Tasmanian School of Business and Economics
University of Tasmania

Discussion Paper Series N 2018-03

The Demand and Supply for Esteem: An experimental analysis

Paul Blacklow
University of Tasmania, Australia

Amy Beth Corman
University of Melbourne, Australia

Hugh Sibly
University of Tasmania, Australia

ISBN 978-1-925646-44-3

The Demand and Supply for Esteem: An experimental analysis*

Paul Blacklow

Tasmanian School of Business and Economics
University of Tasmania
Hobart 7000 TAS
Australia

Paul.Blacklow@utas.edu.au

Amy Beth Corman

Department of Economics
University of Melbourne
Parkville 3010 VIC
Australia

abcorman@unimelb.edu.au

Hugh Sibly

Tasmanian School of Business and Economics
University of Tasmania
Hobart 7000 TAS
Australia

Hugh.Sibly@utas.edu.au

17 August 2018

Abstract

People enjoy judging and receiving the approval of others. They may modify their behaviour in costly ways to obtain such approval. This paper presents an experiment in which some participants can, at a cost, appear to others to have a better performance on a real effort task than they really do. The only motivation for such an action is esteem seeking. The provision of esteem is also recorded. We measure esteem seeking when participants are facing both high and low performing partners. We model our experiment theoretically: individuals generate income party to undertake consumption but also partly to gain esteem. Our results are consistent with theory: those with low marginal utility of consumption engage in esteem seeking.

JEL Codes: C92 D91
Keywords: Esteem, Image, Laboratory Experiment

* We would like to thank Tim Cason, Yan Chen, Lata Gangadharan, Jim Murphy, Katya Sherstyuk, John Tisdell and Tom Wilkening for comments on earlier versions of this work.

1 Introduction

Casual empiricism suggests that people hold strong opinions on the actions of others, and that their actions are influenced by the opinion of others. People appear to, not only enjoy receiving the approval of others but also enjoy evaluating others. Furthermore judgement and esteem seeking behaviour (or disesteem avoidance) appear to be ubiquitous.

In order to gain esteem from judgemental others people may modify their behaviour. For instance, they may buy more fashionable or different style of clothes than required for utilitarian purposes in the anticipation of receiving praise, or avoiding disapproval, from others. Similarly they may buy a more luxurious or stylish car than they would want for purely transportation purposes. Or they may stay longer at their job than is necessary to achieve the required work. These example suggest that esteem seeking will have important influence on a wide range of a person's economic choices. Rigorously studying this behaviour in the field is, however, problematic. It is difficult to isolate esteem seeking from other motivations. For example, if a person buys a luxury car it unclear whether they are esteem seeking or simply like luxury cars.

In this paper we use an economic experiment to identify and model this behaviour within an economic framework. In our experiment some participants can, by buying costly 'credits', appear to another, randomly-matched, participant to have a better performance on a real effort task than they actually have. The only plausible interpretation of the decision to purchase credits is that the participants are esteem seeking. The experimental design allows us to identify whether being matched with a high vs. low scoring participant influences esteem seeking (purchase of credits). Additionally the design allows us to relate esteem seeking behaviour to a participant's willingness to supply effort (which we interpret as being related to the participant's marginal utility of income).

Our theoretical modelling of participant behaviour in our experiment predicts that only those with a relatively low marginal utility of consumption who will purchase credits. We do indeed find experimentally that only some of our participants purchase credits, and having a low marginal utility of income is an important determinant of that choice. We also find that participants are more likely to purchase credits if their matched participant is low scoring. Gender and psychological characteristics are also found to influence the purchase of credits.

To quantify the provision of esteem, we also elicit feedback on a participant's performance from their matched participant. To mirror the examples discussed above, the matched participant can't identify how much the score they view is due to effort or due to the purchase of credits. In line with our theoretical predictions, the positiveness of feedback increases with the score observed by the matched participant. This is the case even though

the matched participant knows that the matched score could be entirely based on the purchase of credits. This suggests that the purchase of credits does indeed buy a participant esteem. We also find that gender and psychological characteristics are also found to influence the nature of feedback.

By design, we use *standard* theoretical and experiment techniques to analyse esteem seeking and provision. The theory we utilise to form hypotheses of participant behaviour in the experiment is also now standard in experimental/behavioural literature. Participant utility is taken as a quasi linear function of monetary income (material consumption) and received esteem. As noted above, the theory predicts that only participants with a relatively low marginal utility of income purchase credits as a way of boosting esteem.

In the experiment itself, following standard practice, participants are anonymous: it was conducted in a computer based lab with terminals in cubicles. The experimental is designed to be context free. Participants in the experiment are paid in piece rates for a real effort task: the slider task by Gill and Prowse (2012). This real effort task is designed to minimised required skill or knowledge. All these attributes of the experiment are hostile to expression and transmission of esteem. One might think that esteem seeking is based on personal identity and that this could be built around demonstrating skill. If we observe esteem seeking in this environment, we can be confident that it will be observed, and most likely to a greater extent, in environments which are more conducive to esteem seeking (such as those described above). Thus the experiment, by adopting such standard techniques, creates a stringent test for the prevalence of esteem seeking. Additionally, by utilising standard experimental techniques, we facilitate the comparison of our results to other experimental results.

Much of the previous interest in esteem seeking stems from how it might be involved in the private provision of public goods, i.e. generate pro-social behaviour. Brennan and Pettit (2004) argue that individual's concern for esteem is fundamental to the functioning of the economy. They emphasise the role esteem play in decisions in all aspects of life, and the constraints individuals face when seeking to obtain esteem.

Benabou and Tirole (2006) develop a formal theoretical model of a market consisting of individuals who gain esteem by being seen to engage in some form of pro-social activity. Gain utility from monetary private income and from providing the public good; but also gain from the esteem of being perceived as pro-social. The relative weight of utility from consuming private income to providing the public good varies across the population. Provision of the public good acts as a signal to an individual's intrinsic pro-sociality. They consider how monetary rewards may crowd out this signal (and thus pro-social behaviour) by casting doubt on individuals' motivation for the provision of a public good. Ellingsen and Johannesson (2008) also develop a signalling model of esteem. They use their signalling

model to explain a number of experimental findings: in particular why control has a hidden cost, and to explain the outcomes of a number of well-studied games.

More recently authors have begun to focus on the role of esteem seeking in relation to private goods. Akerlof (2015) develops a model of esteem to explain how individuals coordinate on values. Akelof's model differs not only in focus from the signalling models of Benabou and Tirole (2006) and Ellingsen and Johannesson (2008) (values vs the provision of public goods) but on the underlying motivation for esteem seeking assumed in the model. For Akerlof, a person's utility arises from values (assumption 1) and is measure by achievement relative to others (assumption 2) and from self-esteem and esteem received from others (assumption 3). Akerlof cites an extensive psychological literature to support these assumptions. In our theoretical modelling to the experiment, we follow Akerlof's assumption 2 and 3 to assume that people make relative evaluations and that they value esteem.

A number of previous papers have looked for experimental verification of status seeking. Status, which represents a person's relative social ranking, is an important avenue through which are able to receive esteem. It potentially provides a platform to give esteem. Charness et al, (2010, 2014) examine, both experimentally and theoretically, status seeking in a setting where participants are paid using piece rates. Repeated status contest within groups of three, where these is explicit winners and losers from the contest. They find that their participants increase effort when informed about their relative performance, pay to increase their apparent performance, and sabotage others' performance measure when given the opportunity.

Like Charness et al, (2010, 2014) we allow participants the opportunity (by purchasing 'credits') to pay to boost their apparent performance (score). In our paper we use pairwise interactions between participants to identify the determinants of individual conferring and receiving esteem. We allow esteem-seeking to endogenously arise rather than construct status within groups. Not providing experimenter determined signal as to status might be expected to lower participant benefit from esteem seeking, resulting in less of this activity. Again this aspect of our design makes the experimental test for the presence of esteem seeking activities more stringent.

A number of other experimental papers consider the role status plays in markets. Eckel and Ball (1998) and Eckel et. al. (2001) show experimentally that markets work in favour of high status individuals. Morozova (2015) utilises the methodology of Eckel et. al. (2001) to study status in an environment where individuals have a group identity.

Tran and Zeckhauser (2012) report on a field experiment in which communicating rank on a test score increased student performance. A recent paper by Burstyn, et al (2017) uses a set of field experiments to isolate status seeking in the demand for tiered status credit cards. They

find that Platinum credit cards are used for social signalling. These field experiment establish the existence of status seeking in real world settings. Indeed commercial institutions are able to exploit this behaviour in their product development. In contrast, our experiment is relatively context free indicates that that esteem seeking behaviour is not reliant on the presence of institutional structures that confer status. People spontaneously esteem seek in our context free experiment, suggesting this effect is indeed ubiquitous.

Relative performance in a competition has been shown to affect behaviour (Azmat and Iriberri, 2010) and wellbeing (Azmat and Iriberri, 2016). In particular, Azmat and Iriberri (2016) present experimental findings in which providing relative performance information, when participants are paid piece-rates, increases both the performance and the inequality in happiness of participants. In our experiment feedback is not only supplying participants with an objective measure of their relative performance, but also provides the participant with their matched participant opinion of their performance. The latter provides participants an observable and clear measure of esteem provision, and one which we use to analyse participant behaviour.

2 The Experiment

2.1 Experimental Design

The experiment was conducted in five stages. In each stages participants undertook a real effort task: the 'sliders' real effort task developed by Gill and Prowse (2012). This task has the advantage that it poses the same level of difficulty to all participants. Participants were provided with a review of the instructions and a quiz prior to the commencement of each stage.

Stages 1, 3 and 5 are designed to calibrate the experimental findings in stages 2 and 4. Stages 1, 3 and 5 have a similar structure. In these three stages all participants are paid a piece rate of 1 Exp$ for each slider correctly placed, where 1 Exp$ is equal to $0.10 Australian dollars. Participants are only aware of their own effort, and therefore income, in these three stages.

Prior to stage 1, in each session participants are randomly allocated to one of two groups of equal size. The two groups are designated group A and group B. The group assignment affects the role of the participant in stages 2 and 4.

In stages 2 and 4 group B participants undertake a contest. Those participants who score in the top half of performers on the real effort task in that stage are given a score and payment of 40. Those in the lower half are given a score and payment of 10.

Each group B participants is randomly matched with a role A participants in rounds 2 and 4. These matchings are different in stages 2 and 4. If a group A participant is matched with a high scoring role group B participant in stage 2 they are be matched with a low scoring role B participant in stage 4, and vice-versa. Role A participants are informed that they are matched with high or low scoring role B participant in rounds 2 and 4 prior to commencing the round.

In stages 2 and 4 group A participants are paid a piece rate for performance on the real effort, as in stages 1,3 and 5. However, in contrast to stages 1, 3 and 5, after completion of the real effort task the group A participants are given the opportunity to purchase credits. Each credit costs the participant 1 Exp$, and they can spend up to the total amount of Exp$ that they have earned in that stage. The credits increase the participants score and reduce their income as follows:

Score = the number of sliders correctly placed + any credits bought

Payment = the number of sliders correctly placed - any credits bought

In Stages 2 and 4 Role B participants are shown the score of their matched role A participant. They are, however, not shown the number of sliders placed or the numbers of credits bought. The role B participant is asked to give feedback on the score on a Likert scale ranging from very bad to very good.

The matching of each role A participant with both a high and low scoring role B participant was designed to test whether and how the differential income of the matched participant influences the participant's choice of effort/income and credits purchased.

After the fifth stage was completed, participants were asked a set of questions related to the outcome in stages 2 and 4. Participants in role A were asked, on a 5 point Likert scale and for each of rounds 2 and 4 (i) whether the prospect of feedback increased their effort, (ii) whether the prospect of feedback caused them to buy credits, (iii) how fair the feedback they received was. Participants in role B were asked on a 5 point Likert scale for both rounds 2 and 4: (i) whether the prospect of giving feedback caused them to increase their effort, and (ii) whether they thought their feedback was fair.

2.2 Experimental Procedures

The experimental sessions lasted approximately 1.5 hours on average. The experiment was conducted in the Experimental Economics Laboratory at the University of Melbourne between November 2015 and April 2016 using z-Tree (Fischbacher, 2007).

The 132 participants were students from the University of Melbourne recruited randomly using ORSEE (Greiner, 2004) from a pool of more than 6000 volunteers. Subjects were from different academic backgrounds including introductory economics.

Upon arrival at the laboratory, participants were seated in partitioned computer terminals. Individuals were then provided a set of instructions, both printed and on the computer screen, which explained in detail the experiment. After reading the instructions, each participant answered a quiz on the structure of the experiments which aimed to help them understand how earnings were determined in the experiment. Participants were then given a practice round of the sliders task to ensure they were familiar with it.

Prior to each stage of the experiment participants were provided with a review of the instructions that were relevant for that stage, and undertook a quiz related to the conduct of that stage. The slider task in each of the five stages of the experiment (and the practice stage) lasted four minutes.

