefits and expands eligibility (Ganong and Liebman, 2013; on Budget and Priorities, 2022). Frequently, individuals who are eligible for government assistance programs are faced with requirements for job training or providing evidence of a job search to continue qualifying for these programs (on Budget and Priorities, 2022). However, with the findings from this study, we note the importance of considering how the response of being low-income may impact the nature of a job search, particularly how the stigma of being benefit-eligible may impede individuals from seeking more competitive opportunities that can improve their own outcomes.

Further research into this area would explore the underlying mechanisms and motivations for entering competition, particularly when a social status is induced. A field study that captures the essence of the findings in this paper could further support our findings and inform policy for individuals who are benefit eligible. Understanding the underlying behavioral responses stemming from benefit-eligibility stigma may enlighten approaches to alleviate poverty.

References

Akerlof, G. A. and Kranton, R. E. (2000). Economics and identity. *The Quarterly Journal of Economics*, 115(3):715–753.

Andrade, C. (2002). The economics of welfare participation and welfare stigma: A review. *Public Finance and Management*, 2(2):294–333.

Ball, S., Eckel, C., Grossman, P. J., and Zame, W. (2001). Status in markets. *The Quarterly Journal of Economics*, 116(1):161–188.

Banker, S., Bhanot, S. P., and Deshpande, A. (2020). Poverty identity and preference for challenge: Evidence from the us and india. *Journal of Economic Psychology*, 76:102214.

Bartling, B., Fehr, E., Maréchal, M. A., and Schunk, D. (2009). Egalitarianism and competitiveness. *American Economic Review*, 99(2):93–98.

Benjamin, D. J., Choi, J. J., and Strickland, A. J. (2010). Social identity and preferences. *American Economic Review*, 100(4):1913–28.

Bolton, G., Dimant, E., and Schmidt, U. (2021). Observability and social image: On the robustness and fragility of reciprocity. *Journal of Economic Behavior & Organization*, 191:946–964.

Bursztyn, L., Egorov, G., and Jensen, R. (2019). Cool to be smart or smart to be cool? understanding peer pressure in education. *The Review of Economic Studies*, 86(4):1487–1526.

Bursztyn, L., Ferman, B., Fiorin, S., Kanz, M., and Rao, G. (2017). Status Goods: Experimental Evidence from Platinum Credit Cards*. *The Quarterly Journal of Economics*, 133(3):1561–1595.

Bursztyn, L. and Jensen, R. (2015). How does peer pressure affect educational investments? *The Quarterly Journal of Economics*, 130(3):1329–1367.

Bursztyn, L. and Jensen, R. (2017). Social image and economic behavior in the field: Identifying, understanding, and shaping social pressure. *Annual Review of Economics*, 9:131–153.

Butera, L., Metcalfe, R., Morrison, W., and Taubinsky, D. (2022). Measuring the welfare effects of shame and pride. *American Economic Review*, 112(1):122–68.

Cheek, J. M. and Buss, A. H. (1981). Shyness and sociability. *Journal of personality and social psychology*, 41(2):330.

Clingingsmith, D. and Sheremeta, R. M. (2018). Status and the demand for visible goods: Experimental evidence on conspicuous consumption. *Experimental Economics*, 21(4):877–904.

Currie, J., Grogger, J., Burtless, G., and Schoeni, R. F. (2001). Explaining recent declines in food stamp program participation [with comments]. *Brookings-Wharton papers on urban affairs*, pages 203–244.

Daponte, B. O., Sanders, S., and Taylor, L. (1999). Why do low-income households not use food stamps? evidence from an experiment. *The Journal of Human Resources*, 34(3):612–628.

Friedrichsen, J., König, T., and Schmacker, R. (2018). Social image concerns and welfare take-up. *Journal of Public Economics*, 168:174–192.

Ganong, P. and Liebman, J. B. (2013). The decline, rebound, and further rise in snap enrollment: Disentangling business cycle fluctuations and policy changes. Working Paper 19363, National Bureau of Economic Research.

Gillies, A. S., Rigdon, M. L., et al. (2019). Plausible deniability and cooperation in trust games. *Review of Behavioral Economics*, 6(2):1–95â.

Hilbert, L. P., Noordewier, M. K., and van Dijk, W. W. (2022). The prospective associations between financial scarcity and financial avoidance. *Journal of Economic Psychology*, 88:102459.

Hoff, K. and Walsh, J. (2018). The whys of social exclusion: Insights from behavioral economics. *The World Bank Research Observer*, 33(1):1–33.

Holt, C. A. and Laury, S. K. (2002). Risk aversion and incentive effects. *American economic review*, 92(5):1644–1655.

Ibarra, H., Carter, N. M., Silva, C., et al. (2010). Why men still get more promotions than women. *Harvard Business Review*, 88(9):80–85.

Kassas, B. and Palma, M. A. (2019). Self-serving biases in social norm compliance. *Journal of Economic Behavior & Organization*.

Major, B. and O'brien, L. T. (2005). The social psychology of stigma. *Annu. Rev. Psychol.*, 56:393–421.

Mani, A., Mullainathan, S., Shafir, E., and Zhao, J. (2013). Poverty impedes cognitive function. *science*, 341(6149):976–980.

Minton, S. and Giannarelli, L. (2019). Five things you may not know about the us social safety net. *Washington, DC: Urban Institute.*

Moffitt, R. (1983). An economic model of welfare stigma. *The American Economic Review*, 73(5):1023–1035.

