Noname manuscript No.
(will be inserted by the editor)

To confess or not to confess? That is the question

Andrés Ramírez Hassan · Alejandro Puerta

Received: date / Accepted: date

1 Introduction

Being dishonest pays, and often pays very well. It is almost impossible to watch news or read newspapers without scandals about dishonest acts. Unfortunately, the academic environment is no an exception to this issue. For instance, Bowers (1964) found that 75% of the students engage in one or more incidents of academic dishonesty. Later, McCabe & Trevino (1997) replicated this survey, and found that this figure ranges between 13% and 95% in American college students. McCabe et al. (2006) illustrate the concern about this issue showing that some business schools are taking into account ethics in their curricula, and giving weight to ethical students' orientation in admission processes. Obviously, dishonesty generates negative consequences in society. But, what can we do about it?

The standard economic theory establishes that dishonest acts are the outcome of rational decisions that trade off external profits and costs (Becker, 1968). In particular, individuals consider the payoff of dishonest acts, against the probability of being caught and the magnitude of punishment if caught.

Andrés Ramírez Hassan
Department of Economics, Universidad EAFIT. Carrera 49 # 7 Sur-50, Medellín, Colombia.
Tel.: +57 4 2619549
E-mail: aramir21@eafit.edu.co

Alejandro Puerta
Department of Economics, Universidad EAFIT. Carrera 49 # 7 Sur-50, Medellín, Colombia.
Tel.: +57 4 2619500
E-mail: apuerta5@eafit.edu.co

However, social experiments have shown that individuals also consider fairness and reciprocity in society (Henrich et al., 1968). So, they have social reference points to compare their behavior. Alignment with these standards generates internal reward, whereas misalignment generates internal punishment through individuals' self–concept (Mazar et al., 2008). To conciliate these two approaches, external and internal motivations, Mazar et al. (2008) argue that individuals find a balance between them through two mechanisms: categorization and attention to moral standards. In the former, individuals categorize their actions into social standards. However, individuals trace their own limits and mold dishonest acts to enbrace them into social standards, such that they do not consider themselves dishonest. So, there is a displacement of own responsibility (Bandura, 2002). In the latter, when individuals are mindful about moral standards, any dishonest action affects negatively their self–concept. Therefore, when moral standards are more accessible, individuals confront themselves to check the moral integrity of their actions (Mazar et al., 2008).

The main objective in this paper is to estimate the short term treatment effects of a series of interventions in the academic environment whose aim is to do moral standards more accessible to students. We establish a setting where getting positive results regarding honest actions (*confessing*) is hard. In particular, we propose an experiment where students should confess that their professor made a mistake that favors their grade in a design where the external costs are null. Therefore, the relative weight of the external reward is very high. In addition, our design promotes a high level of displacement of own responsibility, such that the probability that individuals judge themselves as dishonest is low when they are actually dishonest (*Not confessing*).

The novelty of our paper in the (dis)honesty literature is the balance between experiment design and econometric methodology. First, we design a "structured observation" experiment (Shaughnessy & Zechmeister, 1985) avoiding self–selection, participants' behavioral modifications due to being watched and self–reporting. These characteristics imply better treatment effects estimates. These experimental setting has been used in psychology (Hyman et al., 2010; Valentino et al., 2011; Piaget, 2013), but, to the best of our knowledge, it is not common in economics. Second, we control for socioeconomic, as well as, psychological variables. Therefore, we mitigate the effect of confounding variables. Third, at methodological level, we use hierarchic longitudinal random effects logit models. This approach allows taking into account individual unobserved heterogeneity in unbalanced panel settings, and calculating treatment effects for specific individuals. Fourth, we follow a “non–informative” Bayesian approach for estimation. As a consequence we can easily perform statistical inference in this setting, and obtain the posterior distributions of functions of parameter under an “objective” approach. In particular, we perform statistical inference of marginal effects in non–linear models. Finally, we take into account model uncertainty due to regressors. We account for 29 regressors, which implies 2^{24} (16,777,216) possible models. We perform this task using an evolutionary algorithm based on Deviance Information Criteria (D. J. Spiegelhalter et al., 2002)

