

The Role of Cognitive Ability, Personality Traits and Gender in Gift Exchange Outcomes*

Emel Filiz-Ozbay
Department of Economics
University of Maryland

John C. Ham
Department of Economics
University of Maryland
IFAU, IFS, IRP (UW) and IZA

John H. Kagel
Department of Economics
Ohio State University

Erkut Y. Ozbay
Department of Economics
University of Maryland

January 10, 2013

Abstract

Controlling for cognitive ability, personality characteristics and gender effects in a gift exchange experiment, men offer higher wages compared to women, as do agents with greater cognitive ability and those scoring higher on agreeableness on the Big Five personality scale. In turn, men provide greater effort than women do. For both genders, a one standard deviation increase in agreeableness generates almost the same increase in effort as a comparable increase in wages. Further, conscientiousness plays a statistically and economically significant role in wage rates offered and effort levels provided, but the sign of this effect differs between men and women.

Key words: gift exchange experiment, Big Five personality characteristics, SAT scores.

JEL classification: D03, J16, J22.

We gratefully acknowledge helpful comments from Hal Arkes, Jennifer Cheavens, Sergio Urzua and participants at the Ohio State University Applied Micro Brown Bag seminar, the University of Binghamton, and the University of North Carolina, Chapel Hill . We thank Kristian Lopez-Vargas for programming the experimental software. Kagel's research was partially supported by NSF grants SES-1226460 and SES-0924764. Opinions, findings, conclusions or recommendations in this material are those of the authors and do not necessarily reflect the views of the National Science Foundation. We are responsible for any errors.

In economics, there is growing interest in the effects of personality traits, cognitive ability, gender, and ethnicity, on economic outcomes using field data (Borghans et al., 2008; Heckman et al, 2006). The present paper looks at the impact of these factors, on behavior in a laboratory gift exchange experiment. Results from laboratory gift exchange experiments have implications for the social preference literature, as well as for labor economics, as they can explain various labor market phenomena that are puzzles from the viewpoint of standard economic theory (e.g., wage rigidity, rent sharing, and efficiency wages).¹ The role of cognitive ability and personality traits is a generally unexplored area in experimental economics in general, and in the study of social preferences in particular.

Our experiment focuses on the effects of cognitive ability, as measured by SAT scores, and personality traits, as measured by the Big Five Inventory (BFI) questionnaire (John et al., 2008). SAT scores serve as a readily available measure of cognitive ability with a substantial positive correlation with other measures of cognitive ability such as the Armed Services Vocational Aptitude Battery, along with a variety of traditional IQ measures (Frey and Detterman, 2004). The 44-item BFI was developed with the goal of creating a brief inventory that would allow efficient and flexible assessment of the Big Five personality characteristics when there is no need for more differentiated measurement of its individual elements. The properties of the BFI are discussed in the body of the paper, but for the moment, we simply note that the BFI provides measures of a person's agreeableness, extroversion, conscientiousness, neuroticism and openness.

We report substantial gift exchange with the usual pattern reported in the literature: higher wages result in substantially higher average effort levels, which are mutually profitable for both “managers” and “workers.” For the pooled data where gender differences are captured only by a gender dummy variable, there are significant differences between men and women both as managers and as workers, with women offering lower average wages, other things equal, and providing lower effort at comparable wage rates than men. When we run separate models for men and women, in both cases higher SAT scores are associated with higher wages, with those scoring higher on the agreeableness trait providing higher wages as well. The latter seems intuitively obvious and is consistent with other results reported in the literature (e.g. Becker et al., 2012).

However, the importance of including SAT scores, or some companionable measure of cognitive ability, in exploring the Big Five personality characteristics is evidenced by the fact that agreeableness loses its statistical significance in our wage setting Tobit equations absent SAT scores, being replaced by a statistically significant, negative effect of extroversion on wage rates offered. Moreover, the importance of allowing for gender differences over the full set of personality characteristics is demonstrated by the fact that conscientiousness plays a statistically and economically significant role in wage rates offered and effort levels provided, but the sign of this effect differs between men and women in terms of both wages offered and effort responses. The

¹ See Gächter and Fehr (2001) and Cooper and Kagel (2012a) for elaboration of these points.

impact of personality characteristics on behavior can be quite substantial: A one standard deviation increase in agreeableness increases the effort response index function of men and women by 87.1% and 106.3%, respectively, compared to the effect of a one standard deviation increase in wages of 135.4% and 100.0% for men and women.

There have been scattered prior attempts to experimentally investigate the effect of personality traits on economic outcomes with a strong social preference element.² Kurzban and Houser (2001) allocated individuals into four groups based on their behavior in a voluntary contribution mechanism (VCM) public good game, and then looked at the role of the Big Five personality characteristics, along with other personality measures, on the probability of an individual falling in a given group using a multinomial logit model.³ They found no statically significant relationship between subject type and any of the Big Five characteristics.⁴ They attribute this finding, in part, to their relatively small sample size (57 subjects). Pothos et al. (2010) investigated the individual correlations between the Big Five components on cooperation in a simultaneous move, one-shot prisoner's dilemma game and found that more agreeable types were less likely to defect ($p = 0.054$).⁵ Becker et al. (2012) investigated the relationship between the Big Five personality characteristics and behavior in a variety of one-shot games with strong social preference elements - the trust game, investment in punishment in a public good (VCM) game, and giving in the dictator game. They focus on direct correlations between the Big Five personality measures and measures of the strength of social preferences in these games.⁶ Among all the personality characteristics, agreeableness had the largest and most significant correlations in their study, being positively correlated with second mover returns and first mover allocations in the trust game, along with giving in the dictator game, and negatively correlated with punishment in the VCM game. Openness had the same qualitative (but weaker quantitative) relationship to social preferences in these games. None of these papers controlled for (any measure of) cognitive ability in their analysis, nor do they investigate the effect of gender differences over the full set of characteristics.

Anderson et al. (2011) is the paper in the literature closest to ours. Using a large sample of truck driver trainees, they measure individual risk and time preferences, along with obtaining scores for the Big Five personality characteristics and cognitive ability based on the Cognitive Skill Index. They run regressions investigating the impact of these variables on a number of life outcomes (e.g., filing for bankruptcy, smoking,

² We are confining this short revue of prior work to those papers looking at the Big Five and other personality characteristics to the social preference literature, which our paper deals most directly with. As Borghans et al., (2008) note there has been limited research on personality characteristics on this important domain. For a survey regarding the power of personality traits both as predictors and as causes of academic and economic success, health, and criminal activity see Almlund et al. (2011).

³ There groups are (1) free riders; (2) cooperators; (3) conditional cooperators

⁴ They report that men were more cooperative than women, while those with higher self-esteem (not one of the Big Five) tended to free ride more.

⁵ They employed procedures very different from the typical economic experiment, using loaded terms (e.g. "defect" or "cooperate") with no financial incentives.

⁶ These direct correlations do not control for any of the other Big Five characteristics. They also check for potential non-linear relations which would compromise the correlation analysis, but report minimal nonlinearity's.

and credit scores), while also controlling for socio-economic characteristics (e.g., age, education and marital status). Their overlap with our paper comes when they turn their attention to behavior in a modified (one-shot) trust game in which first movers could send either \$0 or \$5, and second movers responded via the strategy method.⁷ To the best of our knowledge, they are the only other experimental study to control simultaneously for personality characteristics and cognitive ability in a laboratory type setting. They found that more agreeable types and those scoring higher on cognitive skills were more likely to send the \$5, while more conscientious types were less likely to do so. More agreeable and neurotic types were more likely to send money back in response to either a \$0 or \$5 transfer, with higher cognitive ability types sending less (more) back in response to a \$0 (\$5) transfer. In addition, women were less likely than men ($p < 0.10$) to send money back in response to receiving \$5.⁸

Our paper compliments their analysis of social preferences in many ways. First, we investigate social preferences in an experimental environment where subjects gain extensive experience with the contingencies is typically been considered to be important in terms properly evaluating behavioral responses. Second, we investigate the effect of cognitive ability and personality characteristics separately for men and women, finding a number of interesting differences, and consider the bias in estimated Big Five effects that occurs when SAT is omitted. Third, there is some concern with their analysis is that since cognitive ability and personality traits will affect demographic variables such as education and marital status (Almlund et al., 2011), they cannot estimate the full effect of cognitive ability and personality traits on their outcome variables, which we do. Fourth, we use a sample that is much more comparable to those used by previous experimenters, and this allows us to suggest that previous results on personality characteristics are likely to be have been misleading since they did not control for cognitive ability and a full set of gender differences.