At the conclusion of the experiment all participants were then asked a series of demographic questions (age, gender, citizenship, ethnicity, enrolment degree). Participants were asked to answer three questions included in the Self-Assessment Manikin (SAM) developed by Lang (1980). The SAM is a tool to measure a participant's emotional state using the PAD (Pleasure, Arousal and Dominance) psychological analysis developed by Albert Mehrabian and James A. Russell (1974). The SAM aims to identify a respondent's emotional state of mind in 3 scales: happiness, arousal and dominance, where:

- The Happiness (Pleasure-Displeasure) Scale measures how pleasant or unpleasant a respondent feels.
- The Arousal Scale measures how energized as compared to soporific a respondent feels.
- The Dominance Scale measures how controlling and dominant compared to controlled a respondent feels.

This SAM questionnaire is used frequently in social psychology, and has been used in an economic experiment by Azmat and Iriberri (2015).

Participants were also asked a series of questions on their psychological makeup: the big 5 personality traits, locus of control and risk attitude. These questions are drawn from those use for the German Socio-Economic Panel questionnaire in 2005.[1] Developed within the psychology literature, the Big Five Personality Traits are intended to provide a methodology of how personality varies across individuals (John and Srivastava, 1999). As the name suggests, this is a categories the personality of a respondents according to the following five personality dimensions. A brief summary of each measure is given below:

[1] Available at: http://www.eui.eu/Research/Library/ResearchGuides/Economics/Statistics/DataPortal/GSOEP.aspx

- Openness to experience: a high score on this measure indicates that the respondent is intellectually curious and has a preference for novelty, whereas a low score indicates the respondent tends to be consistent and cautious.
- Conscientiousness: a high score on this measure indicates the respondent has self-discipline and is organized whereas a low score indicates a tendency to procrastinate and be impulsive or careless.
- Extraversion: a high score on this measure indicates the respondent is outgoing, energetic and stimulated by social interaction, whereas a low score indicates a tendency to be solitary and introspective.
- Agreeableness: a high score on this measure indicates the respondent is compassionate and cooperative whereas a low score on this measure indicates they are analytical, suspicious and detached.
- Neuroticism: a high score on this measure indicates the respondent is sensitive to negative emotions whereas a low score indicates they are confident and sure of themselves.

Observe that the two traits Extraversion and Agreeableness describe how an individual interacts with other people. By contrast, Conscientiousness, Neuroticism, and Openness to Experience deal with how a person responds emotionally and intellectually to the world in general.

The Locus of Control measure was developed by Rotter (1966). As its name suggests, this scale measures how much an individual believes they have control over their life. A high score on this measure indicates that an individual strongly believes that their success in life are mainly due to their own efforts, while a low score on this measure suggests that an individual strongly believes that their life outcomes are mainly due to luck. We also adopt the question related to risk attitude used in the 2004 German Socio-Economic Panel. This question asks participants, on an 11-point Likert scale, to rate their willing to accept risk. [2]

[2] This question thus attempts to measure how participants feel about risk rather than provide a point estimate of their of risk aversion. Our principle aim here was to capture participants' psychological makeup, to see if and how that influenced behaviour. While additional information on participants' characteristics may have been obtained from incentivised mechanisms such as those for risk aversion and pro-sociality, we were weary of introducing additional incentive measures because of their possible interaction with the experiment proper. Additionally, under the institutional constraints faced when conducting the experiment, such incentivised measurements would have reduced time available to conduct the experiment and reduced payments to participants (and thus the salience of the experiment).

3 Theory

In this section we derive theoretical predictions for the outcome of our experiment on the assumption that participant utility is derived from material consumption and referent-dependent esteem.

3.1 The participant's utility function

To model the experiment theoretically, let e_r the effort taken by participant in stage r (i.e the number of sliders completed). Let and w be the piece rate paid for the real effort task in all rounds for role A participants and rounds 1, 3 and 5 for role B participants. Let K_r be the payment to role B participants in rounds $r = 2, 4$, where $K_r = 10$ if $e_r < \overline{e}_r^B$ $K_r = 40$ if $e_r > \overline{e}_r^B$, where $\overline{e}_r^B$ is the average effort of role B participants in round r. The payoff to participants, π, is given by:

$$\pi = \begin{cases} \pi^A = \sum_{r=1}^{5} we_r - \sum_{r=2,4} c_r & \text{for role A participants} \\ \pi^B = \sum_{r=1,3,5} w\, e_r + \sum_{r=2,4} K_i & \text{for role B participants} \end{cases}$$

Participants have a score, t_r in rounds 2 and 4. For role A participants their score in round $r = 2{,}4$ is given by:

$$t_r = we_r + c_r$$

The score for role B participants in round r is K_r.

Participants receive utility from material consumption (i.e. money earned in the experiment). Let this component of utility be given by:

$$\upsilon\pi$$

where υ measures the magnitude of the marginal utility of income of particular individuals. The parameter υ is assumed distributed across the population. As the income earned by participants from the experiment is very low compared to each participant's annual income, it would seem appropriate to approximate their utility with a linear function.

A participant receives utility from the difference between the esteem they receive (measured by feedback in our experiment), s_r, and s_r^m, the esteem they give to their matched partner. Assume this component of utility takes the form:

$$Yy(s_r - s_r^m)$$

where $y(s_r - s_r^m)$ is common to all individuals and Y is a measure of the intensity of a given participant's marginal utility of esteem. As it is assumed participants prefer higher relative esteem, $y'(s_r - s_r^m) > 0$. It is further assumed that $y''(s_r - s_r^m) < 0$. That is, there is a decreasing marginal utility of relative esteem. This further implies an increasing marginal dis-utility of dis-esteem. While such an assumption is plausible, it should be noted that it is also conceivable that participants have increasing marginal utility of dis-esteem. Such a possibility could also be analysed in the context of the model presented here. However it would introduce non-convexities into the participant's objective function which would in turn introduce the possibility of multiple local maxima. Such a possibility is excluded from consideration here as it would involve very detailed analysis that is tangential to the core direction of this paper.

Note that a participant can increase the component of their utility due to esteem by reducing the esteem offered to others.[3] However we assume that participants also have an intrinsic motivation to make accurate judgements of others. Making an inaccurate judgement causes a reduction in this utility, i.e. it incurs a psychological cost.

The role B participant's judgement is conveyed as either esteem or disesteem. In our experiment participants observe the score of their partner. In making an assessment of the esteem to offer their matched partner, the participant must judge the extent that the score reflects effort. Define participant's estimate of the proportion of their matched participant's score in round r due to effort as, μ_r^m, so the participant estimates their partner's score as $\mu_r^m t_r^m$. Let the benefit from providing esteem to a matched partner with a score t_r^m be given by:

$$F - G(s_r^m - \mu_r^m t_r^m)^2$$

where F is the utility from providing feedback , and G is a parameter which measures the increasing marginal cost of inaccuracy in the participant's judgement. For simplicity of analysis it is assume that G is common to all individuals.

Under the above assumptions the utility of a participant from the experiment, is given by:

$$U(\pi, e, s_2, s_4) = \upsilon\pi - \sum_{r=1}^{5} h(e_r) + \sum_{r=2,4} \left\{ Yy(s_r - s_r^m) - \frac{G}{2}\left(s_r^m - \mu_r^m t_r^m\right)^2 \right\} + 2F$$

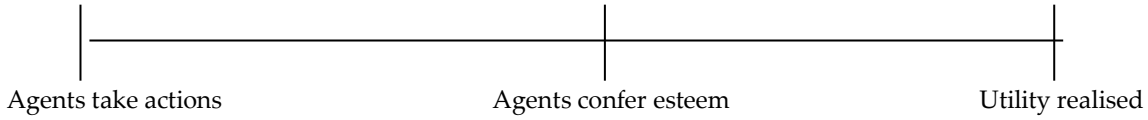
where $h(.)$ is the cost of effort in each stage, where $h'(.) > 0$ and $h''(.) > 0$. We take the cost of effort function, $h(.)$, to be common across all individuals. This assumption reflects that

[3] The participant may not want to discount the score to reflect effort. The analysis would be unaffected if this was the case.

the slider task is intended to be a pure real effort task and minimise any skill differences amongst participants (Gill and Prowse, 2012).[4]

Note that the structure of the payoff is such that the utility derived in each round from a participant's actions and conferring of esteem is independent of the utility derived in other rounds. Thus we can analyse the actions and coffering of esteem in each round independently from other rounds. Assume the following ordering of decisions within each round of the experiment. First, individuals take actions, i.e. choose effort and credits purchase (action stage). After all individuals have taken their actions, their actions are revealed all individuals simultaneously evaluate actions (evaluation stage). Then individuals realise utility. Figure 1 depicts the psychological timeline in each round.

Figure 1:Timeline of decisions within in each round



We now consider the actions of participants in each role. We expect there are two relevant sources of participant heterogeneity of those taking part in the experiment: their marginal utility of income, υ, and their marginal utility of esteem, Y.

3.2 Role B participants' feedback

The key function of role B participants in our experiment is to provide feedback to role A participants. Recalling that objective feedback occurs when $s_r^{mA} = \mu_r^{mA} t_r^{mA}$, where the superscript mA refers to the matched partner being in role A, we have:

> **Proposition 1**: A role B participant's feedback, $s_r^{mA}(\mu_r^{mA} t_r^{mA}, K_r; Y, G)$, is a function of the matched role A participant's score, the participant's own score, their marginal utility of esteem and their marginal cost of misrepresentation. Further:
>
> (i) Role B feedback is lower than the objective level, i.e. $s_r^{mA} < \mu_r^{mA} t_r^{mA}$, and
>
> (ii) Role B's feedback is positively related to the reported score.

[4] It could be the case that participants have a heterogeneous distaste for the slider task. Such a possibility could be incorporated into the theory at the cost of additional complexity. In this event a relatively high distaste for undertaking the slider task would have a quality similar effect to having a relatively low marginal utility of consumption could

When the utility of esteem takes the linear form $(s_r^{mB} - s_r^{mA}) = \theta(s_r^{mB} - s_r^{mA})$, with $\theta>0$, then the esteem given by role B participants to their matched role A participant is:

$$s_r^{mA} = \mu_r^{mA} t_r^{mA} - \frac{\theta Y}{G} \quad (1)$$

and consequently feedback is independent of the role B participant's score.

Role B participants have two forces working on them when providing feedback. They gain utility from providing accurate feedback, but have their utility from esteem falls as their feedback becomes more positive. They optimal trade-off between these two effects is to provide lower than objective feedback. However because accurate feedback provides positive utility to the role B participant, raising the role A participant's score raises the feedback provided by the role B participant. Note also, that when the utility of esteem is linear, the marginal utility of esteem is constant for role B participants (in particular not dependent their score), and therefore their score will not influence their feedback.

3.3 Role A participants' effort and credit purchase

First consider the effort choice of role A participants in rounds 1, 3 and 5. In the absence of esteem considerations the participant will equate the marginal utility of income with the marginal cost of effort. That is, the participant will choose to exert their lone effort, e_i^*, in rounds 1, 3, and 5, where lone effort is defined as:

$$w\upsilon = h'(e_r^*)$$

It is straightforward to demonstrate that, because of increasing marginal cost of effort, a participant's lone effort is positively related to their utility of income

Now consider the role A participant's actions in round 2 and 4. Recall that a role A participant can increase their score by either increasing effort or buying credits. The matched role B participant cannot identify the extent to which a score is due to effort or the purchase of credits. Mathematically we will therefore assume that:

$$w\frac{\partial s_r^{mA}}{\partial(\mu_r t_r)}\frac{\partial(\mu_r t_r)}{\partial c_r} = \frac{\partial s_r^{mA}}{\partial(\mu_r t_r)}\frac{\partial(\mu_r t_r)}{\partial e_r}$$

That is, the matched role B partner cannot tell whether score comes from effort or credit. We assume that $\frac{\partial(\mu_r t_r)}{\partial c_r}$ is a constant based on participant beliefs and that $\frac{\partial s_r^{mA}}{\partial(\mu_r t_r)} \leq 0$. In this case the following proposition holds:

Proposition 2: In rounds 2 and 4, role A participants choose to exert greater level of effort than the lone effort. If the role A participant purchases does not purchase credits effort is given by:

$$h_r^{'}(e_r) = \upsilon w + w\Delta U(Y) \tag{2}$$

where $\Delta U(Y)$ is the role A participant's marginal utility of esteem. If the participant purchases credits effort is given by:

$$h_r^{'}(e_r) = 2w\upsilon \tag{3}$$

and the number of credits purchased, c_r, satisfies:

$$\upsilon = \Delta U(Y) \tag{4}$$

The proof of proposition 2 along with the precise formula for $\Delta U(Y)$ is given in appendix A. The magnitude of the marginal utility of esteem is positively related to the parameter Y. Both equations (2) and (12) state that the participant chooses effort up at the point where the marginal benefit of effort is equal to its marginal cost. To explain this, observe that when the participant does not buy credits, the benefit of additional effort is, υw, the additional income generate plus, $w\Delta U(Y)$, the additional esteem generated from the increase in score. Thus (2) equates the marginal cost of effort, $h_r^{'}(e_r)$, with the marginal benefit of effort. When the participant buys credits, (4) indicates that they optimise by choosing number of credits purchased to the point where their opportunity cost of the purchase, υ, is equal to, ΔU, the marginal utility of the esteem generated. Thus, when the participant optimally purchases credits the marginal esteem is also υ . In this way (12) can also be interpreted as stating that effort is chosen so that the marginal benefit of effort is equal to its marginal cost.