The results in our experiment are hopeful despite that we establish a very tough environment for being honest. In particular, we found that the treatment effect, measured through the semi–elasticity of the odds ratio (*confess vs No confess*), is on average 1.83. The 95% credible interval is equal to (1.37,2.39). In addition, we found for the representative student that participating in the treated group increases the probability of confessing the professor’s mistake on average 26.4%. The 95% credible interval is (2.87%,41.5%). This is evidence that, at least in a short term in the academic environment, moral standards availability pushes internal reward up, and as a consequence, treated individuals had a higher probability to act in an honest way. This is a encouraging outcome in a society that requires more honest individuals.

The remainder of this paper proceeds as follows. Section 2 describes the experiment and the integrity intervention. Section 3 exhibits the econometric methodology. Our main results are presented in section 4. Section 5 concludes.

2 Design: experiment and interventions

Our point of departure is social cognitive theory, which conceives individuals as agentic operators having the power to influence their own actions (Bandura, 1999), but led by the interplay between behavioral patterns (affective), environmental events (biological), and internal personal characteristics (cognitive) (Bandura, 1978). Therefore, the moral reasoning is influenced by self–regulatory mechanisms, which conciliate internal and external rewards. Given this framework, we design a set of interventions, whose main objective is to do moral standards regarding academic integrity more available. Therefore, we try to enhance the internal reward associated with *confessing*, hoping that outperforms the external reward of *not confessing*.

Table 1 shows the list of interventions that we apply. These interventions are based on previous psychological experiments showing positive effects on students' integrity, and comprehend: lectures about academic integrity, discussions, movie fragments, ethical dilemmas, among others.

The setup of our experiment involves 36 undergraduate students of econometrics that were equally split in two groups, treatment and control. Both groups have the same econometrics professor, who lectures them two classes (theory and applications) every week for 4 months. These groups are joined in the weekly theory lecture, and split for the applied class. In the latter, the treatment group, which was randomly selected, is weekly exposed to the interventions during three months (12 in total). Each intervention last 15 minutes, and was done at the end of the class, which last in total 90 minutes including the intervention. The net time for the applied class was 75 minutes in both groups.

The strategy of the identification of the causal effect of the interventions on the students was based on the grading system of the course. In particular, 60% of total grade is based on 12 quizzes that were done each week during three months. Each quiz had 10 multiple choice questions, each correct answer represents 0.5 points, such that a perfect score is equal to 5.0. Each quiz had the same weight, that is 5%. All students were evaluated at the same time with the same quizzes, then the control and treatment groups faced exactly the same conditions. To identify the causal effect, the professor intentionally made mistakes grading the quizzes, but students did not know it. Specifically, he randomly selected 4 students from each group, and gave them 0.5 extra points (one right answer). To reinforce the fact that there was a professor's mistake, and not a kind of bonus, all the solutions were clarified at the moment of delivering the graded quizzes. In addition, the right number of correct answers was written next to the wrong grade (see Figure 1 in Appendix). Our design tries to minimize the situation where students did not notice professor's mistake. However, if this issue is present, it makes sense to assume that is randomly distributed between all students because the control and treatment groups were exposed exactly to the same design. Observe that our "structured observation" experiment (Shaughnessy & Zechmeister, 1985),

Table 1: Interventions

Title	Elicitation	Reference
Lecture: ethics and academic integrity	Reminding academic integrity	(Ayal et al., 2015)
Fragment of the movie Inside Job	Ethical dissonance	(Barkan et al., 2015)
Lecture: Integrity in professional environment	Consequences of dishonest actions	(Bandura et al., 1996)
– Ask explicitly if the grade is OK – Asking about satisfaction in class	– Obligation to lie – Needs satisfaction	– (Gneezy, 2005) – (Ryan & Deci, 2000)
– Institutional campaign: "Atreverse a Pensar" – Signing a honesty commitment	– Euphemistic language – Honor code	– (Bandura et al., 1996) – (Shu et al., 2012)
– Gift: bracelet saying "I dare to think" – Public recognition to a student who *confessed* professor's mistake in a specific class	– Reminding academic integrity – Transforming internal reward into external reward (probable gain vs sure gain)	– (Mazar et al., 2008) – (Ayal et al., 2015)
Brain games episode: "Moral dilemma"	Reminding integrity	(Thaler, 1980)
Professors delivers quizzes in his office, and obtains feedback from students about the class	Needs satisfaction	(Ryan & Deci, 2000)
Lecture: *Bancolombia* case	Ethical dissonance	(Barkan et al., 2015)
Brain games episode: "Moral dilemma"	Reminding integrity	(Thaler, 1980)
Real testimony concerning cheating behaviors	Consequences of dishonest actions	(Bandura et al., 1996)
Classmates's stories about dishonest acts	Licensing and compensation	(Jordan et al., 2011)