The structure of the paper is as follows: Section 1 provides a characterization of the Big Five personality characteristics, focusing on what they are designed to identify, along with a justification for using SAT scores to measure cognitive ability. Section 2 reports the experimental design and procedures. Experimental results are reported in Section 3, starting with the usual measures reported for gift exchange experiments, and then moving on to the main analysis regarding the effects of gender, cognitive ability and the Big Five characteristics on outcomes reported. Section 4 summarizes the main results of the paper and suggests additional areas of study.

⁷ They also look at the relationship between the Big Five and a variety of life outcomes for their sample. Under the strategy method, second movers state their actions contingent on first mover behavior and then are bound by these actions.

⁸ A large portion of their sample of truck drivers were unconditional cooperators, sending money back regardless of whether or not they were sent anything.

1. Big Five Personality Measures and SAT Scores

Prior to the start of each session, subjects filled out the 44-item Big Five Inventory (BFI) questionnaire (John et al., 2008).⁹ The Big Five personality characteristics represent a consensus among personality psychologists on a general taxonomy of personality traits. These personality characteristics do not represent a particular theoretical perspective but are derived from the analysis of the natural language terms people use to describe themselves and others. The focus of the Big Five is on internal consistency rather than predictive ability. The idea behind the Big Five is not that these personality characteristics reflect any intrinsic greatness or that personality differences can be reduced to five traits but that the five dimensions represent personality at a very broad level of abstraction; each dimension summarizes a large number of distinct, more specific, personality characteristics. When more factors than the Big Five have been identified across cultures and studies, they rarely replicate across multiple studies conducted by independent investigators.

The BFI measure consists of 44 short phrases based on trait adjectives known to be prototypical markers for the Big Five. For example, the openness adjective *original* becomes “Is original, comes up with new ideas” in the BFI. A number of different instruments are available to measure the Big Five, none of which are considered the gold standard. The BFI has been used frequently in research settings where time is at a premium, as it typically takes between 10 and 15 minutes to complete. The Big Five personality traits consist of: (1) Agreeableness – contrasts a pro-social and communal orientation to others, and includes traits such as altruism, tender-mindedness, trust and modesty; (2) Extroversion – implies an energetic approach toward the social and material world, including traits such as sociability, activity, assertiveness and positive emotionality; (3) Conscientiousness – describes socially prescribed impulse control that facilitates task- and goal-directed behavior, such as thinking before acting, delaying gratification, following norms and rules, and planning, organizing and prioritizing tasks; (4) Neuroticism – contrasts emotional stability and even-temperedness with negative emotionality, such as feeling anxious, nervous, sad and tense; and (5) Openness – describes the breadth, depth and complexity of an individual’s mental and experiential life. Scoring higher on the scale of each characteristic is associated with the more positive elements of the traits, except for neuroticism, where the high pole is associated with poorer coping with life events.

SAT scores were used as the proxy measure for cognitive ability (denoted by *g*). We use SAT scores as they are readily available through the registrar’s office (with subjects having signed consent forms to this effect)

⁹ These were computerized using zTree (Fischbacher, 2007), the software used to program the experiment. The BFI questionnaire is available at <http://www.outofservice.com/bigfive/>. The material in this section on the Big Five is taken from John et al. (2008) and Broghans et al. (2008).

and the fact that these scores are highly correlated with various measures of cognitive ability.¹⁰ With respect to the latter, Frey and Dutterman (2004) extracted measures of *g* from the Armed Services Vocational Aptitude Battery (ASVAB) and correlated these with SAT scores for 917 subjects, with a correlation of 0.82 (0.86 corrected for nonlinearity).¹¹ Further, they found that simple correlations between SAT and various IQ tests available for subsets of their data ranged from 0.53 to 0.82, with the SAT correlating significantly ($p < 0.05$) with all six of the traditional IQ tests examined.¹² They also investigated the correlation between SAT scores and Raven's Advanced Progressive Matrices among 104 undergraduates, and obtained a correlation of 0.48 (0.72 for restricted range).

The Big Five personality characteristics are on a scale of 1 to 4, with SAT on a scale of 400-1600. In the data analysis that follows, we convert these scores to the percent of maximum possible score (POMP). Specifically for individual *i*

$$\text{POMP}_i = \frac{\text{Observed}_i - \text{Minimum}}{\text{Maximum} - \text{Minimum}}$$

where *Observed_i* is the observed score for individual *i*, *Minimum* is the minimum possible score on the scale, and *Maximum* is the maximum possible score on the scale. POMP scores have a number of desirable characteristics for indexes of this sort, particularly in terms of the regression coefficient estimates that follow, as it puts them on a normalized scale that helps in interpreting the results (Cohen et al., 1999). Since POMP is a linear transformation of the original scores, statistical evaluation of the data remains unchanged. Average POMP scores, along with their ranges and standard deviations, are reported in Table 1.

¹⁰ In our experience, asking for permission to collect SAT scores and other material (e.g., grade point average, high school class rank, major, etc.) available from the registrar's office as part of the consent form has no impact on recruiting subjects.

¹¹ The Armed Services scores were extracted from the 1979 National Longitudinal Survey of Youth.

¹² The authors note that these results must be treated with caution, as some of the sample sizes were small.

Table 1
POMP Scores for the Big Five and SAT
Scale

Group	N	Statistic	Scale					
			SAT	Ope	Con	Ext	Agr	Neu
Total Population	192	Mean	72.7	67.7	66.3	59.1	67.4	41.2
		SD	11.8	15.0	13.9	18.6	16.1	18.9
		Range	44 - 98	22 - 100	25 - 98	12 - 97	13 - 100	3 - 94
Males	105	Mean	74.0	67.6	65.1	58.2	68.2	38.4
		SD	12.4	14.6	14.2	17.7	16.5	18.5
		Range	44 - 98	22 - 100	25 - 95	21 - 97	25 - 100	3 - 94
Females	87	Mean	71.1	67.8	67.8	60.2	66.4	44.5
		SD	10.9	15.5	13.5	19.5	15.6	18.9
		Range	45 - 97	25 - 90	30 - 98	12 - 97	13 - 95	9 - 91

Ope = Openness; Con = Conscientiousness; Ext = Extroversion; Agr = Agreeableness; Neu = Neuroticism

2. Experimental Design and Procedures

After the completion of the BFI questionnaire, subjects were given written instructions, which were also read out loud.¹³ After reading the instructions, subjects were randomly divided into two equal size groups: “managers” and “employees.” Subjects played the same role throughout an experimental session. Each experimental session had 12 periods, with the number of periods announced in advance. At the beginning of each period, each manager was randomly matched with an (anonymous) employee to play the two stage gift exchange game. There were 16 subjects in each session, with the random matching protocol programmed so that no employee was re-matched with the same manager more than twice and never re-matched in two consecutive periods. The anonymity, in conjunction with subjects learning the outcome of only their own match, was designed to generate a sequence of one-shot games.

In stage 1 of each period, managers chose a wage, which had to be an integer value from the interval [0,100]. In stage 2, each employee, after seeing the wage offer, chose an effort level, which also had to be an integer value from the interval [0,100], after which each manager observed the effort level of the employee he

¹³ The instructions of the experiment are available at www.econ.umd.edu/~filizozbay/FHKO_instructions.pdf.

or she was paired with.¹⁴ Payoffs were symmetric and calculated as follows for managers (π_M) and employees (π_E):

$$\pi_M = 100 - w + 5e$$

$$\pi_E = 100 - e + 5w$$

where w is the wage offered and e is the effort level chosen. Payoff functions were common knowledge, with subjects asked to calculate the payoffs of managers and employees in five examples before the experiment started, with the goal of ensuring that subjects understood the payoffs for themselves and the player they were paired with.

Assuming that players care only about own income, the unique subgame perfect Nash equilibrium of the game is to provide zero effort after any wage offer, in anticipation of which wage offers are zero as well. On the other hand, the efficient wage and effort level, which maximize total surplus, is 100 for both.