Participants are assumed to differ in their marginal utility of income, υ, and, Y, the intensity of marginal utility of esteem. We now consider how differences in υ and, Y affect effort and credits choices.

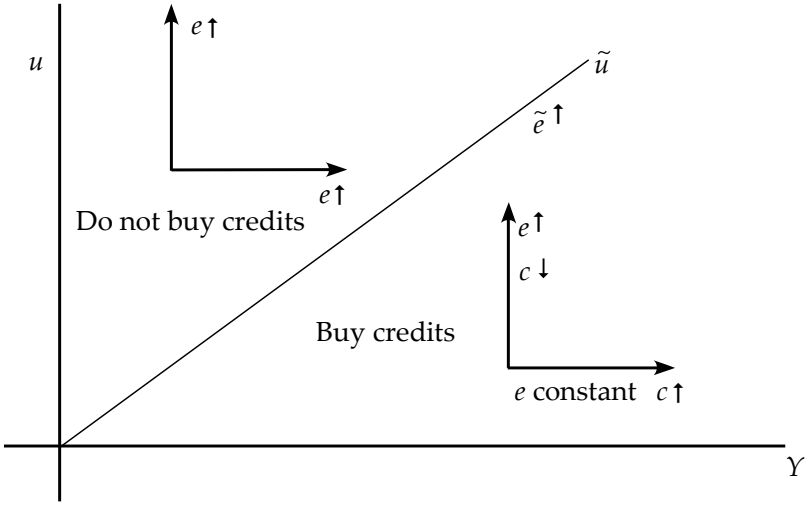
Proposition 3: (i) Suppose a role A participant does not buy credits. Then effort is increased by increased υ and Y. (ii) Suppose a role A participant purchases credits. Increased Y increases the number of credits purchased. Increased υ increases effort and reduces credits purchased.

Suppose the participants do not buy credits. In this case both an increase in the marginal utility of income and the intensity of esteem increase the benefit of supplying effort. Hence effort increases. Suppose the participant buys credits. Then an increase in the marginal utility of income will induce a greater effort. However it will also cause a substitution away from credits as monetary payments have become more valuable to the participant. An

increase in the intensity of esteem causes a substitution of given earnings from other goods (i.e. the monetary payment) towards esteem seeking (credit purchases).

We now consider the question of whether the participant will purchase credits. A participant is indifferent between purchasing credits or not if (2) and (4) hold when $c = 0$. This condition defines the curve $\tilde{\upsilon}(\tilde{Y})$ in Figure 2. Each point on this line represents combinations of the marginal utility of income and esteem for which the individual is indifferent between purchasing credits or not. It is show in the appendix that $\tilde{\upsilon}(\tilde{Y})$ has a positive slope. Intuitively, a higher marginal utility of income means that a role A participant has a higher $\tilde{e}$. Consequently such a participant would require a higher marginal utility from esteem to be indifferent between buying credits and not.

Figure 2: The credit purchase decision as a function of υ and Y



By utilising the definition of $\tilde{\upsilon}(\tilde{Y})$ the following proposition is established:

> **Proposition 4**: (i) A role A participant will buy credits in rounds 2 and 4 if they have a sufficiently low marginal utility of income, i.e. if $\upsilon < \tilde{\upsilon}$, , and will not buy credits if they have a sufficiently high marginal utility of income i.e. if $\upsilon \geq \tilde{\upsilon}$. (ii) A role A participant will buy credits in rounds 2 and 4 if they are sufficiently sensitive to esteem concerns, i.e. if $Y^K > \tilde{Y}^K$, and will not buy credits if they are sufficiently insensitive to esteem concerns, i.e. if $Y^K \leq \tilde{Y}^K$.

As depicted in Figure 2, this result shows that credits will only be purchased by participants who are sufficiently concerned about esteem and have a sufficiently low marginal utility of income. Those who do not buy credits will tend to have a relatively high marginal utility of income and relatively low marginal utility of esteem.

3.4 Key conjectures from theory

In this section we derive four conjectures from the above theory that can be tested with data from our experiment. Proposition 3 and 4 together indicate that a role A participant will increase their propensity to buy credits when they are more sensitive to feelings of esteem (higher Y). Role A participants may be more sensitive to esteem with matched with either a high or low scoring role B participant. However, *a priori*, we do not which, or if either, of the possibilities is true. To test whether participants react to their matched role B participants type (i.e. high or low scoring) we test the following conjecture:

> **Conjecture 1**: The role A participant's purchase of credits is not affected by the type of their matched role B participant.

This conjecture addresses one of the key issues of the paper, whether the status of people an individual associates with affects that individual's esteem seeking.

Proposition 4 shows that the decision to purchase credits depends on υ, the marginal utility of consumption. From its definition, lone effort is dependent only on υ. Thus we are able to use lone as a proxy for υ. We use our experimental data to estimate participant's lone effort as their effort over rounds 1, 3, and 5. Therefore, by inspection of Figure 2 we conjecture:

> **Conjecture 2**: Those role A participants with low lone effort are more likely to buy credits in rounds 2 and 4.

If true, this conjecture would imply that it is those people who have a lower marginal utility of consumption are the ones who are more likely to engage in esteem seeking.

Two key theoretical predictions regarding participant behaviour can be assessed using the feedback data obtained in rounds 2 and 4. These conjecture follow from proposition 1. Proposition 1(i) indicates role B feedback is less than objective. If this is true, we might expect that role A participants view the feedback they receive less favourably than do the role B participants who gave the feedback. Thus we conjecture:

> **Conjecture 3**: Role B participants view their feedback as fairer than their matched role A participants do.

Our fourth conjecture follows directly proposition 1(ii):

> **Conjecture 4**: A role B participant's feedback will be related to their matched role A participant's score.

Conjectures 1,2 and 3 can be assessed using the summary statistics from the experiment. However conjecture 4 requires an econometric analysis in order to be tested.

4 Results

In this section we first give an overview of the key summary statistics of the data from our experiment. We then present an econometric analysis to identify the determinants of esteem seeking and esteem provision.

4.1 Summary statistics

Table 1 provides the number of participants who purchased credits in rounds 2 and 4. Recall that role A participants who have a high scoring partner in round 2 face a low participant in round 4, and visa-versa. Table 1 demonstrates that there is a fall in the number buying credits from round 2 to 4. The average purchase of credits is higher when the participant faces a low scoring participant. The distributions of credits is significantly different at the 10% level when role A participants has a low vs. high matched partner (p=0.07, Kruskal-Wallis test). This suggest that conjecture 1 is false, and:

> **Observation 1:** Role A participants buy more credits when matched with low scoring role B participants.

Overall of the 66 role A participants, 14 participants bought credits when matched with a high scoring role B player while 15 bought credits when facing a low scoring role B participant. This difference is not significant.

Table 1 Credits by round and partner score

	Round 2		**Round 4**	
Partner is:	High	Low	High	Low
Number buying credits / 66	10	7	4	8
Average credits bought	2.8	7	5.25	8.75

In the post-experiment survey role A participants were asked, using a Likert scale, whether knowing they would receive feedback on their score caused them to buy credits in rounds 2 and 4. In Table 12 in Appendix C.1 we calculate the Spearman rank correlation between the decision to buy credits and the answer to this post-experiment survey question. We find a statistically significant correlation between the decision to purchase credits and the tendency to answer in the direction of strongly agree to this question. This provides support for the interpretation of the decision to purchase of credits as one designed to elicit esteem.

Table 2 compares the average effort of those who bought credits and those who did not. The average effort in rounds 1, 3, and 5 is also lower amongst those who buy credits than those who do not (p =0.002). This is consistent with conjecture 2, thus:

> **Observation 2**: Credit purchase is associated with a lower average effort across rounds 1, 3, and 5.

Table 2 Average effort of those who bought and those who did not buy credits

	Average *effort* *r* = 1, 3, 5 un-partnered rounds	**Average *effort* *r* = 2, 4 partnered rounds**
A Not Buy (*r* = 2 or 4)	38.9†	38.4
A Buy (*r* = 2 or 4)	34.3†	34.2
† The average effort in rounds 1, 3, and 5 is significantly lower amongst those who buy credits than those who do not (p =0.002).		

Table 3 gives the averages of role A participants actions in rounds 2 and 4. The participants are divided by those who bought credits and those that did not. Note that the average effort of those who bought credits is lower those who did not. The distributions of effort are significantly different at the 1% level. (p=0.006). Average effort of participants who bought credits does not differ significantly when facing high scoring vs low scoring partners.

Table 3 Average Role A participants' actions in rounds 2 and 4

		Not Buy Credits			**Buy Credits**		
Partner	**Stage**	***r* = 2**	***r* = 4**	**Total**	***r* = 2**	***r* = 4**	**Total**
High	Avg *effort*	36.7	39.6	38.3	32.6	38.5	34.3
	Avg *credits*				2.8	5.3	3.5*
	Avg *score*				35.4	43.8	34.3
Low	Avg *effort*	37.3	38.0	37.6	32.3	35.9	34.2
	Avg *credits*				7.0	8.8	7.9*
	Avg *score*				39.3	44.7	42.1
Total	Avg *effort*	37.0	38.9	38.0†	32.5	36.8	34.2†
	Avg *credits*				4.5	7.6	5.8
	Avg *score*				37.0	44.4	40.0
† The distributions of effort in rounds 2 and 4 for buyers vs non-buyers, are significantly different at the 1%. * The distributions of credits in rounds 2 and 4 for those paired with a low vs high scoring partner, are significantly different at the 10% level.							

However these participants buy on average significantly more credits when facing low scoring partners ($p = 0.07$). These findings are consistent with participants using more credits to buy more esteem when facing low scoring partners compared to high scoring partners. This is reflected in a higher average score for participants who bought credits when facing a low scoring partner compared to a high scoring partner.

Table 4 shows the average effort across all rounds of both role A and role B participants. Note that there is a tendency for average effort to increase across rounds, apart from those role A participants who buy credits in rounds 2 and 4. To compare their relative effort, the ratio of the average effort of role B participants to the average effort of role A participants is shown in the bottom row. In round 1 the ratio is close to 1. The ratio is lower in subsequent rounds, apart from the average efforts in rounds 2 and 4 of those role A participants who buy credits. This observation is consistent with esteem concerns in rounds 2 and 4 raising effort levels. Further this effect appears to have spilt over into rounds 3 and 5.

Table 4 Average Effort by Round

	r = 1	**r = 2**		**r = 3**	**r = 4**		**r = 5**
Role		**Not Buy**	**Buy**		**Not Buy**	**Buy**	
A	34.5	37.0	32.5	38.0	38.9	36.75	40.2
B	33.7	34.7		35.8	36.0		38.1
B/A	0.98	0.94	1.07	0.94	0.93	0.98	0.95

Table 5 shows the average feedback provided by Role B participants to role A participants, and the average assessment by role A participants the fairness of that feedback. The role B participants give feedback on a 5-point Likert scare, where 5 represents very good and 1 very bad. The role A participants assess the fairness of that feedback on a 5-point Likert scale, where 5 represents very fair and 1 very unfair. It can be seen that the ranking of the fairness by participants follows the feedback they receive. Role A participants with low scoring partners who did not buy credits receive the best feedback and also view that feedback as most fair. The role A participants with low scoring partners who did buy credits received the worst feedback and also were the group who viewed that feedback as the least fair. It can also be noted that those participants receive very similar feedback from the role A participants with high scoring partners who do not buy credits. Yet the former group interpret their feedback as less fair. This suggests that those role A participants who purchase credits desires positive feedback arising from their purchase from their low scoring partner. They are subsequently disappointed (and thus view the feedback as relatively unfair) when they do not receive sufficient recognition.

Table 5 Average Feedback and Fairness Scores

Partner type	Question	Buy Credits		
		No	Yes	Total
High	Feedback (B to A)	3.88	4.07	3.92
	Fairness (A to B)	3.63	3.79	3.67
Low	Feedback (B to A)	4.22	3.87	4.14
	Fairness (A to B)	3.90	3.47	3.80
Total	Feedback (B to A)	4.05	3.97	4.03
	Fairness (A to B)	3.77	3.62	3.73

There is also, on average, a difference in fairness perceptions between those that bought credits and those who did not. Those who bought credits viewed their high scoring partner as providing fairer feedback than their low scoring partner. This suggest that the role A partner was expecting more positive affirmation from the low scoring partner. On the other hand, those who didn't buy credits thought the low scoring partner provided fairer feedback.

Table 6 compares the average perception of the fairness of role B's feedback from the role A participant receiving the feedback and the role B participant giving the feedback. It can be seen that:

> **Observation 3**: On average role A participants have a significantly less favourable perception of the feedback they receive than role B participants have of their own feedback.

Observation 3 is consistent with conjecture 3: role B participants will be less than objective in their feedback.