which differs from naturalistic observation in the sense that researchers intervene to exert some control over the events they are observing, prevents against self–selection, behavioral modification of individual due to being watched, and self–reporting. These are desirable characteristics which help to identify in a better way causal effects due to moral integrity interventions.

The professor delivered the graded quizzes in our experiment. Therefore, it is supposed by the students that he did not have any proves to corroborate his mistakes. So, the probability that students are caught is zero, and as consequence, there is not external punishment. In addition, the experiment is based on professor's mistakes. This implies a high probability that students displace their responsibility due to *Not confessing* being not considered fraud (Bandura, 2002). However, it is a dishonest action because the grade does not correspond to what students deserve. This is a very tough environment to be honest, which in turn is based entirely on internal reward.

Observe that in our design students did not know that there were participating in an experiment. This situation may raise ethics issues. However, in experimental economics, explicit deception, that is lying, is usually banned, but, implicit deception, being "economical whit the truth", is not (McDaniel & Starmer, 1998; Hersch, 2015). Two main issues arise when deceiving: methodological (Bonetti, 1998), because it could make data to be invalid, and ethical, in the sense that we could harm others. Regarding the latter, Kelman (1967) suggests that the primary way of counteracting negative effects is post–experimental feedback. So, we informed students at the end of the experiment, and asked for their approval. They liked the experiment, and its main objective. Then, they agreed to participate in a focus group to obtain qualitative insights, and that we use the outcomes for designing a moral integrity campaign, and publishing the experiment outcomes. In any case, we always keep students' anonymity. In addition, our experiment design always benefits students with a higher grade. There were apparently not negative consequences for them. Regarding the methodological issues, we kept high confidentiality standards, and students told us in the posterior focus group that they did not suspect about our academic exercise.

3 Methodology

In order to obtain conditional causal effects, we performed a survey that included psychological, socioeconomic, educational and living habits variables.

Regarding the psychological variables, we set a moral dilemma where students had to respond how much they agreed with several assertions. Then, we calculate the C–index, which measures the degree to which individuals let their judgment behavior be determined by moral concerns or principles, rather than by other psychological forces (Lind, 1999). Calculation is based on the Moral Judgment Test, which is an instrument to measure individual's moral–judgment competence besides assessing their moral attitudes. This means to measure the ability of individuals to judge arguments in controversial moral problems on the basis of their own moral principles (Lind, 2008).

The construction of the test has two parts (j), each one having six stages (l) corresponding to each moral stage according to the Kohlberg's hierarchy (REFERENCE). The first part comprehends the judgments supporting an action, whereas the second the judgments criticizing it. The C-Index is calculated as follows

$$\frac{SS_{stage}}{SS_{dev}} \times 100$$

where $SS_{stage} = \frac{1}{2}\sum_{l=1}^{6}(\sum_{j=1}^{2} X_{lj})^2 - \left(\frac{1}{2}\frac{1}{6}\sum_{l=1}^{6}\sum_{j=1}^{2} X_{lj}\right)^2$, and $SS_{dev} = \sum_{l=1}^{6}\sum_{j=1}^{2} X_{lj}^2 - \left(\frac{1}{2}\frac{1}{6}\sum_{l=1}^{6}\sum_{j=1}^{2} X_{lj}\right)^2$. This index ranges between 0 and 1, and higher values imply that individual's judgment is based on moral concerns rather than other psychological issues.