Twelve sessions were conducted at the Experimental Economics Laboratory at University of Maryland. All sessions were computerized using zTree software (Fischbacher, 2007). Subjects were recruited from the undergraduate population of University of Maryland using an online recruitment system. Sessions lasted about 60 minutes, including answering the BFI questionnaire. Subjects were paid privately and individually at the end of each session at a rate of 250 experimental currency units (ECUs) to 1 US dollar along with a \$6 participation fee. Average earnings were approximately \$21.75 for employees and \$14.40 for managers (including the \$6 show-up fee).

3. Experimental Results

Aggregate outcomes are reported first in the way they are commonly reported for experiments of this sort. More detailed analysis follows, accounting for the impact of the BFI personality traits and SAT scores on behavior. Results are summarized in terms of a series of conclusions following the data reported.

3.1 Overview of Experimental Results

Figure 1 shows average wage offers and effort levels over time for the pooled data. Average wage offers were 50.7 ECUs with an average effort of 25.2 ECUs. Figure 2 shows effort as a function of the wage rate offered, with average effort clearly increasing with increases in the wage rate.¹⁵ Although there was some heterogeneity between sessions, similar patterns are reported in all sessions. Observe that average wage offers are persistently higher than average effort levels, with neither of them close to the strictly selfish, own income maximizing, equilibrium of zero wage and zero effort. This was true for all sessions with only 12.0% of all wage offers and 37.3% of all effort choices equal to zero for the pooled data. Note that the vast majority of zero wage offers

¹⁴ In addition to providing payoff information for the current period, each subject's computer screen reported the whole history of that subject's play and payoffs.

¹⁵ Following the standard for reporting in this literature, we include zero wage offers here. However, zero wage offers are overwhelmingly met with zero effort, so we delete the zero wage observations when analyzing responses using the Tobit analysis in sections 3.2 and 3.3.

(92.0%) were met with zero effort.¹⁶ Excluding these zero wage offers, 29.9% of all effort choices were zero for the pooled data.

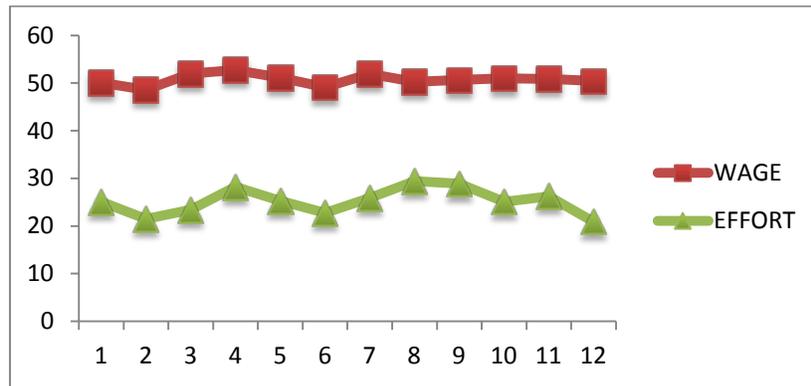


Figure 1. Average wage and effort level per period

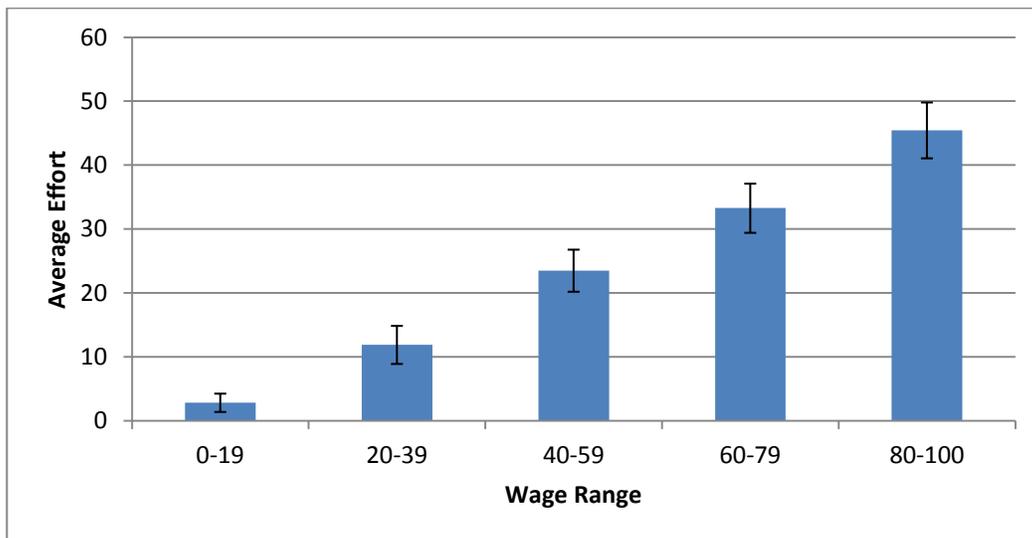


Figure 2. Effort level over each wage range for aggregate data (bars indicate 95% confidence interval for the mean)

Conclusion 1: Aggregate wage and effort level data exhibit the typical pattern reported for gift exchange games of this sort: Average wages are higher than average effort levels, with more effort provided in response to higher wages. Further, the data show minimal, or non-existent, end game effects.

Figure 3 reports average managers’ payoffs in terms of the wage rate offered, net of the base rate of 100 ECUs in the managers’ payoff function (π_M) (where error bars indicate the 95% confidence interval for mean payoffs). As is typically the case, managers who offer higher wages are rewarded with substantially higher average incomes as a result of the substantially higher effort levels workers provide. But, as Figure 4 shows,

¹⁶ Of the 138 zero wage offers, 127 were met with zero effort.

these higher incomes are not without risk, as 20% or more of the effort responses at higher wage rates are met with zero effort.

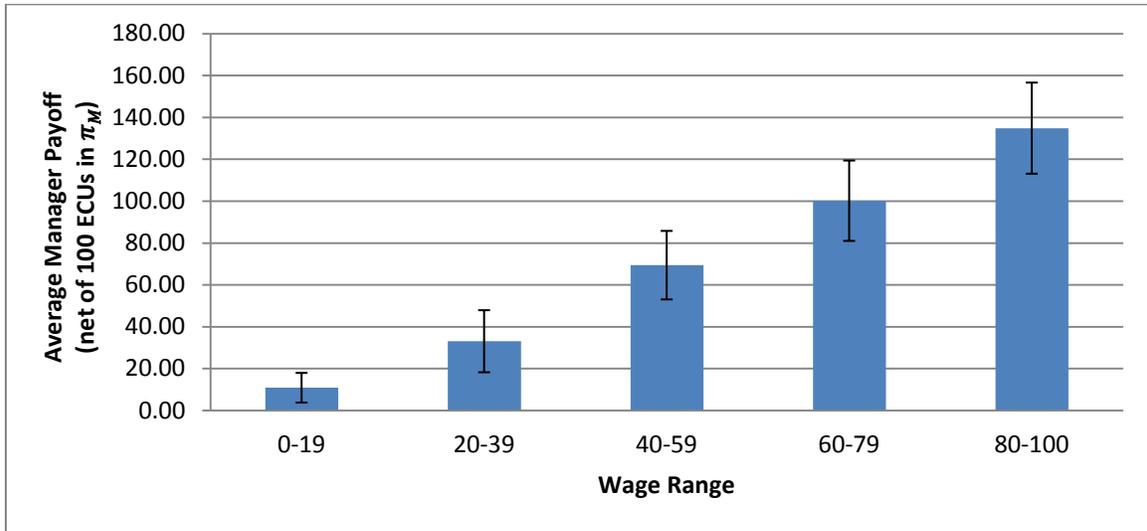


Figure 3: Average income of managers at each wage interval (net of the 100 ECUs included in π_M)