Table 6 Average Fairness of Feedback

Role	r = 2	r = 4	Total
A	3.77	3. 70	3.73†
B	4.02	4.15	4.08†

† The mean and distributions are significantly different at the 5% level. (p=0.02 for the two sample t-test and p=0.02 for the Kruskal-Wallis test)

4.2 Econometric Analysis

In this section we use the above theory to inform our specification of our estimating equations.

4.2.1 Role A Participants: effort and credits

We first outline the specification of the econometric model used to estimate the effort and credit choice of role A participants. In rounds 2 and 4 of the experiment role A participants must first decide whether or buy credits (*buy*=1) or not (*buy*=0). We model the decision of a role A participant to buy credits, *buy* via a latent variable, *buy** for each player *i* in round *r* = 2 and *r* = 4, such that player buys credits *buy*=1 if *buy**>0 and does not buy credits when *buy*=0 if *buy**>0 as:

$$buy^{*}_{i,r} = \alpha_0 + \alpha_{r4} r4_{r,i} + \alpha_l \, low_{r,i} + \alpha_{\bar{e}} \, \overline{effort135}_i + \alpha_e \, effort_{r,i} + \boldsymbol{\alpha}_d \mathbf{dempsy}^R_i + \varepsilon^{buy*}_{r,i} \qquad (5)$$

where *r*4 is a dummy for round 4, *low* is a dummy variable for being matched with a low scoring role B participant, $\overline{effort135}$ is the participant's average effort in rounds 1, 3 and 5, *effort* is the participants effort in the round, and $\mathbf{dempsy}^R$ is a vector of demographic and psychological variables. As shown in the theory, there will be a critical level of marginal utility of income above which the role A participant does not buy credits. In the theory, a participant's marginal utility is directly related to their lone effort. Lone effort in our experiment is measured by $\overline{effort135}$. It is expected from the theory above that, as higher $\overline{effort135}$ indicates a higher marginal utility of income, that $\alpha_{\bar{e}} < 0$. The specification (5) also allows for the role B participant's score, through α_l to influence the purchasing decision.

As noted above, the average effort of participants in rounds 1, 3 and 5 is determined by marginal utilities of income. These may, in turn, be determined by demographic and psychological factors. Additionally those who buy might exhibit different behaviour to those who do not. To allow for these possibilities we specify average effort in rounds 1, 3, and 5, $\overline{effort135}$ for each player *i* and rounds *r* = 2 and *r* = 4 as:

$$\overline{effort135}_i = \zeta_0^{buy} + \boldsymbol{\zeta}_d^{buy} \mathbf{dempsy}_i + \varepsilon^{\bar{e},buy}_{r,i} \quad \text{for } buy_{r,i} = 0,1 \qquad (6)$$

In rounds 2 and 4 effort has an additional benefits relative to rounds 1, 3 and 5 for role A participants: it increases their score (esteem) in addition to their income. In the absence of these effects the effort in rounds 2 and 4 would be predicted by the effort expended by the participant's effort in rounds 1, 3, and 5 (which in turn we postulate depends on the marginal utility of income). The extent to which the esteem concerns cause the participant to increase effort may depend on whether their role B partner has a high or low score. There may also be ordering effects as suggested by the summary statistics. This is captured by

including the dummy variable for round 4 ($r4$). Thus we specify the participant's effort in rounds 2 and 4 as:

$$effort_{r,i} = \psi_0^{buy} + \psi_{r4}^{buy} r4_{r,i} + \psi_l^{buy} low_{r,i} + \psi_{\bar{e}}^{buy} \overline{effort135}_i + \psi_c^{buy} credits_{r,i} + \psi_d^{buy} \mathbf{dempsy}_i^R + \varepsilon_{r,i}^{e,buy} \quad \text{for } buy_{r,i} = 0,1 \text{ and } \psi_c^0 = 0 \tag{7}$$

where *credits* is the number of credits purchased by the participant in round r. If $\psi_l^{b=0} >(<)0$, then a role A participant will exert more effort when matched with a low (high) score type B participant.

If a role A participant decides to buy credits, they must decide on *credits*. We adopt for each player i and rounds $r = 2$ *and* $r = 4$:

$$credits_{r,i} = \beta_0^b + \beta_{r4} r4_{r,i} + \beta_l low_{r,i} + \beta_{\bar{e}} \overline{effort135}_i + \beta_e effort_{r,i} + \boldsymbol{\beta}_d \mathbf{dempsy}_i^R + \varepsilon_{r,i}^c \quad \text{if } buy_{r,i} = 1 \tag{8}$$

An observed deviation from average effort by the role A participant could be due to them having a low marginal utility of income or a high marginal utility from esteem. Thus if participants buy credits we expect that there is a positive correlation between the number purchased and an effort deviation.

The model includes 3 endogenous variables, *buy**, *effort*, *credits* and $\overline{effort135}$. The latter is a function of only exogenous variables and can be treated as pre-determined or weakly exogenous. There are two exogenous treatments, *low*, *r*4 and fifteen exogenous demographic variables in **dempsy**, four of which are excluded in $\mathbf{dempsy}^R$ from all equations other than (6) to ensure identification. Further details of the estimation are given in appendix B.

Table 7 shows the estimation results of the reduced form of equation (5) by a Probit by omitting the endogenous *effort* (see appendix B for details). The regression is consistent with conjecture 2: a low average effort in rounds 1, 3 and 5 implies higher probability that the role A participant purchases credits. Further, after controlling for controlling for effort 1, 3 and 5 in this way, females are less likely to purchase credits. Both low and round 4 insignificant. Thus there is no evidence of the score of partner influencing the purchasing decisions. Similarly there is no evidence of an ordering effect in the decision to purchase. Of the emotional and psychological variables, only neuroticism is significant. Recall that neuroticism measures a participant's sensitivity to emotions. It seems plausible that those who are sensitive to the emotions associated with the receipt of feedback/esteem would be those who purchase credits. The results in Table 7 are consistent with there being the dichotomy in behaviour suggested by theory as summarised by Figure 2. This in turns

supports the separate econometric modelling the behaviour of those who did and didn't buy credits.

Table 7 Choice to buy credits – Reduced Form - Probit

Variable	Estimate (Std Err)	Variable	Estimate (Std Err)
*r*4	-0.29 (0.28)	Risk	0.06 (0.06)
low	0.08 (0.27)	Happiness	-0.05 (0.10)
$\overline{effort_{135}}$	-0.13*** (0.03)	Arousal	-0.07 (0.07)
Constant	5.49*** (1.93)	Dominance	0.04 (0.09)
Female	-0.99*** (0.34)	Extraversion	0.03 (0.05)
Age	-0.06 (0.05)	Agreeableness	0.00 (0.05)
		Conscientiousness	0.03 (0.07)
		Neuroticism	0.12*** (0.04)
		Openness	-0.15* (0.08)
R-squared(McFadden)	0.227	Log likelihood	-53.72

The econometric model given by equations (5), (6), (7) and (8) by can solved and the reduced form for all of the endogenous variables as specified by equations (15), (16), (17), and (18) in appendix B.2. The result of the SUR estimation of the reduced form is given in Table 8 on the following page. The first row in Table 8 confirms that if a participant buys credits, they buy significantly more credits when facing a low scoring role B partner, given their demographics and personality. Thus conjecture 1 is rejected in this analysis: role A participants buy significantly more credits when matched with low scoring role B participants.

Table 8 Reduced Form SUR Estimates

	Probit	$buy = 1$			$buy = 0$	
	$\Phi(buy^*)$	$\overline{effort_{135}}$	*effort*	*credits*	$\overline{effort_{135}}$	*effort*
r4	-0.29		2.10**	-1.00		2.10**
	(0.26)		(0.83)	(1.33)		(0.99)
low	0.04		-0.92	3.00**		-0.37
	(0.26)		(0.73)	(1.37)		(1.00)
Constant	0.04	29.46***	20.24**	-46.24***	41.68***	49.54***
	(1.48)	(3.75)	(8.76)	(10.51)	(5.46)	(5.73)
Female	-0.14	-2.73***	-0.35	5.09**	-4.00***	-4.97***
	(0.34)	(0.74)	(1.89)	(2.00)	(1.17)	(1.33)
Age	-0.02	-0.87***	-0.39	1.64***	-0.40***	-0.51***
	(0.04)	(0.16)	(0.30)	(0.35)	(0.14)	(0.16)
Risk	0.01	1.94***	1.15***	0.04	-0.29	-0.09
	(0.06)	(0.25)	(0.35)	(0.60)	(0.21)	(0.24)
Happiness	-0.16	1.25**	0.22	0.97	1.22***	1.12***
	(0.10)	(0.53)	(1.18)	(1.26)	(0.32)	(0.39)
Arousal	0.02	0.84***	1.30***	1.34***	-0.65***	-0.74***
	(0.07)	(0.18)	(0.36)	(0.50)	(0.21)	(0.24)
Dominance	-0.01	0.39	0.46	-2.71***	0.68*	0.19
	(0.08)	(0.32)	(0.51)	(0.78)	(0.40)	(0.45)
Extraversion	0.04	-0.27**	-0.14	-0.42*	0.28	0.15
	(0.05)	(0.12)	(0.21)	(0.22)	(0.20)	(0.22)
Agreeableness	-0.03	0.68***	0.23	-1.24***	-0.09	-0.48**
	(0.05)	(0.16)	(0.25)	(0.35)	(0.19)	(0.20)
Conscientiousness	0.04	-0.14	-0.18	0.43	0.18	0.31
	(0.06)	(0.10)	(0.21)	(0.33)	(0.22)	(0.26)
Neuroticism	0.11***	-0.33	-0.83*	-3.20***	0.23	0.05
	(0.04)	(0.30)	(0.43)	(0.86)	(0.16)	(0.17)
Openness	-0.07	0.25	1.17	5.99***	-0.28	-0.37
	(0.07)	(0.54)	(1.06)	(1.52)	(0.30)	(0.36)
Business	-0.06	4.32***	4.87**	10.76***	4.35**	5.50***
	(0.45)	(1.04)	(2.20)	(2.31)	(1.73)	(1.89)
Eco/Eng/Arc	0.03	9.09***	14.02***	27.29***	7.35***	7.27***
	(0.46)	(1.71)	(2.51)	(5.03)	(2.08)	(2.31)
Med/Vet	-0.69	-6.67***	-2.67	-7.22***	2.21	3.58*
	(0.57)	(1.07)	(3.53)	(2.73)	(1.98)	(2.18)
Sci/Oth/None	-0.73*	9.42***	13.16***	29.56***	2.63*	2.08
	(0.41)	(1.48)	(1.21)	(4.88)	(1.43)	(1.50)
R-Squared	0.14	0.97	0.87	0.90	0.33	0.33

The second row indicates that participants increase their effort in round 4 relative to round 2, by the same amount regardless of whether a participant buys credits. The constant terms in the regressions for average effort in rounds 1, 3 and 5 ($\overline{effort_{135}}$) in columns 3 and 6 indicate that average effort is 11 sliders lower for buyers than non-buyers given role B demographics

and psychology. This difference is significant at the 10% level using a standard Wald test (see Appendix B.2.1) which provides some evidence that buyers have a significantly lower marginal utility given their demographics and personality. The constant in effort of buyers in rounds 2 and 4, is almost 30 points lower for buyers in column 4 of Table 8 compared to non-buyers in column 7, and significantly lower at the 1% level of significance, (again see Appendix B.2.1).

Females have lower effort in rounds 1, 3 and 5 irrespective whether they buy credits. In terms of the theory this would suggest that females have a lower marginal utility of effort. However, it is possible that they might have found the slider task more distasteful than men did. Females who buy credits purchase more credits than males who buy credits. This finding is consistent with females having a greater concern for the opinion of others, though it is not possible to assess this explanation with the available data.

Two key issues not addressed in the reduced form model are the role of the endogenous variables, *effort* for buyers and non-buyers, and *credits* for buyers. To address this we estimate equations (6) through to (8), using three stage least squares (3SLS) for buyers and non-buyers. The details of the estimation technique are given in the appendix B.3 and the results of the estimation are given in Table 13 in appendix C.2. It is found that *effort* significantly increases the number of *credits* purchased by buyers and that increases in the number of credits purchased by buyers increases their *effort*. The result suggests that *effort* and *credits* for buyers are complementary in generating esteem. The 3SLS estimation finds that when a role A participant faces a low scoring partner (*low*), they buy significantly more credits than when facing a high scoring partner, (after controlling for $effort$, $\overline{effort}_{135}$, in addition to demographics and personality).

A final matter for which the 3SLS does not consider is to simultaneously account of the decision to buy credits or not, together with the other endogenous variables of *effort* and *credits* for buyers. To examine this issue we use a General Structural Equation Model (GSEM), the details of which are provided in Appendix D together with the estimation results in Table 16. The GSEM results are consistent with our 3SLS and reduced form results.

4.2.2 Role B participants' feedback

Theory predicts that the role B's feedback will be related to their matched role A participant's reported score. The cost of deviating from objective scoring may also differ across role B participants according to their demographic characteristics. We thus estimate, using an ordered probit, role B feedback as a function of reported score and demographic characteristics. The full results of the estimation are reported in the first column of Table 14

and Table 15 in the appendix C.3. The significant estimates of coefficients are also shown in Table 9.