In addition, we control for a cognition measure related to students' motivation using Likert scale associated with each set of questions. In particular, we asked students questions regarding personal intrinsic and extrinsic motivation, and self-efficacy (personal mastery orientation, and science deep–level strategies). Anderman et al. (1998) indicate that cheating behaviors and beliefs in science are associated with motivational orientations, and Finn & Frone (2004) conclude that students who were performing well were less likely to cheat when they had high self-efficacy, but were more likely to cheat when they had low self-efficacy, suggesting that they had little confidence in their ability to maintain high grades.

Regarding the socioeconomic, educational and living habits variables, we control for variables that have been used in previous researches, such as household income Bowers (1964), parents' education (Bowers, 1964; McCabe & Trevino, 1997; Grimes & Rezek, 2005; Khodaie et al., 2011), age (Haines et al., 1986; Lipson & McGavern, 1993; McCabe & Trevino, 1997; Anderman & Midgley, 2004; Finn & Frone, 2004; Grimes & Rezek, 2005; Kanat-Maymon et al., 2015), gender (Bowers, 1964; Lipson & McGavern, 1993; McCabe & Trevino, 1997; Finn & Frone, 2004), religion (Grimes & Rezek, 2005), GPA (Bowers, 1964; Haines et al., 1986; Lipson & McGavern, 1993; McCabe & Trevino, 1997; Finn & Frone, 2004; Grimes & Rezek, 2005), and membership to college representative groups Bowers (1964); Haines et al. (1986); McCabe & Trevino (1997). In addition, we control for other variables that, to the best of our knowledge, have not been used previously in this literature, such as race, high school type (public or private), having scholarship, class attendance, out of class study hours (econometrics and other subjects), and living habits (smokers and alcohol drinkers).

Our main aim is to identify the conditional causal effect of moral integrity interventions, whose aim is to do moral standards more accessible to students, on the probability of confessing. Taking into account that students skip some interventions, we propose to measure the causal effects associated with the accumulated number of interventions that the student attended at the moment of its confession. In addition, we control for a dummy indicating control group against treatment group.

The dependent variable in our econometric specification is *to confess* ($y_{it} = 1$) or *not to confess* ($y_{it} = 0$) the professor's mistake. However, we should take into account in our econometric specification that the same student can be randomly selected many times. In Table 5 in Appendix can be seen the probabilities of being selected different number of times in the experiment. In addition, we should take into account unobserved heterogeneity. Therefore, we estimate a hierarchic longitudinal random effects logit model.

In particular, $y_{it} \sim Bernoulli(\theta_{it})$ such that $logit(\theta_{it}) = \frac{\theta_{it}}{1-\theta_{it}} = x_{it}'\beta + b_i + \epsilon_{it}$ where $b_i \sim \mathcal{N}(0, V_b)$ captures individual unobserved heterogeneity, β is a k–dimensional vector of fixed effects parameters, x_{it} are regressors, and $\epsilon_{it} \sim \mathcal{N}(0, \sigma^2)$ are stochastic errors.

The likelihood contribution of each individual is hard to evaluate in this specification due to the presence of unobserved heterogeneity,

$$f(y_{it}|\beta, b_i, \sigma^2, V_b) = \prod_{i=1}^{n} \int \left[\prod_{t=1}^{n_i} \left[\Lambda(x_{it}'\beta + b_i + \epsilon_{it})\right]^{y_{it}} \left[1 - \Lambda(x_{it}'\beta + b_i + \epsilon_{it})\right]^{1-y_{it}} \right] \phi(b_i|0, V_b) db_i$$

where $\Lambda(u) = \frac{exp(u)}{1+exp(u)}$, $\phi(u|\mu, \Sigma) = \frac{1}{(2\pi)^{\frac{p}{2}}|\Sigma|^{\frac{1}{2}}} exp\left[-\frac{1}{2}(u-\mu)'\Sigma^{-1}(u-\mu)\right]$, n is the total sample size, and n_i is the sample size associated with each individual i.