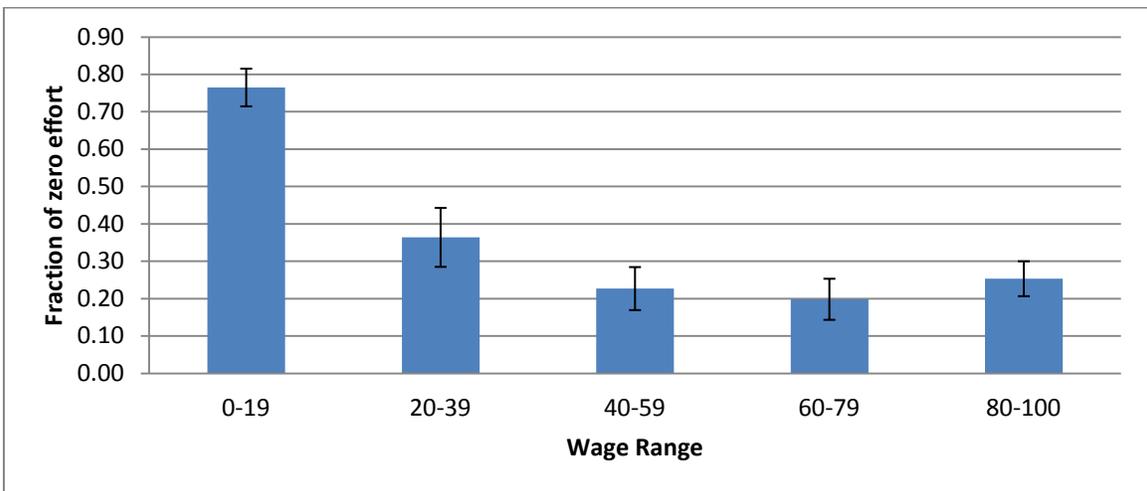


Figure 4: Fraction of zero effort at each wage interval

Conclusion 2: Offering higher wages yields greater expected income for managers. But there are risks involved on account of a persistent percentage of zero effort responses regardless of the wage rate offered.

Figure 5 shows that although there are minimal differences in effort responses between men and women at lower wages, at middle and higher wage rates men consistently provide greater average effort than women. Figure 6 shows that men also tend to offer higher wages than women with most of this tendency accounted for as a result of men providing wages offers in the interval 80-100 substantially more often than women: 39.8% of all men’s wage offers versus 16.9% for women (see Figure 6).

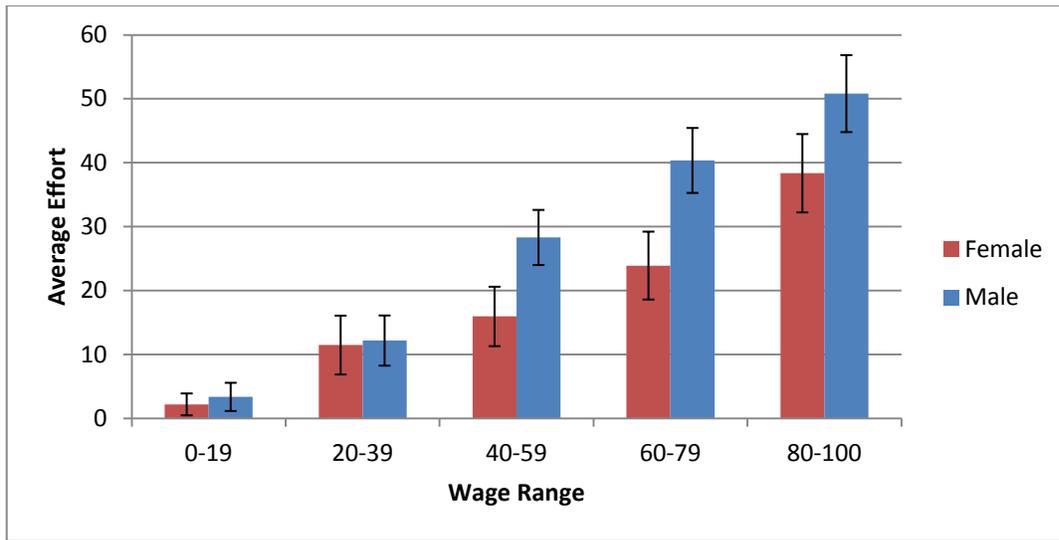


Figure 5. Effort level over wage intervals for men and women

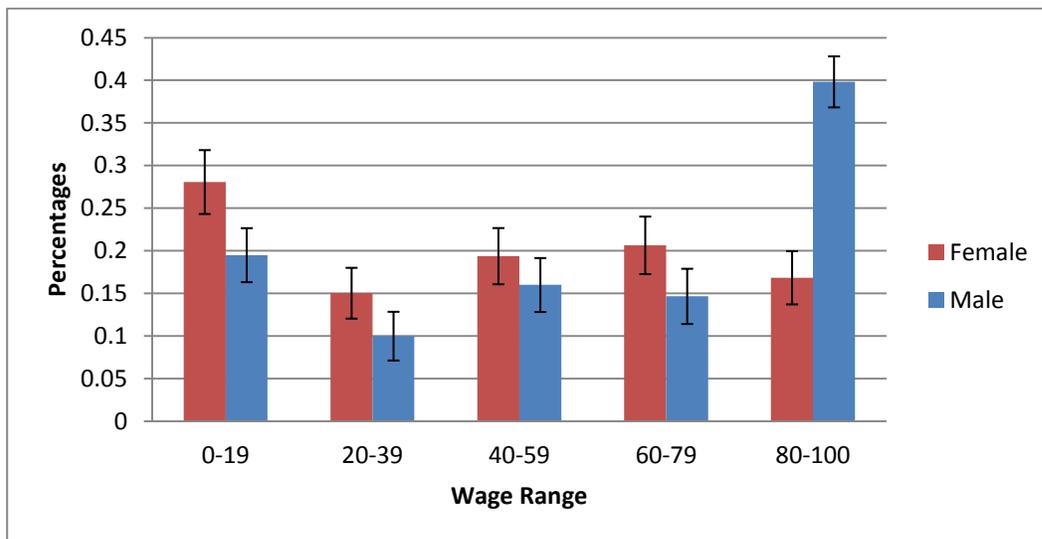


Figure 6. Distribution of wage offers for men and women.

Conclusion 3: Men tend to offer higher wages than women, and to respond with greater effort at higher wage rates as well.

Firm statistical support for Conclusion 3, along with the impact of the BFI and SAT scores on wage offers and effort responses, are provided in the regression analysis reported on below.

3.2 Statistical Analysis Including the Big Five and SAT: Wage Offers

Table 2 reports regression results for wage offers in relationship to subjects' SAT scores and the BFI. Since actual wage offers are bounded by zero from below and 100 from above, a random effect two-limit Tobit model is used for the statistical analysis. Specifically, we assume that the index function for offered wages for individual i in period p is

$$w_{ip}^* = \beta X_{ip} + \alpha_i + e_{ip}. \quad (1)$$

Further, we assume that observed wage offers are determined by

$$\begin{aligned}
 w_{ip} &= 0 & \text{if } w_{ip}^* < 0, \\
 w_{ip} &= 100 & \text{if } w_{ip}^* > 100, \\
 w_{ip} &= w_{ip}^* & \text{otherwise.}
 \end{aligned}
 \tag{2}$$

In (1) α_i is a random effect error term, which is *iid* across i and distributed as $N(0, \sigma_\alpha^2)$, while e_{ip} is an idiosyncratic error term, which is *iid* over i and p and distributed as $N(0, \sigma_e^2)$ (We also assume that α_i and e_{jp} are independent for all i, j and p .) We considered a version of (1) that contained session fixed effects, but this did not have any effect on the standard errors, and the session effects were not close to being significant at standard levels.¹⁷ All of the specifications included dummy variables for race – Caucasian, African-American, and Asian – with Other serving as the excluded category.¹⁸ In almost all cases the coefficients on these variables failed to be statistically significant and are suppressed due to space considerations. Those few cases where these variables proved to be statistically significant are discussed below in the text or in related footnotes. It is worth emphasizing that β_k represents the effect of increasing X_k on the index function w^* , not on the expected value of the observed wage $E(w) \equiv E(w | X)$. However, $\beta_k \geq 0$ implies that $\partial E(w) / \partial X_k \geq 0$ and vice-versa. Moreover, if β_k is statistically significant, then $\partial E(w) / \partial X_k$ will also be significant. Finally, the larger β_k is in absolute value, the larger $\partial E(w) / \partial X_k$ will be in absolute value as long as $E(w)$ is in the open interval $(0, 100)$.¹⁹ Hence in what follows we will use ‘the effect on the wage’ and the ‘effect on the offered wage index function’ interchangeably for qualitative results, but will use the latter when reporting quantitative results.²⁰

Columns 1 and 2 of Table 2 report estimates of the wage index function for the pooled data with and without the SAT variable respectively. Results absent the SAT scores are reported because past studies have often looked at the impact of the Big Five while not having information on cognitive ability, and thus will produce biased coefficients for (1) if any of the Big 5 variables has a nonzero partial correlation with SAT (see the online Appendix).²¹ Comparing the results with and without SAT indicates the magnitude of these biases, if any. The remaining columns repeat the analysis in columns (1) and (2) separately for men and women, since a likelihood ratio test for differences in coefficient values between the two rejects the null hypothesis that men

¹⁷ We use session fixed effects instead of session random effects since the number of sessions is small (Imbens and Wooldridge 2009).