The results shown in shown in Table 9 yield:

> **Observation 4**: Feedback by role A participants is significantly related to their matched role A participant's reported score.

This observation is consistent with conjecture 4.

Table 9 Determinants of Feedback

Variable	Estimated Coefficient (Standard Error)
Score of A	0.10*** (0.02)
Female	0.48* (0.26)
Arousal	-0.08* (0.05)
Dominance	-0.11* (0.06)
Neuroticism	0.06** (0.03)
Sci/Oth/None `	0.61* (0.34)

There is also evidence that the psychological makeup of the participant affects feedback. Two of the dimensions of emotional state of individual are significant at the 10% level: the more alert one is the lower the feedback and the more controlling one feels the lower the score. These latter findings are consistent with our theoretical assumption that a psychological cost to deviating from objective reporting, and that this cost is inversely related to the strength of emotions. Furthermore, of the personality traits, neuroticism is significant at the 5% level: the more sensitive an individual is, the more positive the feedback.

There is no evidence that the score (high vs low) of the role B participant, K_i, influences the feedback. This finding is consistent with the utility of esteem, $y(.)$, being a linear function. Additionally there is no evidence of an ordering affect, so that experience of providing feedback does not change the feedback. This is consistent with role B participants having well-formed perceptions of performance.

4.2.3 Role A Participants: Fairness

In the post-experiment survey the role A participants indicate their perception of the fairness of their matched partner's feedback. We now consider the determinants of role A participants' perception of the fairness of the feedback they receive.

Assuming that individuals care about feedback, it would be expected that the perception of fairness is positively related to the level of feedback. Furthermore, the effort a participant makes may influence the participant's perceptions of fairness. Indeed there may be an interaction between feedback and effort. A participant who provides little effort and receives good feedback (and visa-versa) may feel the feedback is unfair. Additionally the perception of fairness may be influenced by the number of credits purchased. The purchase of credits may lead a participant to expect high feedback, or might make the participant feel their score is fraudulent. These former effect would make credit purchase increase the fairness perception, which the latter would see credit purchases reduce fairness perceptions.

We model A's perception of how fair their evaluation was using an ordered probit with controls for demographics, personality, round 4, their partner's effort, and their actions thought their effort and credits purchased. The full results of the estimation are reported in Table 14 and Table 15 in appendix C.3. Table 10 provides the coefficients estimated to be significant from the ordered probit of role A's perception of the fairness of their feedback.

Table 10 Determinants of fairness of feedback

Variable	Estimated Coefficient (Standard Error)
Feedback=1 "very poor"	-1.77*** (0.54)
Feedback=4 "good"	0.82*** (0.31)
Feedback=5 "very good"	1.90*** (0.37)
Extraversion	0.08** (0.04)
Agreeableness	0.12*** (0.04)
Business	0.86*** (0.31)

The results in Table 10 indicate that role A participants' assess their feedback as fairer when they receive "good" feedback and even more fair when they receive "very good" feedback compared to neutral feedback. Furthermore fairness assessment is significantly lower for "very poor" feedback compared to neutral feedback.

This is consistent with the view that, at least some matched role B participants who offer low feedback, reduce their feedback below the objective level.

The full estimation in Table 14 and Table 15 in appendix C.3 shows that effort , credits and partner type are not significant in A's assessment of how fair their feedback was.

The personality traits Extraversion and Agreeableness significantly increase role A's perception of how fair their evaluation was. These are the two traits related to inter-individual interactions. Thus it is not implausible that these variable play a role in the assessment of feedback and, furthermore, the sign of the effect is in the expected direction. There is no evidence that role A participant's emotional state, their other personality variables, gender and age play a role in their perception of the fairness of the feedback they receive.

4.2.4 Role B participant's view of their own fairness

Role B participants were asked to indicate how fair their feedback was. Role B participants' perception of fairness might be related to their characteristics, the score of their matched role A partner, or the feedback they give. We regress Role B fairness perception is regressed on these variables using an ordered probit. The full results of the estimation are reported in the second columns of Table 14 and Table 15 in appendix C.3.

Table 15 shows that only the personality trait of openness is significant at the 5% level. That is, people who are intellectually curious are more likely to believe their feedback is fair. Such people may feel they are better informed, and therefore believe they make better judgements. There is also some evidence that people who give 'very bad' as feedback are more likely to believe the feedback they provide is unfair (significant at the 10% level). Furthermore, people who give 'very good' as feedback are more likely to believe the feedback they provide is fair (significant at the 5% level). Table 15 also shows Economics, engineering and architecture students believe their feedback is less fair (significant at the 10% level).

No other variable is statistically significant as a determinant of the role B participant's perception of their own fairness. Importantly this includes the matched role A participant's score and the feedback they give other than 'very bad'. Taken with the results in **Table 6** this suggest that most role B participants systematically believe their feedback is not unfair.

The results in Table 14 and Table 15 reinforce the message from **Table 6**: there is a difference in the formulation of the perception of the fairness of the feedback between role A and role B participants. It is thus consistent with the motivation behind conjecture 3: if participants care about their relative esteem, then matched role A and role B participants each face different psychological incentives when assessing feedback.

5 Discussion

We present an experiment which is motivated by the hypothesis that people are motivated to seek esteem in additional to material goods. Our central finding is that approximately 22% of role A participants in our experiment purchased credits. As credits provide no material benefit to the participant who buys them, their purchase must then be designed to elicit esteem from their partner. This interpretation is further supported by the high correlation between the participants' actual purchase of credits with their response to the post experiment question on the motivation for buying credits.

We theoretically model participant behaviour in our experiment on the assumption that they gain utility receiving and giving esteem, in addition to material goods. Our experimental findings are broadly consistent with predications from this theory. Importantly, those role A participants whose lone score is low are more likely to buy credits than those with a high lone score. From theory, a high lone score is interpreted as the individual having a higher marginal utility of consumption: thus role A participants with a high lone score are relatively more concerned with material goods and less concerned with receiving esteem.

We further observe that role A participants buy more credits when facing a low scoring partner. This suggests that, when people are seeking esteem, they prefer to turn to people who have a similarly low performance on the real effort talk. Such a finding is reminiscent of Festinger's (1957) definition of a person's referents as people who are "close to [their] own ability". Such an interpretation is reinforced by the finding that low effort role A participants are more critical of the feedback provided by low scoring partners compared to high scoring partners. It seems they are wanting more recognition from low scorers (who are like themselves) than high scorers (who are not like themselves).

Our theoretical modelling suggests that role B feedback should be positively related to role A's performed. This is observed. Theory also suggests that, because esteem assessments are relative, there should be a gap between objective assessments and In this case we would expect a gap between role B assessment of the fairness of their own feedback and the assessment of that feedback by role A participants. Again this is observed in the summary statistics. Further it appears that the processes by which role A and role B participants arrive at their fairness assessment differs. Role B participants fairness assessments are not influenced by the feedback they provide. However Role A participants' view of the fairness of feedback is related to the feedback they receive: better feedback is viewed as more fair. This finding is consistent with Role B participants under providing lower than objective feedback – and asserting this is fair.

These findings taken together supports the hypothesis that there is a taste for esteem. The findings further suggest that there is a considerable heterogeneity in participants' strength of

demand for esteem over material goods. Demographic and psychological characteristics are important determinants of this heterogeneity. The purchase of credits (which we interpret as clear evidence of esteem seeking) is more likely amongst those with lower marginal utility of income, females and who exhibit the personality trait of neuroticism. Within that group a higher purchase of credits is observed amongst older people, females, those with emotional states of high arousal and low dominance, and with psychological states of high openness, low neuroticism and low agreeableness. We similarly find that feedback (which we interpret as the supply of esteem) is related to the personality trait neuroticism (with some evidence that gender, and emotional state may play a role).

The findings of this paper point the way for future work. Notably, to provide a stringent test for the hypothesis that people have a taste for receiving and supplying esteem, the experiment was conducted under anonymous conditions. In reality, of course, esteem seeking most frequently occurs in an open social context. The experimental environment we adopted is thus not conducive to esteem seeking. It may be expected that when participants are identified as particular individuals, the benefit from esteem seeking would increase. In this case our theory suggests engagement in esteem seeking would also increase. While this conclusion seems plausible, it does require experimental verification. Similarly, the type of activity may be important in the extent to which participants esteem seek. Even under anonymous conditions, a person may be more willing to esteem seek when purchasing a new car than when buying credits to improve their score on a real effort task.[5] Again, while this seems plausible, it does require empirical investigation.

The paper illustrates this behaviour is likely to be widespread yet can be understood and studied using very standard economic techniques. If esteem seeking can indeed be modelled using standard techniques, it opens up scope for analysis of a broader range of economic phenomena. For example, a person's demand for employment may depend on both the income and the stream of esteem it generates. Rigorously investigating the economic implications of this, and other similar scenarios, is yet another avenue for fruitful future work.

[5] On this topic, it is worth keeping in mind that Akerlof (2017) argues that the activities which yield esteem are endogenous.

References

Akelof, R. (2017), "Value Formation: The Role of Esteem," Games and Economic Behavior, 102, 2017,

Azmat, G. and Iriberri, N. (2010) "The Importance of Relative Performance Feedback Information: Evidence from a Natural Experiment using High School Students", The Journal of Public Economics, 2010, vol. 94 (7-8), pp. 435-452

Azmat, G. and Iriberri, N. (2016), The Provision of Relative Performance Feedback: An Analysis of Performance and Satisfaction. *Journal of Economics & Management Strategy*, 25: 77–110. doi:10.1111/jems.12151

Benabou, R. and Tirole, J. (2006) "Incentives and Prosocial Behavior", American Economic Review, December, v. 96, iss. 5, pp. 1652-78

Brennann, G. and Pettit, P. (2004) The economy of esteem: An essay on civil and political society, Oxford and New York: Oxford University Press

Bursztyn, L., Fermany, B., Fiorinz, S., Kanzx, M., Rao, G. "Status Goods: Experimental Evidence from Platinum Credit Cards"

Charness, G, Masclet, D. Villeval, M.C. (2010) "Competitive Preferences and Status as an Incentive: Experimental Evidence" IZA DP No. 5034

Charness, G, Masclet, D. Villeval, M.C. (2014) "The Dark Side of Competition for Status" Management Science, 60, January, pp. 38-55

Eckel, Catherine C. and Ball, Sheryl B., (1998) "The Economic Value of Status", *Journal of Socio-Economics*, Vol. 27, No. 4, p. 495, 1998.

Eckel, Catherine C. and Ball, Sheryl B. and Grossman, Philip J. and Zame, William R., (2001) "Status in Markets". Quarterly Journal of Economics, Vol. 116, No. 1, pp. 161-187.

Ellingsen, T. and Johannesson, J., (2008) "Pride and Prejudice: The Human Side of Incentive Theory", American Economic Review, June 2008, v. 98, iss. 3, pp. 990-1008

Festinger, L., (1954) "A Theory of Social Comparison Processes", *Human Relations*, 7, 117-140.

Fischbacher, U., (2007) "z-Tree: Zurich Toolbox for Ready-Made Economic Experiments", Experimental Economics, 10(2), 171-178.

Gill, D. and Prowse, V. (2012) 'A Structural Analysis of Disappointment Aversion, in a Real Effort Competition.' American Economic Review, 102(1): 469--503

John, O., and S. Srivastava (1999): "Big Five Trait Taxonomy. History, Measurement, and Theoretical Perspectives," in Handbook of Personality. Theory and Research, ed. by L. Pervin, and O. John, pp. 102–139. Guilford, New York.

Lang, P. J. (1980) "Behavioral treatment and bio-behavioral assessment: computer applications", In J. B. Sidowski, J. H. Johnson, & T. A. Williams (Eds.), Technology in mental health care delivery systems (pp. 119-l 37). Norwood, NJ: Ablex.

Morozova, Annie, (2015) "What Is in the Stars? The Effect of Status on Social Preferences". Available at SSRN: https://ssrn.com/abstract=2617648 or http://dx.doi.org/10.2139/ssrn.2617648

Mehrabian, A and J. A. Russell (1974). *An approach to environmental psychology* (1 ed.). Cambridge, Mass.: MIT Press.

Rotter, J. (1966): "Generalized expectancies for internal versus external control of reinforcement," *Psychological Monographs: General & Applied*, 80 (1), 1–28.

Tran A. and Zeckhauser, R. (2012) "Rank as an inherent incentive: Evidence from a field experiment", *Journal of Public Economics*, 96, pp. 645–650.

Veblen, Thorstein (1899). The Theory of the Leisure Class. Project Gutenberg.