So we follow a Bayesian approach using latent variables to estimate this model (Chib & Carlin, 1999). In particular, the latent variable is given by the logit link function, and as this function is linear in parameters, we can see this model as a Gaussian model such that we can use standard conjugate priors, that is, $\beta \sim \mathcal{N}(\beta_0, B_0)$, $\sigma^{-2} \sim \mathcal{G}(\alpha_0, \delta_0)$ and $V_b^{-1} \sim \mathcal{W}(v_0^{-1} R_0, v_0)$. We set "non–informative" priors in our estimation. In particular, we centered at zero the fixed effect parameters, and use an over–disperse diagonal covariance matrix with elements equal to 1.0E6. In addition, we set $\alpha_0 = \delta_0 = 0.001$ (D. Spiegelhalter et al., 2003), $v_0 = q = 3$ and $R_0 = diag\{0.1\}$. Then

$$\pi(\beta, b_i, \sigma^2, V_b|y_{it}) \propto f(y_{it}|\beta, b_i, \sigma^2, V_b)\pi(\beta)\pi(\sigma^{-2})\pi(V_b^{-1})$$

Under this framework, we can obtain conditional posterior distributions for the parameters, and as consequence, Gibbs sampling algorithms can be used to obtain posterior chains. In particular, the fixed effects parameters and the random effects parameters are distributed multivariate normal, the stochastic error variance is inverse–gamma, and the covariance matrix of the random effects is inverse–Wishart (Zeger & Karim, 1991; Chib & Carlin, 1999). In addition, the Metropolis-Hastings algorithm is used to obtain the posterior chain of the latent variables using as a proposal a normal distribution, and getting the implicit probability through the inverse function of the logit function for building the ratio of probabilities, proposal against actual, using the product of the Bernoulli and Gaussian distributions (Martin et al., 2017).

In addition, we performed different robustness checks using hierarchical longitudinal random effects linear probability models and linear probability models.

Other issue that we handle in our methodological approach is model uncertainty regarding regressors. In particular, we have 24 possible controls, this implies 2^{24} (16,777,216) possible models. Hence, we use the Deviance Information Criterion in order to choose the best model. We implement the following algorithm:

Algorithm A1 Model selection

1: Given an actual model (M_t^a), $t = 1, 2, \ldots, S$ where $S = 10,000$ is the total number of iterations in our application.
2: Propose a candidate model (M^c) with the same regressors as M_t^a, but adding one additional regresors with probability 0.5 or deleting one regressor with probability 0.5.
3: Calculate the DIC for the actual model (DIC^a) and the candidate model DIC^c.
4: Set $\alpha = Min\{1, DIC^a/DIC^c\}$.
5: Draw $u \sim \mathcal{U}(0, 1)$
6: If $u < \alpha$, then $M_t^a = M^c$, other case $M_t^a = M^t$.

To calculate the DIC (D. J. Spiegelhalter et al., 2002), we need to get the deviance, that is $D(\beta, b_i, \sigma^2, V_b) = -2logf(y_{it}|\beta, b_i, \sigma^2, V_b)$, and the expected deviance, $\overline{D} = E_\theta[D(\theta)]$, where $\theta = \{\beta, b_i, \sigma^2, V_b\}$. Then, we calculate the effective number of parameters as $p_D = \overline{D} - D(\theta^*)$, where θ^* is usually the posterior mean estimate. Finally, we get $DIC = p_D + \overline{D}$.

4 Results

We can see in Table 2 descriptive statistics and tests for mean differences between the control and treatment groups. In general, we observe that there are not significant statistical differences between the control and treatment groups, except for the probability of confessing, which is our main objective variable, mother education, which is higher for the control group, and extrinsic motivation, which is also higher for the control group.