¹⁸ Other consisted of sixteen subjects: 7 Hispanics and 9 of unknown ethnicity.

¹⁹ The difference between β_k and $\partial w / \partial X_k$ will be smaller here than in the standard Tobit model, since we have truncation at both ends.

²⁰ Labor economists refer to the terms $\partial E(w) / \partial X_k$ as ‘partial effects’ and often discuss these effects quantitatively. Here we focus on the Tobit coefficients because this approach follows the convention in experimental economics.

²¹ The online Appendix is available at www.econ.umd.edu/~filizozbay/FHKO_appendix.pdf. If SAT did not affect the wage index function there would be no bias, but from column (1) it is clear that SAT does indeed affect this index function.

and women have the same coefficients ($p < 0.01$) conditional on them having different intercepts.²² This rejection is not surprising given the differences between men and women reported in Figure 6.

Table 2
Random Effects Estimates of the Wage Index Function
(Standard Errors in Parentheses)

Variable	Pooled data (men and women)		Men	Women	Men	Women
SAT	1.697*** (0.291)		1.530*** (0.360)	1.667*** (0.464)		
Agreeableness	0.529** (0.206)	0.372 (0.239)	0.648** (0.280)	0.453* (0.255)	0.461 (0.324)	0.300 (0.284)
Openness	-0.052 (0.196)	0.058 (0.227)	-0.046 (0.296)	0.157 (0.223)	0.061 (0.347)	0.228 (0.251)
Neuroticism	0.084 (0.181)	0.002 (0.210)	-0.093 (0.265)	0.190 (0.206)	-0.115 (0.312)	0.085 (0.231)
Conscientiousness	0.191 (0.245)	0.315 (0.284)	1.147*** (0.355)	-1.042*** (0.339)	1.477*** (0.406)	-1.23*** (0.378)
Extroversion	-0.102 (0.194)	-0.563*** (0.205)	-0.032 (0.272)	0.343 (0.247)	-0.414 (0.298)	-0.012 (0.256)
Female	-15.23** (6.132)	-16.49** (7.149)				
Period	-0.0906 (0.240)	-0.0902 (0.240)	0.0035 (0.386)	-0.143 (0.295)	0.007 (0.387)	-0.140 (0.294)
Constant	-97.73** (38.05)	45.23 (33.78)	-142.7*** (50.53)	-91.40* (54.60)	-29.23 (49.96)	63.11* (38.03)
Log-Likelihood	-4406.7	-4421.4	-2167.0	-2214.5	-2174.7	-2220.2

*Statistically significant at the 10% level. ** Statistically significant at the 5% level. *** Statistically significant at the 1% level.

For the pooled data, SAT, agreeableness and the dummy variable for women are all statistically significant at conventional levels. Subjects with higher SAT scores offer higher wages, as do more agreeable types. The agreeableness outcome is not surprising in light of past research, particularly experiments showing that more agreeable types are more trusting and more reciprocal in the trust game (Becker et al., 2012). Regarding the strong positive relationship between SAT and wages, one explanation that immediately comes to

²² The likelihood ratio test statistic equals 50.4, which is substantially larger than the critical value for $\chi^2_{99\%}(10)$.

mind is that higher cognitive ability is associated with a greater tolerance for risk (Dohmen et al., 2010; Burks et al., 2009): As noted earlier, there is a persistent 20% chance, regardless of the wage level, of a zero effort response. But as Figure 3 shows, the expected return increases substantially at higher wages rates even after accounting these zero effort responses. So although it is riskier to provide a wage offer in the interval of 80-100 compared to a zero wage offer with its guaranteed return of 100, the expected value of a wage offer in this interval (235 ECUs) far exceeds the guaranteed 100 ECUs for a zero wage offer.²³ Of course, this is not the only possible explanation for the strong positive relationship between wages and SAT scores. Others include the possibility that subjects with greater cognitive ability will more readily recognize the possibility of mutual beneficial cooperation. There is some evidence for this effect: Jones (2012) finds that subjects with higher SAT scores are more likely to cooperate in a repeated play prisoner dilemma game with a random end point.²⁴ However, one argument against this explanation is that men with higher SAT scores tend to offer lower effort, other things equal (see below).²⁵

Dropping SAT from the Tobits for the pooled data eliminates the statistical significance of agreeableness. Further, the extroversion characteristic, which is negative and not close to being statistically significant with SAT included, stays negative, increases substantially in absolute value, and becomes significant at the 0.01 level with SAT excluded. Using the specification error results in the Appendix for intuition, this change in value for extroversion reflects the fact that extroversion has a statistically significant negative coefficient in a regression of SAT on all of the other independent variables.²⁶ Finally, omitting SAT has a substantial effect on the intercept, as it goes from statistically significant and negative to positive but insignificant.²⁷

²³ Obviously small, nonzero effort responses also make it risky to offer high wages. We focus on the zero responses for expositional ease here.

²⁴ But this result does not account for the impact of any of the Big Five on the behavior reported.

²⁵ One suggested explanation that did not check out is that higher SAT types would be more likely to experiment with different wage rates. But there is no evidence to this effect in the data.

²⁶ Its coefficient value is -0.142 ($p < 0.01$).

²⁷ Again, the results in the Appendix can be used for intuition here.

Table 3
The Effect on the Wage Index Function of a One Standard
Deviation Increase in Key Explanatory Variables
(Change as a Percent of the Mean Value of the Wage Index Function in Parentheses)

		Change in Mean Wage Index Function Value Resulting from a One Standard Deviation Increase		
	Mean Wage (w^*) Index Function Value	SAT	Agreeableness	Conscientiousness
Pooled Data	51.6	19.8 (38.4%)	8.0 (15.5%)	2.5 (4.8%)
Men	62.4	19.2 (30.8%)	9.9 (15.9%)	14.6 (23.4%)
Women	40.8	17.5 (42.9%)	6.6 (16.2%)	-13.4 (-32.8%)

To obtain a picture of the magnitude of the effects implied by the estimates in column (1) of Table 2, Table 3 presents the effect on the wage index function of a one standard deviation increase in SAT and agreeableness (These are the variables that are statistically significant in column (1); we also include conscientiousness since it becomes important once we separate men and women.) A one standard deviation increase in the SAT score increases the pooled wage index function by 19.8ECUs (38.4%), while a one standard deviation increase in agreeableness only increases the wage index function by 8.0 ECUs (15.5%).

Conclusion 4: For the pooled data, subjects with higher SAT scores and more agreeable types have significantly higher wage index functions, with a one standard deviation increase in SAT scores increasing this index function by a greater amount than a one standard deviation increase in agreeableness, with d regression estimates indicating that men have significantly higher wage offer index functions than women. Dropping SAT from the regression leads to agreeableness becoming insignificant and extroversion becoming statistically significant, indicating the importance of including a measure of cognitive ability when assessing the impact of personality characteristics on behavioral outcomes.

Looking at separate estimates of the wage index function for men and women with SAT included (columns 3 and 4 of Table 2), the truly interesting result here is that conscientiousness is now statistically significant for both genders, but *opposite* in sign: It is positive for men but negative, with a comparable absolute value, for women. Although this opposite reaction by gender to increased conscientiousness is unexpected, some immediate justification for it can be found in observed differences in the degree of *reciprocity* resulting from increased conscientiousness found in the estimated effort response index functions reported below. There, other things equal, greater conscientiousness in men results in a modest but positive increase in the effort

response (positive reciprocity), but in women it results in a modest *decrease* in the effort response. Thus, more conscientious men, thinking from their own perspective, would be likely to offer higher wages, other things equal, while women thinking from their perspective would not.²⁸ The SAT coefficient value is comparable between men and women, as is agreeableness. Finally, dropping SAT (see columns 5 and 6) from these separate gender specifications has larger effects on the agreeableness and extroversion coefficients for women than for men, which is consistent with the fact that these variables significantly affect SAT for women but not for men.