Appendix A. Additional Mathematical Results and Proofs

A.1 Proof of Proposition 1

Proof: (i) The utility of the role B participants is given by:

$$U^B = \upsilon\pi^B - \sum_{r=1}^{5} h(e_r) + \sum_{r=2,4} \left\{ Yy(s_r^{mB} - s_r^{mA}) - \frac{G}{2}\left(s_r^{mA} - \mu_r^{mA} t_r^{mA}\right)^2 \right\} + 2F$$

The optimal choice of esteem supplied to their partner is given by:

$$Yy'(s_r^{mB} - s_r^{mA}) = G\left(\mu_r^{mA} t_r^{mA} - s_r^{mA}\right) \tag{9}$$

By the implicit function theorem the esteem offered by the role B participant to their role A partner is given by:

$$s_r^{mA}(\mu_r^{mA} t_r^{mA}, K_r; Y, G) \tag{10}$$

Follows from form (9) by noting that $y'(.) > 0$. Taking the total derivative of (10) gives:

$$y'(K_r - s_r^{mA})dY + Yy''(K_r - s_r^{mA})[dK_r - ds_r^{mA}] = G\left(d(\mu_r^{mA} t_r^{mA}) - ds_r^{mA}\right)$$

or:

$$\left[G - Yy''(K_r - s_r^{mA})\right]ds_r^{mA} = Gd(\mu_r^{mA} t_r^{mA}) - y'(K_r - s_r^{mA})dY - Yy''(K_r - s_r^{mA})dK_r$$

The result follows if $y''(.) < 0$.

A.2 Role B effort choice

Define a participant's lone effort, e_r^*, by:

$$w\upsilon v'(\pi^B) = h'(e_r^*)$$

Then:

> **Proposition 5**: In round *r* = 1, 3 and 5 , the role B participant chooses the lone effort level. Lone effort in these rounds increases with the participant's marginal utility of income. In round *r* = 2 and 4 the role B participant chooses effort level 0 if their marginal utility of effort or marginal utility of esteem is sufficiently low, and effort level $\overline{e}^B$ otherwise, where:

Proof: Observe that:

$$\frac{de_i^*}{d\upsilon} = \frac{wv'(\pi^B)}{\left[h''(e_i^*) - w^2\upsilon v''(\pi^B)\right]} > 0$$

Assume that $\overline{e}_2^B = \overline{e}_4^B = \overline{e}^B$. In rounds 2 and 4 the role B participant must choose between effort level 0 with a payment of 10 or effort level $\overline{e}^B$ with a payment of 40. A participant will choose effort 0 (40) if with their combination of $\{\upsilon, Y\}$ satisfies:

$$2h(\overline{e}^B) - 2h(0) + \frac{G}{2}\left[\left((\mu_r^{mA}\hat{t}_r^{mA}, 40) - \mu_r^{mA}\hat{t}_i^{mA}\right)^2 - \left(s_r^{mA}(\mu_r^{mA}\hat{t}_r^{mA}, 10) - \mu_r^{mA}\hat{t}_r^{mA}\right)^2\right]$$

$$< (>)60\upsilon + Y\left[y\left(40 - s_r^{mA}(\mu_r^{mA}\hat{t}_r^{mA}, 40)\right) - y\left(10 - s_r^{mA}(\mu_r^{mA}\hat{t}_r^{mA}, 10)\right)\right]$$

Participants who are indifferent between effort 0 and effort $\overline{e}^B$ have marginal utility of income, $\overline{\upsilon}$, and marginal utility of esteem $\overline{Y}$ which satisfy:

$$2h(\overline{e}^B) - 2h(0) + \frac{G}{2}\left[\left((\mu_r^{mA}\hat{t}_r^{mA}, 40) - \mu_r^{mA}\hat{t}_r^{mA}\right)^2 - \left(s_r^{mA}(\mu_r^{mA}\hat{t}_r^{mA}, 10) - \mu_r^{mA}\hat{t}_r^{mA}\right)^2\right]$$
$$= 60\overline{\upsilon} + \overline{Y}\left[y\left(40 - s_r^{mA}(\mu_r^{mA}\hat{t}_r^{mA}, 40)\right) - y\left(10 - s_r^{mA}(\mu_r^{mA}\hat{t}_r^{mA}, 10)\right)\right]$$

It must be the case that $\overline{e}^B$ has a magnitude sufficient to ensure that half the role B participants receive payment of 10 and the other half a payment of 40.

A role B participant will only provide effort above the lone effort level if doing so lifted them to the upper half of the effort distribution. A role B participant whose effort is in the upper half of the effort distribution will have no incentive to increase it beyond level $\overline{e}^B$. A role B participant whose lone effort is in the lower half of the distribution would only increase their effort levels beyond their lone level if the benefit outweighed the cost. This would mean that their lone effort must be sufficiently close to average effort. The participant with marginal utility $\underline{\upsilon}$ has the lowest.

A.3 Proof of Proposition 2

Define the marginal utility of esteem for a role A participant, $\Delta U(Y)$, as:

$$\Delta U(Y) = \delta Y y'(s_r^{mA} - s_r^{mB}) \qquad (11)$$

where:

$$\delta \equiv \frac{\partial s_r^{mA}}{\partial(\mu_r t_r)}\frac{\partial(\mu_r t_r)}{\partial c_r}$$

ΔU measures the increase of utility of esteem from an extra dollar added to the role A participant's score. Then Proposition 2 follows directly from the following proposition:

Proposition 6: If at the optimum $\upsilon \geq \Delta U(Y)$ then c_r=0 and e_r satisfies:

$$h_r^{'}(e_r) = \upsilon w + w\Delta U(Y) \tag{2}$$

Otherwise e_r sastisfies:

$$h_r^{'}(e_r) = 2w\upsilon \tag{12}$$

and c_r satisfies

$$\upsilon = \Delta U(Y) \tag{4}$$

when e_r is given by (12). In both cases e_r^* is less than the lone effort level.

Proof of proposition 6: Role A participants have the following utility function:

$$\begin{aligned} U^A(e, c_2, c_4, s_2^{mB}, s_4^{mB}) \\ &= \upsilon\left(\sum_{r=1}^{5} we_r - c_2 - c_4\right) - \sum_{r=1}^{5} h(e_r) \\ &+ \sum_{r=2,4} Yy(s_r - s_r^{mB}) - \frac{G}{2}(s_r^{mB} - K_r)^2 + 2F \end{aligned}$$

In rounds r = 2 and 4, the esteem provided by the role A participant to their matched role B participant is:

$$Yy^{'}(s_r - s_r^{mB}) = G(K_r - s_r^{mB})$$

where the participant uses (10) to estimate the esteem from their partner based on their own actions as:

$$s_r = s_r^{mA}(\mu_r^{mA} t_r^{mA}, K_r; \hat{Y}, G)$$

where $\hat{Y}$ is the expected value of the marginal utility of esteem. Equation (1) follows if the utility of esteem is linear. The role A participant's utility in the action stage of round $r \in \{2,4\}$ is thus:

$$U^A(e_r, c_r) = \upsilon(we_r - c_r) - h_r(e_r) + Yy\left(s_r^{mA}(\mu_r^{mA} t_r - s_r^{mB}\right) - \frac{G}{2}(s_{ri}^{mB} - K_{ri})^2 + 2F$$

The Lagrangian for utility maximisation is:

$$L = \upsilon(\pi - c_2 - c_4) - \sum_{r=1}^{5} h_r(e_r) + \sum_{r=2,4} Yy\left(s_r^{mA}(K_r, \mu_r^{mA} t_r^{mA}) - s_r^{mB}(\mu_r^{mA} t_r^{mA}, K_r)\right) - g(s_r^{mB}(\mu_r^{mA} t_r^{mA}, K_r) - K_r) + 2F + \lambda_1 c_r + \lambda_2(\pi_r - c_r)$$

where $\lambda_1 \geq 0$ and $\lambda_2 \geq 0$ are Lagrange multipliers and $g\left(s_r^{mB}\left(\mu_r^{mA} t_r^{mA}, K_r\right) - K_r\right) = \frac{G}{2}(s_r^{mB} - K_r)^2 - F$. The Kuhn Tucker conditions give:

$$\frac{\partial L}{\partial e_r} = w\upsilon + Yy'\left(s_r^{mA}(K_r, \mu_r^{mA} t_r^{mA}) - s_r^{mB}(\mu_r^{mA} t_r^{mA}, K_r)\right) \frac{\partial s_r^{mA}}{\partial(\mu_r t_i)} \frac{\partial(\mu_r t_r)}{\partial e_r} - h_r'(e_r) + w\lambda_2 = 0$$

$$\frac{\partial U}{\partial c_r} = -\upsilon + Yy'\left(s_r^{mA}(K_r, \mu_r^{mA} t_r^{mA}) - s_r^{mB}(\mu_r^{mA} t_r^{mA}, K_r)\right) \frac{\partial s_r^{mA}}{\partial(\mu_r t_r)} \frac{\partial(\mu_i t_r)}{\partial c_r} + \lambda_1 - \lambda_2 = 0$$

Note we have used the envelope theorem, specifically (A), to eliminate terms.

As noted in the text, assume that:

$$w \frac{\partial s_r^{mA}}{\partial(\mu_r t_r)} \frac{\partial(\mu_r t_r)}{\partial c_r} = \frac{\partial s_r^{mA}}{\partial(\mu_r t_r)} \frac{\partial(\mu_r t_r)}{\partial e_r}$$

That is, the matched role B partner cannot tell whether score comes from effort or credit.

Suppose $\lambda_2 = 0$. The FOC become:

$$\upsilon - h_r'(e_r)/w = -\Delta U$$

$$\upsilon = \Delta U + \lambda_1$$

If there is an internal solution, i.e. $\lambda_1 = 0$ and $\lambda_2 = 0$, then optimal effort is chosen by:

$$2w\upsilon = h_r'(e_r) \tag{12}$$

and optimal credits are:

$$\upsilon = \Delta U(Y) \tag{4}$$

If c_r=0, or $\lambda_1 \geq 0$

$$\upsilon - h_r'(e_r)/w = -\Delta U$$

Again, this implies that e_i is greater than the participant's lone level. Further

$$\upsilon = \Delta U + \lambda_1$$

or:

$$2\upsilon - h_r^{'}(e_r)/w = -\lambda_1$$

Hence $u^{'}(we_r) \geq \Delta U$ and $2u^{'}(we_r) \geq h_r^{'}(e_r)/w$.

A.3 Proof of Proposition 3

Now

$$\mathrm{d}\upsilon = \delta y^{'}(s_r^{mA} - s_r^{mB})\mathrm{d}Y + Yy^{'}(s_r^{mA} - s_r^{mB})\frac{\partial(\mu_r t_r)}{\partial c_r}\frac{\partial^2 s_r^{mA}}{\partial(\mu_r t_r)^2}\mu_r[dc_r + wde_r]$$
$$+ \delta^2 Yy^{''}(s_r^{mA} - s_r^{mB})[dc_r + wde_i]\frac{\partial s_r^{mA}}{\partial(\mu_r t_r)}$$

and:

$$2w\mathrm{d}\upsilon = h_r^{''}(e_r)de_r$$

Thus $\frac{de_r}{dY} = 0$, and $\frac{de_r}{\mathrm{d}\upsilon} > 0$:

$$\mathrm{d}\upsilon = \delta y^{'}(s_r^{mA} - s_r^{mB})\mathrm{d}Y + \left[\delta^2 Yy^{''}(s_r^{mA} - s_r^{mB}) + Yy^{'}(s_r^{mA} - s_r^{mB})\frac{\partial(\mu_r t_i)}{\partial c_r}\frac{\partial^2 s_r^{mA}}{\partial(\mu_r t_r)^2}\right]\left[dc_r + \frac{2w^2\mathrm{d}\upsilon}{h_r^{''}(e_r)}\right]\frac{\partial s_i^{mA}}{\partial(\mu_r t_r)}$$

or:

$$\left[1 - \delta^2 \mathrm{Y}y^{''}(s_r^{mA} - s_r^{mB})\frac{\partial s_i^{mA}}{\partial(\mu_r t_r)}\left(\frac{2w^2}{h_r^{''}(e_r)}\right) - Yy^{'}(s_r^{mA} - s_r^{mB})\frac{\partial(\mu_r t_r)}{\partial c_r}\frac{\partial^2 s_r^{mA}}{\partial(\mu_r t_r)^2}\right]\mathrm{d}\upsilon$$
$$- \delta y^{'}(s_r^{mA} - s_r^{mB})\mathrm{d}Y = \delta^2 \mathrm{Y}y^{''}(s_r^{mA} - s_r^{mB})\frac{\partial s_r^{mA}}{\partial(\mu_r t_r)}dc_r$$

Hence:

$$\frac{dc_r}{\mathrm{d}Y} = -\frac{\delta y^{'}(s_r^{mA} - s_r^{mB})}{\delta^2 \mathrm{Y}y^{''}(s_r^{mA} - s_r^{mB})\frac{\partial s_r^{mA}}{\partial(\mu_r t_r)}dc_i} > 0$$

and:

$$\frac{dc_r}{d\upsilon} = \frac{\left[1 - \delta^2 \mathrm{Y} y''(s_r^{mA} - s_r^{mB}) \frac{\partial s_r^{mA}}{\partial(\mu_r t_r)} \left(\frac{2w^2}{h_r''(e_r)}\right) - \mathrm{Y} y'(s_r^{mA} - s_r^{mB}) \frac{\partial(\mu_r t_r)}{\partial c_r} \frac{\partial^2 s_r^{mA}}{\partial(\mu_r t_r)^2}\right]}{\delta^2 \mathrm{Y} y''(s_r^{mA} - s_r^{mB}) \frac{\partial s_r^{mA}}{\partial(\mu_r t_r)} dc_r} < 0$$