Table 2: Descriptive statistics: Difference in mean Welch's test

Variable	Mean Control	Mean Treatment	t– statistic
Proportion of confessing	0.25	0.08	2.12
Gender (women)	0.44	0.60	-0.88
Age	20.50	20.87	-0.72
Catholic	0.81	0.60	1.28
Income 0[a]	0.00	0.13	-1.47
Income 1[b]	0.19	0.33	-0.90
Income 2[c]	0.37	0.33	0.23
Income 3[d]	0.44	0.20	1.42
Caucasian	0.75	0.87	-0.80
Education father	16.87	15.33	1.12
Education mother	17.44	15.13	2.03
Private school	0.94	0.87	0.64
GPA	3.98	3.86	1.06
Scholarship	0.06	0.27	-1.52
College groups	0.37	0.40	-0.14
Smoker	0.06	0.20	-1.11
Alcohol consumer	0.56	0.46	0.52
Study for other subjects (hours)	3.62	3.27	0.79
Econometrics study (hours)	5.87	5.33	0.76
Intrinsic motivation	22.31	21.40	0.85
Extrinsic motivation	12.62	10.73	3.14
Self–efficacy	12.00	10.80	0.87
C-Index	40.69	31.41	1.20

[a] Less than 632 USD monthly
[b] 632-1580 USD monthly
[c] 1580-3160 USD monthly
[d] More than 3160 USD monthly

Given that our econometric specification is non-linear, we can calculate the marginal effects for specific individuals. In Table **??** can be seen the outcomes associated with our representative student. In particular, she is 21 years

Table 3: Marginal effect of attended interventions: Representative student.

	Quantiles						
Variable	2.5%	50%	97.5%	Cog^a	$Educ^b$	SE^c	H^d
Amount	0.94	2.00	3.72	✓	X	X	X
	2.08	2.99	4.04	✓	✓	X	X
	2.27	3.25	3.53	✓	✓	✓	X
	2.05	3.25	4.01	✓	✓	✓	✓

[a] Cognitive controls: C-index, motivation and self–efficacy.
[b] Educational controls: Parents' education, school (private or public), GPA, scholarship (yes or not), Belonging to college groups and study hours (econometrics and others).
[c] Socioeconomic controls: Age, gender, Catholic, monthly income and Caucasian.
[d] Habits controls: alcohol drinker and smoker.

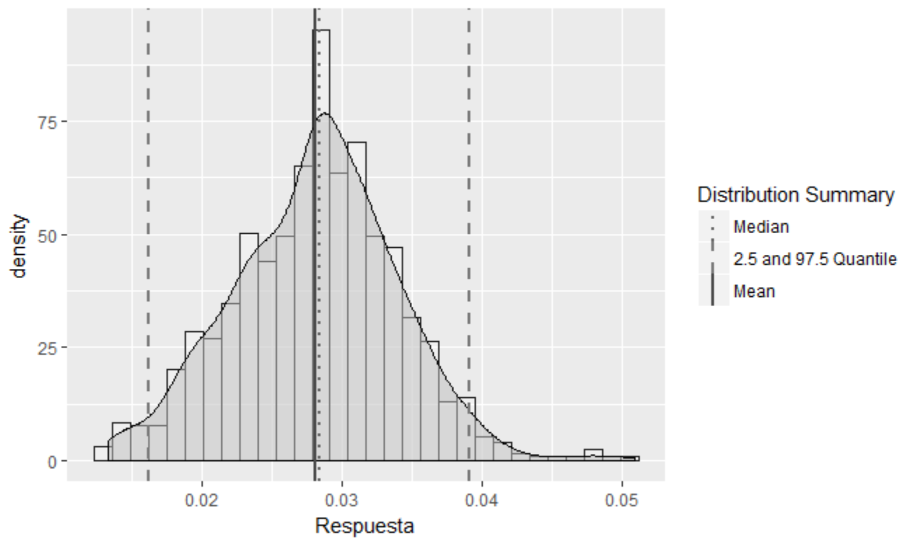
old, Catholic, Caucasian, her household monthly income is between 1,580 and 3,160 USD, both her mother and father have 16 years of education, she is graduated from a private school, without scholarship, GPA equal to 3.9, a C-index equal to 35.2, average motivation (intrinsic and extrinsic) and self–efficacy (22, 11.8, and 11.5). She studies econometrics and other subject 5.8 and 3.4 hours per weak, respectively. She is no smoker and no alcohol consumer.

We observe in Table **??** that the marginal effect associated with the amount of attended interventions is robust, and statistically significant under different set of controls. In particular, we observe that one additional integrity intervention attended by the representative student implies an average increase in the probability of confessing equal to 3.35%, with a 95% credible interval from 2.05% to 4.01% using all the set of controls. We show in the Appendix that these outcomes are robust regarding econometric methodology (see Table **??**).