Rows (2) and (3) of Table 3 indicate the effect of a one standard deviation increase of key explanatory variables on the offered wage index function for men and women respectively. Here a change in conscientiousness is predicted to result in a 14.6 ECU *increase* in the wage index function for men versus a 13.4 ECU *decrease* for women (a 23.4% increase compared to a 32.8% decrease from the mean of the index function). For men, this increase is almost as large as the impact of a one standard deviation increase in SAT scores (19.2 ECUs) and is larger than the impact of a one standard deviation increase in agreeableness (9.9 ECUs).

Conclusion 5: Men and women differ substantially in their mean wage offer index functions, but show comparable effects in terms of a one standard deviation increase in SAT and agreeableness. However, the impact of conscientiousness is positive for men and negative for women, with both effects statistically significant and of comparable absolute value. This differential impact of conscientiousness is consistent with the differential impact of conscientiousness on the effort response index functions for men and women reported on below.

3.3 Statistical Analysis Including the Big Five and SAT: Effort Responses

Since actual effort levels also are bounded by zero from below and 100 from above, we again use a random effects Tobit model for our statistical analysis. We assume that the index function for effort responses for individual i in period p is given by

$$\begin{aligned} E_{ip}^* &= \delta_1 X_{ip} + \delta_2 w_{ip} + \delta_3 (w_{ip} * X_{ip}) + \gamma_i + u_{ip} \\ &= \delta Z + \gamma_i + u_{ip}. \end{aligned} \tag{3}$$

Further, observed effort response is given by

$$\begin{aligned} E_{ip} &= 0 \quad \text{if } E_{ip}^* < 0, \\ E_{ip} &= 100 \quad \text{if } E_{ip}^* > 100, \\ E_{ip} &= E_{ip}^* \quad \text{otherwise.} \end{aligned} \tag{4}$$

²⁸ A potential explanation for this differential effect is discussed in the conclusions section of the paper.

In (3) γ_i is a random effects error term, which is *iid* across i and distributed as $N(0, \sigma_\gamma^2)$, while u_{ip} is an idiosyncratic error term, which is *iid* (over i and p) and distributed as $N(0, \sigma_u^2)$ (We assume that γ_i and u_{jp} are independent for all i, j and p .) We considered a version of (3) that included session fixed effects but again this did not affect the standard errors, and the session effects were not close to being significant at standard levels. We allow for interaction terms between the explanatory variables and the offered wage, since the null hypothesis of no interactions was decisively rejected ($p < 0.01$ in all cases).²⁹

Table 4 reports the estimated effort response index functions in the same format as those for wage offer index functions. We restrict the analysis to the case where subjects face a positive wage offer, since zero wage offers are overwhelmingly met with zero effort. As such cognitive and non-cognitive characteristics essentially play no role in mediating responses to zero wage offers, so to include them would bias the estimates.³⁰

Column (1) of Table 4 presents the results for the pooled data when we include SAT.³¹ The coefficient on the wage by SAT interaction term is positive and significant at the 10% level. This coefficient, in conjunction with the negative sign for the SAT coefficient itself, implies that subjects with higher SAT scores have a lower effort responses at all wages (other things equal), with this negative effect diminishing at higher wages.³² The wage by agreeableness interaction term is positive and statistically significant at the 1% level, which, in conjunction with the positive coefficient for agreeableness, implies that more agreeable types offer higher effort at all wage rates. The coefficient for conscientiousness is positive and statistically significant, with the wage by conscientiousness negative and statistically significant; these coefficient values imply that more conscientious types provide higher effort throughout, with this positive effect diminishing at higher wages. Finally, the female by wage interaction effect is negative and statistically significant at the 1% level. This in conjunction with the small positive (but statistically insignificant gender dummy), implies no significant differences in effort levels at lower wages, with women providing significantly less effort than men at higher wages.³³ Column (2) shows that the principle impact of dropping the SAT variable is a dramatic increase in the magnitude and statistical significance of the coefficient on the wage variable by itself.

²⁹ The relationship between the index function E^* and the expected value of effort is analogous to that between the index function w^* and the expected value of the wage.

³⁰ The bias arises from mixing the effort index function (3) with the very different effort response index function that applies to a zero offered wage. Note that the offered wage is exogenous to the responders, hence omitting the responses for zero wage offers does not create any selection bias.

³¹ For the pooled data, there are no significant ethnicity effects for any of the ethnicity dummies or for the wage by ethnicity interaction effects.

³² $\partial E_{ip}^* / \partial SAT_i = -.601 + .0059w_{ip}$, and only becomes positive for $w_{ip} > 101.9$, which is outside the [0,100] interval. A statistical test for SAT and SAT*wage together is not statistically significant at the 10% level or better. As will be shown below this results from the fact that SAT and SAT*wage is only significant at conventional levels for men, but not for women, so that in the pooled data the combined effect is masked.

³³ The difference in the effort index function between a woman and a man who have identical characteristics is $5.961 - .227w$, which is negative for $w > 26.26$.

Table 4
Random Effects Estimates of the Effort Response Index Function
(Standard Errors in Parentheses)

Variable	Pooled data (men and women)		Men	Women	Men	Women
	Wage	0.240 (0.380)	0.652** (0.306)	-0.042 (0.621)	0.277 (0.579)	0.665 (0.557)
SAT	-0.601 (0.382)		-1.525*** (0.550)	0.097 (0.527)		
Agreeableness (Agr)	0.058 (0.274)	0.065 (0.274)	0.565 (0.408)	0.314 (0.367)	0.327 (0.402)	0.289 (0.344)
Openness (Ope)	0.355 (0.295)	0.299 (0.293)	0.416 (0.463)	-0.013 (0.323)	0.291 (0.465)	0.001 (0.316)
Neuroticism (Neu)	0.107 (0.221)	0.076 (0.220)	0.319 (0.324)	-0.146 (0.269)	0.242 (0.326)	-0.140 (0.269)
Conscientiousness (Con)	0.641** (0.296)	0.631** (0.296)	0.770* (0.403)	0.230 (0.378)	0.769* (0.406)	0.234 (0.377)
Extroversion (Ext)	-0.004 (0.225)	0.016 (0.235)	-0.294 (0.379)	0.476* (0.266)	-0.287 (0.382)	0.466* (0.260)
Female	5.961 (8.491)	7.765 (8.44)				
Period	-0.405 (0.254)	-0.422* (0.253)	-0.516 (0.389)	-0.157 (0.303)	-0.558 (0.390)	-0.153 (0.303)
Wage*SAT	0.0059* (0.0032)		0.0134*** (0.0050)	-0.0145 (0.0046)		
Wage*Agr	0.0122*** (0.0023)	0.0121*** (0.0023)	0.0087** (0.0036)	0.0139*** (0.0033)	0.0109*** (0.0035)	0.0143*** (0.0030)
Wage*Ope	-0.0035 (0.0024)	-0.0029 (0.0024)	-0.0002 (0.0042)	-0.0051* (0.0028)	0.0012 (0.0041)	-0.0054* (0.0028)
Wage*Neu	0.0021 (0.0018)	0.0025 (0.0018)	0.0012 (0.0030)	0.0060** (0.0024)	0.0018 (0.0029)	0.0059** (0.0024)
Wage*Con	-0.0062*** (0.0024)	-0.0061** (0.0024)	-0.0092** (0.0036)	-0.0053* (0.0031)	-0.0088** (0.0035)	-0.0053* (0.0031)
Wage*Ext	-0.0013 (0.0018)	-0.0015 (0.0018)	0.0020 (0.0033)	-0.0028 (0.0023)	0.0016 (0.0033)	-0.0026 (0.0023)
Wage*Female	-0.227*** (0.070)	-0.247*** (0.068)				
Constant	-55.03 (46.18)	-97.31*** (37.24)	-59.27 (65.76)	-64.36 (66.12)	-139.1** (59.60)	-56.22 (47.64)
Log Likelihood	-3303.7	-3305.7	-1821.3	-1443.1	-1826.0	-1443.2

*Statistically significant at the 10% level. ** Statistically significant at the 5% level. *** Statistically significant at the 1% level.

Row (1) of Table 5 reports the quantitative effects of a one standard deviation increase in the key explanatory variables on the effort index function. For the pooled data, wage plays the dominant role in determining effort levels, with a one standard deviation increase raising the effort index function by 22.4 ECUs (over 100%). A one standard deviation increase in agreeableness increases the effort index function by 12.3 ECUs (65.1%).