Niederle, M. and Vesterlund, L. (2007). Do women shy away from competition? do men compete too much? *The quarterly journal of economics*, 122(3):1067–1101.

on Budget, C. and Priorities, P. (2022). A quick guide to snap eligibility and benefits.

Palma, M. A., Ness, M. L., and Anderson, D. P. (2017). Fashionable food: a latent class analysis of social status in food purchases. *Applied Economics*, 49(3):238–250.

Senechal, T., McDuff, D., and Kaliouby, R. (2015). Facial action unit detection using active learning and an efficient non-linear kernel approximation. In *Proceedings of the IEEE International Conference on Computer Vision Workshops*, pages 10–18.

Shih, M., Pittinsky, T. L., and Ambady, N. (1999). Stereotype susceptibility: Identity salience and shifts in quantitative performance. *Psychological science*, 10(1):80–83.

Veblen, T. (1899). Mr. cummings's strictures on" the theory of the leisure class". *Journal of Political Economy*, 8(1):106–117.

Zhang, P., Zhang, Y., and Palma, M. (2020). Social norms and competitiveness: My willingness to compete depends on who i am (supposed to be). *Available at SSRN 4010949.*

A Additional Tables

	(1) will take benefit	(2) will take benefit
low status	-0.022	-0.042
	(0.059)	(0.057)
high status	-0.043	-0.049
	(0.070)	(0.050)
plausible deniability × low status	-0.000	0.011
	(0.059)	(0.063)
plausible deniability × high status	0.087	0.083
	(0.064)	(0.056)
plausible deniability × medium status	0.087**	0.042
	(0.044)	(0.040)
shyness scale		-0.080***
		(0.022)
expectation of group take-up		0.483***
		(0.087)
risk aversion		-0.007
		(0.007)
aheadness aversion		-0.019
		(0.037)
behindness aversion		-0.032
		(0.039)
female		0.007
		(0.033)
year of study		-0.004
		(0.010)
constant	0.891***	0.812***
	(0.038)	(0.118)
observations	276	276
number of groups	92	92
number of sessions	31	31
r^2	0.038	0.259

*** p<0.01, ** p<0.05, * p<0.1.

Notes: All three regressions use random effects terms for subject group and cluster-robust standard errors at the session level.

Table A.1: Regression analysis of indicated willingness to take benefit (strategy method) on treatment interacted with status, stage 1 quiz performance, elicited preferences, and demographics.

	(1) will take benefit	(2) will take benefit	(3) will take benefit
plausible deniability	0.000	0.004	0.011
	(0.059)	(0.059)	(0.065)
percentile of quiz score		-0.002	-0.001
		(0.002)	(0.002)
shyness scale			-0.136**
			(0.053)
expectation of group take-up			0.393**
			(0.182)
riskaversion			-0.009
			(0.017)
aheadness aversion			0.057
			(0.057)
behindness aversion			-0.118**
			(0.059)
female			0.008
			(0.060)
year of study			0.028
			(0.022)
constant	0.870***	0.905***	0.951***
	(0.043)	(0.049)	(0.306)
observations	92	92	92
number of groups	92	92	92
number of sessions	31	31	31
r^2	0.000	0.010	0.258

*** p<0.01, ** p<0.05, * p<0.1.

Notes: All three regressions use random effects terms for session, and use cluster-robust standard errors at the session level.

Table A.2: Regression analysis of indicated willingness to take benefit (strategy method) on treatment, stage 1 quiz performance, elicited preferences, and demographics. The analysis is restricted to low status subjects only.

	stage 2 tournament belief			
stage 1 assigned status	1st tercile	2nd tercile	3rd tercile	total
low	3	29	14	46
medium	2	51	22	46
high	1	17	28	46
total	6	68	64	138

Table A.3: Two-way contingency table of general knowledge quiz status (stage 1) and expected performance in 3-person tournament (stage 2), control treatment only. Binning significantly differs from random ($\chi^2 = 8.831$, $p < 0.1$).

stage 1 assigned status	stage 2 tournament belief 1st tercile	2nd tercile	3rd tercile	total
low	2	30	14	46
medium	1	29	16	46
high	1	18	27	46
total	4	77	57	138

Table A.4: Two-way contingency table of general knowledge quiz status (stage 1) and expected performance in 3-person tournament (stage 2), Plausible Deniability treatment only. Binning significantly differs from random ($\chi^2 = 9.112$, $p < 0.1$).

	Overall	Control	Plausible Deniability
positive emotional factor	0.000	-0.007	0.007
	(0.777)	(0.687)	(0.857)
valence	-6.783	-7.147	-6.424
	(13.552)	(12.606)	(14.422)
joy	1.506	1.253	1.755
	(6.552)	(5.735)	(7.263)
surprise	3.647	3.831	3.464
	(6.738)	(6.695)	(6.777)
contempt	2.747	2.747	3.077
	(7.886)	(7.886)	(8.471)
anger	0.526	0.413	0.637
	(2.599)	(1.561)	(3.317)
sadness	1.047	0.858	1.234
	(4.214)	(3.444)	(4.851)
disgust	1.051	1.142	0.960
	(2.099)	(2.387)	(1.765)
fear	3.752	3.780	3.724
	(8.122)	(8.082)	(8.166)
pupil	2.950	2.966	2.934
	(0.409)	(0.378)	(0.438)
observations	2,691	1,338	1,353
subjects	249	124	125
groups	92	46	46
sessions	31	14	17

Table A.5: Summary statistics of recorded emotion variables overall and separated by treatment. Variables were recorded at up to 11 distinct points of the experiment for each subject (see table 9).