A.4 Proof of Proposition 4

To demonstrate proposition 4 it is necessary to demonstrate that the locus $\{\tilde{\upsilon}, \tilde{Y}\}$ has a positive slope. Proposition 4 then follows from Proposition 3. The role A participant would be indifferent between purchasing credits or not purchasing credits, $\tilde{e}_r$, satisfies:

$$2w\tilde{\upsilon} = h'(\tilde{e}_r) \tag{13}$$

Taking the total derivate of (13) gives:

$$h''(\tilde{e}_r)\mathrm{d}\tilde{e}_r = 2w d\tilde{\upsilon}$$

Thus

$$\frac{d\tilde{e}_r}{d\tilde{\upsilon}} = \frac{2w}{h''(\tilde{e}_r)} > 0$$

The marginal utility of esteem which makes the participant indifferent between purchasing credit and not, $\tilde{Y}^K$, satisfies:

$$\tilde{\upsilon} = \delta y'\left(s_r^{mA}(\mu_r^{mA}\tilde{e}_r, K_r; Y, G) - s_r^{mB}\right)\tilde{Y} \tag{14}$$

The participant purchases credits if $Y^K > \tilde{Y}^K$ and not if $Y^K \leq \tilde{Y}^K$. Observe that also:

$$h_r'(\tilde{e}_r) = 2\mathrm{w}\delta y'\left(s_r^{mA}(\mu_r^{mA}\tilde{e}_r, K_r; Y, G) - s_r^{mB}\right)\tilde{Y}$$

Thus:

$$\frac{d\tilde{Y}}{d\tilde{e}_r} = \frac{h_r''(\tilde{e}_r) - 2\mathrm{w}\delta^2 y''\left(s_r^{mA}(\mu_r^{mA}\tilde{e}_i, K_r; Y, G) - s_r^{mB}\right)\tilde{Y}}{2\mathrm{w}\delta y'\left(s_r^{mA}(\mu_r^{mA}\tilde{e}_r, K_r; Y, G) - s_r^{mB}\right)} > 0$$

if $y''\left(s_r^{mA}(\mu_r^{mA}\tilde{e}_r, K_i; Y, G) - s_r^{mB}\right) > 0$.

Appendix B. Estimating Models and Techniques

To test the robustness of our results we present here estimates of the model using alternative econometric techniques.

The data collected from the experiment for 132 participants over 5 rounds was arranged in panel format. Since participants only interact in two rounds (2 and 4) no panel estimation techniques were used. Instead a dummy *r4* for round 4 (round 2 being the base) was included and demographic and personality variables included in **dempsy.**

$$\mathbf{dempsy}^{R} = [\ Female\ \ Age\ \ \ Risk\ \ \ Happiness\ \ Arousal\ \ \ Dominance\ \ \ Extraversion\ \ \ Agreeableness\ \ \ Conscientiousness\ \ Neuroticism\ \ \ Openness\ ]$$

$$\mathbf{dempsy} = \left[\mathbf{dempsy}^{R}\ \ \ Business\ \ \ Eco/Eng/Arc\ \ \ Med/Vet\ \ \ Sci/Oth/None\right]$$

The base case for all estimations that include $\mathbf{dempsy}^{R}$ are age zero, males who have zero for all their personality scores. The base case for all estimations that include **dempsy** are age zero, males, who study Arts or Education ($Arts/Educ = 1$) who have zero for all their personality scores.

B.1 Estimation of the decision to buy credits

Equation (5) in section 4.2.1 contains the two potentially endogenous variables in *credits* and *effort*. Since $\overline{effort}_{135}$ is a function of only exogenous variables it can treated as pre-determined or weakly exogenous. If $\alpha_e = 0$ in equation (5), that is there is no role for *effort* in buying credits, then *buy** is only a function of exogenous variables and the pre-determined $\overline{effort}_{135}$ and so may be estimated by a standard probit. The results of estimating equation (5) under these assumptions over the 132 observations for Role A participants over rounds 2 and 4 are presented in Table 7 in section 4.2.1.

If $\alpha_e \neq 0$ then equation (5) and (7) could be estimated by IV probit but only if only if the amount credits purchased does not affect *effort*, $\psi_c^b = 0$, so that equation (7) is only a function of exogenous and the pre-determined $\overline{effort}_{135}$. These estimation results are available from the authors on request. If $\alpha_e \neq 0$ and $\psi_c^b \neq 0$ then equation (5) can only validly estimated by generalised structural equation modelling as explained in Appendix D.

B.2 Reduced Form Estimation

Solving equations (5), (6), (7), (8) for $buy^{*}_{r,i}$, $effort_{r,i}$, $credits_{r,i}$ and $\overline{effort135}_{i}$ gives the reduced form of the model for $i = 1$ to n individuals and rounds $r = 2$ and $r = 4$.

$$buy^{*}_{r,i} = \theta^{b}_{0} + \theta^{b}_{r4} r4_{r,i} + \theta^{b}_{l} low_{r,i} + \boldsymbol{\theta}^{b}_{d}\mathbf{dempsy}_{i} + \varepsilon^{b}_{r,i} \quad (15)$$

$$\overline{effort135}_{i} = \theta^{\bar{e},buy}_{0} + \boldsymbol{\theta}^{\bar{e},buy}_{d}\mathbf{dempsy}_{i} + \varepsilon^{\bar{e},buy}_{r,i} \quad \text{for } buy = 0, 1 \quad (16)$$

$$effort_{r,i} = \theta^{e,buy}_{0} + \theta^{e,buy}_{r4} r4_{r,i} + \theta^{e,buy}_{l} low_{r,i} + \boldsymbol{\theta}^{e,buy}_{d}\mathbf{dempsy}_{i} + \varepsilon^{e,buy}_{r,i} \quad \text{for } buy = 0, 1 \quad (17)$$

$$credits_{r,i} = \theta^{c}_{0} + \theta^{c}_{r4} r4_{r,i} + \theta^{c}_{l} low_{r,i} + \boldsymbol{\theta}^{c}_{d}\mathbf{dempsy}_{i} + \varepsilon^{c}_{r,i} \quad \text{for } buy = 1 \quad (18)$$

Equations (15), (16), (17) and (18) may be estimated individually by OLS. If the error terms in equations (15) to (18) are contemporaneously correlated with each other, then estimation can be made more efficient by using SUR estimation. We estimate equation (15) over 132 observations together with equations (16) and (17) for subset of 103 non-buyer observations and with equations (16) to (18) for 29 buyer observations using SUR. The estimation results are presented and discussed in Table 8 in section 4.2.

B.2.1 Reduced Form – Tests on Parameters

A test of the null hypothesis that $\theta^{\bar{e},buy=1}_{0} = \theta^{\bar{e},buy=0}_{0}$, the average effort, given demographics, personality, study and partner type is different between buyers and non-buyers of credits, gives a chi-squared statistic of $\chi^2 = 3.25$ and p-value $\Pr(\chi^2 > 3.25) = 0.0716$ and is rejected at the 10% level of significance.

A test of the null hypothesis that $\theta^{e,buy=1}_{0} = \theta^{e,buy=0}_{0}$, the effort, given demographics, personality, study and partner type is different between buyers and non-buyers of credits , gives a chi-squared statistic of $\chi^2 = 11.12$ and a p-value of $\Pr(\chi^2 > 11.12) = 0.0009$ and is rejected at the 1% level of significance.

B.3 Three Stage Least Squares Estimation

Equations (6), (7) and (8) are estimated together for those that buy credits (buy_i=1) using Three Stage Least Squares (3SLS) over 29 observations. Similarly equation (6) and (7) are estimated by 3SLS for those that do not buy credits (buy_i =0) over 103 observations. The results of these two systems of equations with 3SLS are presented in Table 13 in Appendix C.2.

Appendix C. Additional Tables of Statistics and Econometric Results

C.1 Summary statistics

The following table indicates how average effort changes with rounds.

Table 11 Average Effort by Rounds, Roles and Buying

Role:	r = 1	r = 2	r = 3	r = 4	r = 5
A non-buyer	35.4	37.0	39.4	39.7	41.9
A buyer	32.3	33.1	34.5	35.4	36.2
A total	34.5	35.8	38.0	38.5	40.2
B total	33.7	34.7	35.8	36.0	38.1
B total / A non-buyer	0.95	0.94	0.91	0.91	0.91
A non-buyer/ A buyer	1.04	1.05	1.04	1.02	1.05
B total /A total	0.97	0.97	0.94	0.94	0.95

The table below provides evidence that role A participants that bought credits were motivated to do so by being evaluated.

Table 12 Proportion who bought credits by Feedback on Credits

	"Do you think that knowing you would get feedback from RECTANGLE caused you to buy Credits in Stage *r*"				
	Strongly disagree	Disagree	Neither agree or disagree	Agree	Strongly agree
r = 2	0.000	0.118	0.700	0.571	0.800
r = 4	0.000	0.091	0.400	0.600	0.833
Total	0.000	0.103	0.600	0.583	0.818

The spearman rank correlation between the 0/1 decision to buy and feedback on credits is 0.6285 (0.643 in round 2 and 0.602 in round 4) and significant at the 1% level of significance.

C.2 3SLS Estimations

The results of the 3SLS estimation are given in Table 13 on the following page. The first row in Table 13 confirms that if a participant buys credits, they buy significantly more credits when facing a low scoring role B partner, given their demographics and personality. The second row indicates that participants increase their effort in round 4 relative to round 2, by the approximately the same amount regardless of whether a participant buys credits. Table 13 shows when allowing for *effort*, $\overline{effort_{135}}$ in addition to demographics and personality, that role A participants buy significantly more credits when facing a low scoring partner (*low*). This result was also found in the reduced form estimation in Table 8, but the inclusion of *effort* and $\overline{effort_{135}}$ has reduced the size of the by about 40%. In Table 13 the effect of *low* is now significant in reducing the effort of buyers in rounds 2 and 4 (*effort*) given $\overline{effort_{135}}$ and *credits* in addition to demographics and personality, unlike in Table 8 when $\overline{effort_{135}}$ and *credits* where not controlled.

The proxy for the marginal utility of income, $\overline{effort_{135}}$, also significantly reduces the number of credits purchased for buyers, while *effort* significantly increases the number of credits purchased for buyers. The latter result suggests that *effort* and *credits* for buyers are complementary in generating esteem. The system estimates in Table 13 also show that $\overline{effort_{135}}$ also is significant in determining *effort* of those that do and do not buy credits. The effect is stronger on those that do not buy credits, with the number of credits purchased by buyers, *credits*, also playing a significant complementary role to their *effort*.

Table 13 3SLS Estimates

	buyer ($buy = 1$)			non-buyer ($buy = 0$)	
	$\overline{effort135}$	*effort*	*credits*	$\overline{effort135}$	*effort*
r4		2.67***	-4.71***		2.89***
		(0.92)	(1.77)		(0.61)
low		-2.68***	4.75***		-0.63
		(0.88)	(1.57)		(0.59)
$\overline{effort135}$		-0.28	0.48		1.05***
		(0.35)	(0.56)		(0.17)
effort			1.74***		
			(0.32)		
credits		0.58***			
		(0.10)			
Constant	28.58***	51.79***	-89.08***	41.84***	4.85
	(3.55)	(17.11)	(20.88)	(5.07)	(8.32)
Female	-2.75***	-4.22**	7.27***	-4.09***	-0.34
	(0.61)	(1.97)	(2.78)	(1.14)	(1.09)
Age	-0.84***	-1.44***	2.48***	-0.40**	-0.07
	(0.12)	(0.50)	(0.60)	(0.15)	(0.12)
Risk	1.95***	1.75**	-3.00***	-0.28	0.18
	(0.18)	(0.74)	(1.10)	(0.21)	(0.14)
Happiness	1.33***	0.34	-0.59	1.22***	-0.14
	(0.40)	(0.85)	(1.51)	(0.36)	(0.25)
Arousal	0.83***	0.74*	-1.26**	-0.65***	-0.06
	(0.18)	(0.38)	(0.63)	(0.24)	(0.17)
Dominance	0.33	1.91***	-3.30***	0.68**	-0.51**
	(0.26)	(0.60)	(1.01)	(0.30)	(0.25)
Extraversion	-0.25**	0.11	-0.19	0.29	-0.16
	(0.11)	(0.25)	(0.44)	(0.17)	(0.11)
Agreeableness	0.69***	1.15***	-1.99***	-0.09	-0.41***
	(0.12)	(0.39)	(0.57)	(0.19)	(0.12)
Conscientiousness	-0.14	-0.48*	0.82*	0.18	0.13
	(0.11)	(0.27)	(0.49)	(0.23)	(0.16)
Neuroticism	-0.29	1.08***	-1.86***	0.23*	-0.18*
	(0.22)	(0.37)	(0.61)	(0.13)	(0.10)
Openness	0.21	-2.40***	4.15***	-0.29	-0.04
	(0.42)	(0.84)	(1.33)	(0.25)	(0.17)
Arts/Educ	base			base	
Business	4.33***			4.55***	
	(0.60)			(1.68)	
Eco/Eng/Arc	8.90***			7.27***	
	(1.27)			(1.90)	
Med/Vet	-6.58***			2.46	
	(1.16)			(1.81)	
Sci/Oth/None	9.73***			2.51*	
	(1.16)			(1.48)	
R-Squared	0.97	0.76	0.64	0.41	0.79