Table 5
Change in the Effort Index Function Resulting from One Standard Deviation
Increases in the Key Explanatory Variables
(Change as a Percent of the Mean Value of the Effort Index Function in Parentheses)

		Change in Mean Effort Index Function Value as a Result of a One Standard Deviation Increase			
	Mean Effort (E^*) Index Function Value	Wage	SAT	Agreeableness	Conscientiousness
Pooled Data	18.9	22.4 (118.5%)	-3.0 (-15.9%)	12.3 (65.1%)	4.2 (22.5%)
Men	19.8	26.8 (135.4%)	-8.9 (-44.9%)	17.3 (87.4%)	3.8 (19.2%)
Women	16.9	16.9 (100.0%)	0.1 (0.6%)	18.0 (106.5%)	-1.1 (-6.5%)

Conclusion 6: The effort response Tobits for the pooled data show that higher SAT scores result in modestly lower effort, while more conscientious and agreeable types provide higher effort. Of these three factors, a one standard deviation increase in agreeableness has by far the greatest impact (a 65% increase) on the effort response function, with a one standard deviation increase in wages having the largest impact. Men have higher effort levels over most of the wage interval, with this difference increasing at higher wages.

The remaining columns in Table 4 report separate estimates of the effort response index function for men and women, since a likelihood ratio test decisively rejects the null hypothesis that men and women have the same coefficients except for the intercept.³⁴ Columns (3) and (4) refer to our preferred specification where SAT is included. Note that a statistically significant impact for SAT is confined exclusively to men, where the coefficient on the SAT variable by itself is negative and statistically significant at the 1% level, with the wage by SAT variable coefficient positive and significant ($p < 0.01$) as well. The estimates indicate that higher SAT scores reduce effort for men throughout the range of possible wages, with its major impact confined to lower wage offers; but the net effect of a one standard deviation increase in wages still results in a substantial increase

³⁴For example, when we include SAT the test statistic is $78.6 > 37.57 = \chi_{99\%}^2(20)$.

in the average effort response of higher scoring SAT men.³⁵ In contrast, for women, neither of the two SAT variables is individually significant, nor are they jointly significant at conventional test levels.³⁶ The net effect is a substantially larger impact of a one standard deviation increase in SAT on the male effort index function, *reducing* it by 44.9%, whereas for women the impact is negligible – compare rows (2) and (3) of Table 5. One possible interpretation of the negative male SAT effect is that men with greater cognitive ability are more sensitive to the one-shot nature of the interactions inherent in the structure of the game. However, it is not clear why women would be less sensitive in this dimension than men given that they have comparable SAT scores.³⁷

The coefficient on the conscientiousness variable by itself is positive for both men and women, but for men the coefficient value is considerably larger and statistically significant. The coefficient on the conscientiousness by wage variable is significantly negative for both men and women, but again the effect for men is bigger (in absolute value). Again, considering rows (2) and (3) of Table 5, these coefficient estimates imply that for men increases in conscientiousness leads to a higher effort response index function, particularly at lower wages, while for women increased conscientiousness has no significant impact on effort at lower wages, and results in lower effort levels at higher wage rates. As noted earlier, this differential effort response by men and women with respect to conscientiousness is internally consistent with the effect this variable has on male and female wage offers. That is, both genders seem accurately to predict own gender effort response with respect to conscientiousness and to act accordingly in setting wages, although the nature of the effort response differs between the genders. Note there is independent evidence for this sort of effect in the psychology literature.³⁸ In terms of the remaining Big Five coefficients, only the women’s (positive) wage by neuroticism interaction and (positive) extroversion coefficients are statistically significant.

Rows (2) and (3) of Table 5 indicate that the effect of a one standard deviation increase in wages on the effort response function is substantial for both men and women, although it is substantially larger for men (26.8 ECUs vs. 16.9 ECUs); an increase of 135.4% for men versus 100.0% for women.³⁹ The large positive impact of increased wages on effort for men, in conjunction with the negative effect of SAT on effort that diminishes at higher wages, suggests mixed motives at work for higher SAT type men: Like everyone, they tend to be more reciprocal at higher wages, but they are somewhat less responsive than lower SAT men or than women, perhaps

³⁵ For men, $\partial E_{ip}^*/\partial SAT_i = -1.525 + .0134w_{ip}$ and becomes positive only for $w > 113.8$, which is outside the [0,100] interval.

³⁶ The test statistic is 0.10 with a p-value of 0.951.

³⁷ One conjecture that we investigated is that perhaps there are more men drawn from economics and business majors than women, with business and economics majors having more exposure to repeated versus one-shot games. Although we do not have major field of study for the entire sample, the data we do have shows no difference on this score. (We have data for 47.6% of the men in the sample and for 48.3% of the women.) However, there is some evidence that it takes women longer to learn to act strategically in some game theoretic settings than men (see Casari, Ham and Kagel, 2007 and Cooper and Kagel, 2012b).

³⁸ In the psychology literature, this is referred to as “consensus bias”: the overuse of self-related knowledge in estimating the prevalence of attributes in a population (Ross, Green and House, 1977; Kruger and Clement, 1994).

³⁹ We use the mean of the other characteristics for men when calculating the wage effects for men, and do the analogous procedure for women. However, since the mean values in Table 1 are close for men and women, it does not make any real difference if we use the pooled means for both men and women when calculating wage effects.

on account of the one-shot nature of the interactions in the game. Finally, a one standard deviation increase in agreeableness raises the effort index function slightly less for men than for women (17.3 ECUs vs. 18.0 ECUs), Comparing the impact of a one standard deviation in agreeableness to a one standard deviation increase in wages, agreeableness has a slightly greater impact for women than a comparable increase in wages. In contrast, for men, the impact of increased agreeableness although impressive, has a smaller impact than a comparable increase in wages. It is reasonably impressive (and not completely anticipated) that one of the personality characteristics – agreeableness – can have such a large quantitative impact on the effort response index function compared to the core economic variable wage paid.

Dropping the SAT variable from the separate specifications impacts the wage coefficient (by itself) for both men and women, but the standard errors are large, so these changes may just reflect sampling error. For men, dropping SAT results in a sharp reduction in the constant, but again the standard error is large. A second order effect of dropping SAT is found in the race by wage interaction effects (not reported in the table but included in the specifications). With SAT in, the only statistically significant race effect is that the dummy variable for Asian men is positive and significant at the 10% level. With SAT out, the wage by African-American interaction effect is negative and statistically significant at the 5% level for men. (With SAT in, the African-American dummy variable is not significant at conventional levels.) These ethnicity effects suggest that it will be important to control for SAT in any study examining ethnic (as opposed to gender) differences in experiments.

Conclusion 7: There are major differences in effort responses of men and women: Men with higher SAT scores have reduced effort levels, particularly at lower wages, with SAT scores having minimal impact on the effort supply of women. More conscientious men supply greater effort, particularly at lower wages, while conscientiousness has no significant impact on effort levels for women at lower wages, and results in less effort at higher wages. A one standard deviation increase in average wages increases the effort response index function for men by 135% compared to 100% for women, with a one standard deviation in agreeableness having substantial effects for both men and women as well. Dropping SAT does not affect the Big Five coefficient estimates but has a substantial impact on the ethnicity coefficients.

Summary and Conclusions

We report results from a one-shot gift exchange experiments accounting for the effects of gender, cognitive ability, and the Big Five social characteristics on outcomes. We find substantial impacts on behavior for each of these typically neglected factors.⁴⁰ On average, women offer lower wages than men do when they are evaluated at the same level of the explanatory variables, with women offering less effort than men in

⁴⁰ In turn these hidden personality and cognitive ability factors may help account for the high degree of individual subject variability in effort responses and wage offers reported in the typical gift exchange experiment.

response to the same wage offers. These results add to the growing literature on gender effects in economics and in particular on gender effects in economic experiments (see Croson and Gnezy, 2009 for a survey).