C.3 Feedback and Fairness Estimation

Theory suggests that the role B's feedback, d_i^*, will be related to their matched role A participant's reported score. The cost of deviating from objective scoring may also differ across role B participants according to their demographic characteristics. Hence we adopt the following estimating equation for feedback given by B:

$$d^*_{r,i} = \phi^B_o + \phi^B_1 r4_{r,i} + \phi^B_2 low_{r,i} + \phi^B_3 \overline{effort135}_i + \phi^B_4 effort_i + \phi^B_5 score^A_{r,i} + \phi^B_6 \mathbf{dempsy}_i + \varepsilon^{d^*}_{r,i} \tag{19}$$

Role A participant's perception of fairness $f^A_{r,i}$ of the feedback they receive is modelled as a function of demographic and psychological characteristics **dempsy**, *r4*, *low*, *effort* $\overline{effort135}$ and *credits* and dummy variables for the ordinal discrete feedback they receive $d_{r,i} \in \{1,2,3,4,5\}$.

$$\begin{aligned} f^{*A}_{r,i} = \varphi^A_0 + \varphi^A_1 r4_{r,i} + \varphi^A_2 low^B_{r,i} + \varphi^A_3 \overline{effort135} + \varphi^A_4 effort_{r,i} + \varphi^A_5 credits_{r,i} \\ + \sum_{k=1}^{5} \varphi^A_{6+k} 1(d_{r,i} = k) + \boldsymbol{\varphi}^A_{12} \mathbf{dempsy}_i + \varepsilon^{f^{A^*}}_{r,i} \end{aligned} \tag{20}$$

Role B participant's perception of fairness of the feedback they gave A is modelled in similar way to Role A's in equation (20) but with the addition of the score of their matched role A partner as given in equation (21).

$$\begin{aligned} f^{B^*}_{r,i} = \varphi^B_0 + \varphi^B_1 r4_{r,i} + \varphi^B_2 low_i + \varphi^B_3 \overline{effort135}_i + \varphi^B_4 effort_{r,i} + \varphi^B_5 score^A_{r,i} \\ + \sum_{k=1}^{5} \varphi^B_{6+k} 1(d_{r,i} = k) + \boldsymbol{\varphi}^B_{12} \mathbf{dempsy}_i + \varepsilon^{f^{B^*}}_{r,i} \end{aligned} \tag{21}$$

Equation (19) and (21) are estimated for 61 Role B participants over rounds 2 and 4 for a total of 132 observations. Equation (20) is estimated for 61 Role A participants over rounds 2 and 4 for a total of 132 observations. The results of the estimating equations (19), (20) and (21) as ordered probits using maximum likelihood via Stata's modified Newton-Raphson algorithm with robust standard errors are reported in Table 14 and Table 15.

Table 14 Feedback and Fairness - Ordered Probits – Part A

	B's Feedback	B's Fairness	A's Fairness
r4	-0.01	0.22	-0.12
	(0.20)	(0.21)	(0.22)
low	0.03	-0.09	-0.04
	(0.28)	(0.27)	(0.20)
$\overline{effort}_{135}$	-0.02	0.04**	0.03
	(0.02)	(0.02)	(0.04)
effort	0.01	-0.03	-0.01
	(0.02)	(0.02)	(0.04)
credits			0.01
			(0.03)
$score^{A}$	0.10***	-0.02	
	(0.02)	(0.02)	
feedback=1		-2.37*	-1.77***
		(1.22)	(0.54)
feedback=2		0.34	-0.74
		(0.73)	(0.49)
feedback=3		base	base
feedback=4		0.19	0.82***
		(0.26)	(0.31)
feedback=5		0.89**	1.90***
		(0.37)	(0.37)
cut1	0.59	-1.87	1.27
Constant	(1.21)	(1.29)	(1.43)
cut2	1.50	-0.88	2.05
Constant	(1.25)	(1.21)	(1.43)
cut3	2.62**	-0.03	2.93**
Constant	(1.29)	(1.24)	(1.41)
cut4	4.06***	1.92	4.72***
Constant	(1.29)	(1.26)	(1.43)
R-squared(pseudo)	0.20	0.13	0.23
Loglikelihood(pseudo)	-132.62	-120.21	-137.68

Table 15 Feedback and Fairness - Ordered Probits – Part B

	B's Feedback	B's Fairness	A's Fairness
Female	0.48*	-0.26	-0.03
	(0.26)	(0.22)	(0.25)
Age	0.01	-0.01	0.06
	(0.02)	(0.02)	(0.04)
Risk	0.04	-0.05	-0.01
	(0.05)	(0.05)	(0.04)
Happiness	-0.03	0.07	-0.01
	(0.06)	(0.07)	(0.08)
Arousal	-0.08*	0.03	-0.05
	(0.05)	(0.05)	(0.05)
Dominance	-0.11*	0.09	-0.03
	(0.06)	(0.07)	(0.07)
Extraversion	0.01	0.02	0.08**
	(0.04)	(0.04)	(0.04)
Agreeableness	-0.04	0.04	0.12***
	(0.05)	(0.05)	(0.04)
Conscientiousness	0.05	0.02	-0.04
	(0.04)	(0.04)	(0.05)
Neuroticism	0.06*	-0.01	-0.02
	(0.03)	(0.04)	(0.03)
Openness	0.02	0.12**	0.02
	(0.06)	(0.06)	(0.07)
Arts/Educ	base	base	base
Business	0.49	-0.55	0.86***
	(0.35)	(0.37)	(0.31)
Eco/Eng/Arc	-0.34	-0.53*	0.22
	(0.33)	(0.32)	(0.37)
Med/Vet	0.34	-0.13	0.29
	(0.54)	(0.38)	(0.31)
Sci/Oth/None	0.61*	-0.26	0.16
	(0.34)	(0.34)	(0.29)

Appendix D. Alternative estimation strategy: Generalised Structural Equation Model (GSEM)

The reduced form estimates do not control for the endogenous variables and while the 3SLS systems for buyers and non-buyers allow for $\overline{effort135}$, *effort* and *credits* to be endogenously related, they do not allow for the decision to buy to be endogenous. In order to simultaneously cater for the endogenous variables, *buy**, *effort* and *credits* and the contemporaneous correlation cross the error terms that it creates, latent variables $L1_i$ and $L2_i$ are added to the model as described below.

$$buy_i^* = \alpha_0 + \alpha_{r4} r4_i + \alpha_l low_i + \alpha_{\bar{e}} \overline{effort135}_i + \alpha_e effort_i + \boldsymbol{\alpha}_d \mathbf{dempsy}_i^R + \alpha_{L1} L1_i + L2_i + \varepsilon_i^b \quad (22)$$

$$\overline{effort135}_{r,i} = \zeta_0 + \boldsymbol{\zeta}_d \mathbf{dempsy}_{r,i} + L1_{r,i} + \varepsilon_{r,i}^{\bar{e}} \quad (23)$$

$$effort_{r,i} = \psi_0^{buy} + \psi_{lr4}^{buy} r4_{r,i} + \psi_l^{buy} low_{r,i} + \psi_{\bar{e}}^{buy} \overline{effort}_{r,i} + \psi_c^{buy} credits_i + \psi_{L2}^{buy} L2_i + \boldsymbol{\psi}_d^{buy} \mathbf{dempsy}_i^R + \varepsilon_i^{e,buy} \quad \text{for } buy = 0, 1 \quad (24)$$

$$credits_{r,i} = \beta_0 + \beta_{r4} r4_{r,i} + \beta_{\bar{e}} \overline{effort}_{r,i} + \beta_e^b effort_{r,i} + \beta_l^b low_{r,i} + \boldsymbol{\beta}_d^b \mathbf{dempsy}_i^R + \beta_{L2} L2_{r,i} + \varepsilon_{r,i}^c \quad \text{if } buy = 1 \quad (25)$$

We estimate the generalised structural equation model (GSEM) given by equations (22) through to (25) with maximum likelihood using Stata's mean–variance adaptive Gauss–Hermite quadrature algorithm with the results given in Table 16 on the following page. The GSEM model has the advantage over the two 3SLS estimations in that the decision to buy credits is explicitly included the model. This allows *buy** to effect $\overline{effort135}$, *effort* for non-buyers, and *effort* and *credits* for buyers and vice versa, through the latent variables *L1* and *L2*. The coefficient estimate for *L1*, α_{L1} in Table 16 is statistically significant and negative, indicating that there is a significant negative endogenous relationship between $\overline{effort135}$ and *buy**. The coefficient estimates for *L2* in Table 16 are statistically significant in *credits* equation for buyers, but insignificant in buyers *effort* equation. The significant positive estimate for β_{L2} in credits indicates that there is a significant positive relationship between *buy** and *credits*.

The other results of the GSEM estimation results in Table 16 are broadly in line with the previous reduced form results in Table 7 and Table 8 and 3SLS results in Table 13. The significant negative effect of $\overline{effort135}$ and the insignificant role of *effort* in the decision to buy credits (*buy* in column 2) in Table 16, confirms our earlier semi-reduced form estimation results in Table 7. Recall $\overline{effort135}$ indicates a participants' marginal utility of income in the

absence of esteem, while *effort*, where role A participants are evaluated and may include the effects of esteem. Thus Table 16 confirms that $\overline{effort135}$, the marginal utility of income in the absence of esteem significantly reduces the probability of buying credits, whereas *effort* which may include the effects of esteem, is not significant in the decision to buy credits.

The main difference in the 3SLS estimates in Table 13 and the GSEM estimates in Table 16 is for the role of $\overline{effort135}$ our proxy for the marginal utility of income in the absence of esteem, on the *effort* and credits for *buyers*. The 3SLS estimates that did not consider the decision to buy as endogenous, did not find any significant role for $\overline{effort135}$ in the *credits* or *effort* of buyers, only a significant positive effect on buyers *effort*. The GSEM estimates in Table 16 show that by considering the decision to buy endogenously, $\overline{effort135}$ is significant in increasing the *effort* of both buyers and non-buyers and in reducing the number of *credits* purchased by buyers.

Table 16 GSEM Estimates

	$\overline{effort_{135}}$	*buy*	$effort^{nb}$	$effort^{b}$	*credits*
r4		-1.45	2.89***	2.51***	-0.77
		(1.39)	(0.56)	(0.86)	(2.43)
low		0.39	-0.64	-1.31	3.12*
		(1.25)	(0.58)	(1.00)	(1.71)
$\overline{effort_{135}}$		-0.46**	0.97***	0.93**	0.87
		(0.23)	(0.06)	(0.45)	(0.75)
effort		-0.02			0.61
		(0.20)			(0.67)
credits				-0.00	
				(0.30)	
Constant	41.20***	20.29**	8.41**	-14.14	-81.52***
	(4.85)	(8.02)	(3.87)	(26.43)	(23.12)
Female	-4.23***	-4.15***	-0.73	2.98	8.89***
	(1.09)	(1.45)	(0.72)	(3.35)	(3.29)
Age	-0.33**	-0.24	-0.10	0.51	2.57***
	(0.13)	(0.20)	(0.10)	(0.82)	(0.81)
Risk	0.10	0.28	0.16	-0.19	-2.07**
	(0.22)	(0.28)	(0.12)	(0.70)	(0.88)
Happiness	1.03***	-0.37	-0.09	0.18	0.74
	(0.33)	(0.45)	(0.22)	(0.51)	(0.95)
Arousal	-0.53**	-0.28	-0.10	0.05	-1.34***
	(0.22)	(0.32)	(0.13)	(0.48)	(0.42)
Dominance	0.52	0.11	-0.44**	0.80	-2.03**
	(0.38)	(0.43)	(0.22)	(0.63)	(0.91)
Extraversion	0.09	0.16	-0.15	0.19	0.05
	(0.17)	(0.21)	(0.11)	(0.20)	(0.32)
Agreeableness	0.07	-0.01	-0.42***	0.06	-1.33**
	(0.17)	(0.25)	(0.11)	(0.43)	(0.54)
Conscientious	0.01	0.57***	-0.16*	0.39	-0.97*
	(0.13)	(0.18)	(0.08)	(0.29)	(0.57)
Neuroticism	-0.31	-0.66**	-0.06	-0.60	2.31**
	(0.28)	(0.31)	(0.19)	(0.79)	(1.11)
Openness	-0.05	0.18	0.14	-0.35	0.11
	(0.18)	(0.30)	(0.17)	(0.23)	(0.35)
Business	3.26**				
	(1.48)				
Eco/Eng/Arc	5.39***				
	(1.50)				
Med/Vet	2.21				
	(1.94)				
Sci/Oth/None	2.52**				
	(1.27)				
LatentVar1	1.00	-3.42***			
	(.)	(0.40)			
LatentVar2		1.00		0.30	1.85***
		(.)		(0.57)	(0.54)