The major impact of cognitive ability on outcomes is that both men and women with higher SAT scores offer higher wages than their counterparts with lower SAT scores. We conjecture that the primary factor behind this outcome has to do with the growing evidence that higher cognitive ability is associated with less risk aversion (Dohmen et al., 2010; Burks et al., 2009): Although higher wages have substantially higher expected returns, particularly at the highest wage rates, they are much riskier than lower wage offers due to a consistent cluster of zero effort responses, regardless of the size of the wage offer. Dropping SAT from the Tobit regressions for wages has several effects: For the pooled data, the coefficient value for agreeableness goes from being positive and statistically significant at better than the 5% level, with SAT in, to no longer being statistically significant. Further extroversion becomes statistically significant at the 1% level with SAT out, indicating the importance of having a measure of cognitive ability when investigating the impact of personality characteristics on economic behavior. Further, dropping SAT from the effort equation impacts the size and significance of ethnicity effects, indicating the potential importance of including some measure of cognitive ability when investigating ethnicity differences in effort responses.

At times, the Big Five personality characteristics have as large an impact on the wage offer and effort response index functions as cognitive ability and economic variables (wages). As in most experiments of this sort, higher wage offers are met with a higher effort response. A one standard deviation increase in agreeableness has, for women, the same impact on the effort response index function as does a one standard deviation increase in wages, while having an impact on male effort just under two-thirds the impact of a comparable wage increase. On the wage side, for men, the impact of a one standard deviation increase in conscientiousness increases the wage offer index function by about the same amount as a one standard deviation increase in SAT scores. The surprising issue here is that, for women, conscientiousness has the *opposite* impact on the wage offer index function (of roughly the same absolute value) as it does for men. This differential effect of conscientiousness on wages is consistent with best responding to its effect on effort where, at low wages, increased conscientiousness leads to increases in the effort response index for men but essentially the same or a modest negative effect for women.⁴¹ One possible explanation for this differential effect of conscientiousness is as follows: One element of the conceptual definition of conscientiousness is “following norms and rules” (John et al., 2008, Table 4.2). With this in mind, note that there is some evidence suggesting that, for women, explicit monetary payments tend to drive out social preferences more than for men.⁴² In this

⁴¹ Note that subjects do not know the gender of the person they are coupled with in any given round of play.

⁴² Mellström and Johannesson (2008) found that paying people to donate blood reduced women’s donations, while men’s donations were unaffected.

case, more conscientious women would be more likely to have lower responsiveness to wage offers, with women wage givers best responding to these beliefs. In contrast, if men are less sensitive, or immune, to this crowding out effect, and more accepting of the notion of explicit monetary benefits for reciprocal responses, more conscientious men would be more likely to take account of this fact and offer higher wages.

The results of this experiment have obvious and immediate implications for the social preference literature in economics. We believe that they have implications for the labor economics literature as well. On this last point, one of the interesting questions will be to extend the analysis of the role of the Big Five personality characteristics, gender, and cognitive ability to gift exchange games in which agents can develop reputations through repeated or longer term contracts. In this case, we would expect a very different effect of cognitive ability on effort responses for both men and women, with agents with greater cognitive ability providing greater effort in response to higher wages, motivated by the potential for cooperation inherent in repeated interactions.

References

- Almlund, Mathilde, Angela Lee Duckworth, James J. Heckman, and Tim Kautz, 2011. 'Personality Psychology and Economics' in *Handbook of the Economics of Education Vol. 4*, E. A. Hanushek, S. Machin, and L. Woessman (Eds.), North Holland:Elsevier: 1-181.
- Anderson, Jon, Stephen Burks, Colin De Young, and Aldo Rustichini, 2011. 'Toward the Integration of Personality Theory and Decision Theory in Explanation of Economic Behavior,' *University of Minnesota working paper*.
- Becker, Anke, Thomas Deckers, Thomas Dohmen, Armin Falk, and Fabian Kosse, 2012. 'The Relationship Between Economic Preferences and Psychological Personality Measures,' working paper, University of Bonn.
- Borghans, L., A.L. Duckworth, J.J. Heckman, and B.T. Weel, 2008. 'The Economics and Psychology of Personality Traits,' *Journal of Human Resources*, Vol. 43, pp. 972-1059.
- Burks, S.V., J. Carpenter, L. Götte, A. Rustichini, 2009. 'Cognitive Skills Affect Economic Preferences, Social Awareness, and Job Attachment,' *Proceedings of the National Academy of Science*, Vol. 106, pp. 7745-7750.
- Casari, Marco, Ham, John, and Kagel, John H., 2007. "Selection Bias, Demographic Effects and Ability Effects in Common Value Auction Experiments," *American Economic Review*, vol. 97 (4), pp. 1278-1304.
- Cohen, Patricia., Jacob Cohen, Leona S. Aiken and Stephen G. West, 1999. 'The Problem of Units and the Circumstance for POMP,' *Multivariate Behavioral Research*, Vol. 34, pp. 315-346.
- Cooper, David J. and John H. Kagel, 2012a. "Other Regarding Preferences: A Selective Survey of Experimental Results," to appear in *The Handbook of Experimental Economics*, vol. 2, John H. Kagel and Alvin E. Roth (Eds), Princeton University Press.
- _____ and _____ 2012b. "A Failure to Communicate: An Experimental Investigation of the Effects of Advice on Strategic Play," Ohio State University working paper.
- Croson, Rachel and Uri Gneezy, 2009. 'Gender Differences in Preferences,' *Journal of Economic Literature*, Vol. 47, pp. 448-474.
- Dohmen, Thomas and Armin Falk, 2011. 'Performance Pay and Multidimensional Sorting: Productivity, Preferences, and Gender,' *American Economic Review*, Vol. 101, pp. 556-590.
- Dohmen, Thomas, Armin Falk, David Huffman, and Uwe Sunde, 2010. 'Are Risk Aversion and Impatience Related to Cognitive Ability?' *American Economic Review*, Vol. 100, pp. 1238-1260.
- Fishbacker, Urs, 2007. "z-Tree: Zurich Toolbox for Ready Made Economic Experiments," *Experimental*

- Economics*, Vol. 10, pp. 171-178.
- Frey, Meredith C. and Douglas K. Detterman, 2004. 'Scholastic Assessment or g?: The Relationship Between the Scholastic Assessment Test and General Cognition Ability,' *Psychological Science*, Vol. 15 (6), pp 373- 378.
- Gächter, Simon and Ernst Fehr, 2001. "Fairness in the Labour Market –A Survey of Experimental Results" <http://ideas.repec.org/p/zur/iewwpx/114.html>
- Heckman, James J., Jora Stixrud, and Sergio Urzua, 2006. 'The Effects of Cognitive and Noncognitive Abilities on Labor Market Outcomes and Social Behavior,' *Journal of Labor Economics*, Vol. 24 (3), pp. 411 - 482.
- Imbens, Guido and Jeffrey Wooldridge 2009. *Lecture Notes for Applied Microeconometrics Workshop*, Institute for Research on Poverty, University of Wisconsin, Madison, August 2009. Available at <http://www.irp.wisc.edu/newsevents/workshops/appliedmicroeconometrics/schedule1.htm> .
- John, Oliver P., Naumann, Laura, P., and Christopher J. Soto 2008. "Paradigm Shift: to the Integrative Big Five In Oliver P. John, R. W. Robbins, and L. A. Pervin (Eds.), *Handbook of Personality, Theory, and Research*, N. Y. Guilford Press, pp. 114-158.
- Jones, Matthew T. 2012. "Strategic Complexity and Cooperation: An Experimental Study," SSRN working Paper. http://papers.ssrn.com/sol3/papers.cfm?abstract_id=2025659
- Kruger, Joachim and Russell W. Clement 1994. "The Truly False Consensus Effect: An Ineradicable and Egocentric Bias in Social Perception,," *Journal of Personality and Social Psychology*, Vol. 67 (4), pp. 596-610
- Kurzban, Robert and Daniel Houser, 2001. 'Individual Differences in Cooperation in a Circular Public Goods Game,' *European Journal of Personality*, Vol. 15, pp. 37-52.
- Mellström, Carl and Magnus Johannesson, 2008. 'Crowding Out in Blood Donations: Was Titmuss Right?' *Journal of the European Economic Association*, Vol. 6, pp.845-863.
- Pothos, Emmanuel M., Gavin Perry, Philip J. Corr, Mervin R. Matthew, and Jerome R. Busemeyer, 2010. 'Understanding Cooperation in the Prisoner's Dilemma Game,' *Personality and Individual Differences*, Vol. 51, pp. 210-216.
- Ross, Lee, Greene, David, and Pamela House. 1977. "The "False Consensus Effect": An Egocentric Bias in Social Perception," *Journal of Experimental Social Psychology*, Vol. 13, pp. 279-301.