We perform Algorithm A1 to select a final model in our application. We observe in Table 4 the marginal effects associated with the regressors in the selected specification.

Table 4: Marginal effects: Selected model using Algorithm A1.

	Quantiles			
Variable	2.5%	50%	97.5%	Mean
Mother educ	0.31	1.50	2.33	1.48
Religious	-2.70	13.20	20.77	11.99
Caucasic	2.71	11.67	21.17	11.65
Number	1.62	2.84	3.91	2.81
Treatment	4.04	14.98	24.50	14.74
Income 2	-0.96	4.70	15.02	5.79
GPA	4.29	6.32	7.64	6.11
Scholarship	16.95	23.78	31.94	24.19
VCV	5.5e-04	1.32e-02	4.32e-02	1.59e-02
σ^2	4.01e-04	2.48e-03	1.7e-02	3.94e-03



5 Conclusions

References

Anderman, E. M., Griesinger, T., & Westerfield, G. (1998). Motivation and cheating during early adolescence. *Journal of Educational Psychology*, *90*(1), 84.

Anderman, E. M., & Midgley, C. (2004). Changes in self-reported academic cheating across the transition from middle school to high school. *Contemporary Educational Psychology*, *29*(4), 499–517.

Ayal, S., Gino, F., Barkan, R., & Ariely, D. (2015). Three principles to revise people's unethical behavior. *Perspectives on Psychological Science*, *10*(6), 738–741.

Bandura, A. (1978). The self system in reciprocal determinism. *American psychologist*, *33*(4), 344.

Bandura, A. (1999). Social cognitive theory of personality. *Handbook of personality: Theory and research*, 154–196.

Bandura, A. (2002). Selective moral disengagement in the exercise of moral agency. *Journal of moral education*, *31*(2), 101–119.

Bandura, A., Barbaranelli, C., Caprara, G. V., & Pastorelli, C. (1996). Mechanisms of moral disengagement in the exercise of moral agency. *Journal of personality and social psychology*, *71*(2), 364.

Barkan, R., Ayal, S., & Ariely, D. (2015). Ethical dissonance, justifications, and moral behavior. *Current Opinion in Psychology*, *6*, 157–161.

Becker, G. (1968). Crime and punishment: An economic approach. *Journal of Political Economic*, *76*(2), 169–217.

Bonetti, S. (1998). Experimental economics and deception. *Journal of Economic Psychology*, *19*(3), 377–395.

Bowers, W. J. (1964). Student dishonesty and its control in college.

Chib, S., & Carlin, B. (1999). On MCMC sampling in hierarchical longitudinal models. *Statistics and Computing*, 17–26.

Finn, K. V., & Frone, M. R. (2004). Academic performance and cheating: Moderating role of school identification and self-efficacy. *The Journal of Educational Research*, *97*(3), 115–121.

Gneezy, U. (2005). Deception: The role of consequences. *The American Economic Review*, *95*(1), 384–394.

Grimes, P. W., & Rezek, J. P. (2005). The determinants of cheating by high school economics students: a comparative study of academic dishonesty in the transitional economies. *International Review of Economics Education*, *4*(2), 23–45.

Haines, V. J., Diekhoff, G. M., LaBeff, E. E., & Clark, R. E. (1986). College cheating: Immaturity, lack of commitment, and the neutralizing attitude. *Research in Higher education*, *25*(4), 342–354.

Henrich, J., Boyd, R., Bowles, S., Camerer, C., Fehr, E., Gintis, H., & McElreath, R. (1968). In search of homo economicus: behavioral experiments in

15 small-scale societies. *The American Economic Review*, *91*(2), 73–78.

Hersch, G. (2015). Experimental economics' inconsistent ban on deception. *Studies in History and Philosophy of Science Part A*, *52*, 13–19.

Hyman, I. E., Boss, S. M., Wise, B. M., McKenzie, K. E., & Caggiano, J. M. (2010). Did you see the unicycling clown? inattentional blindness while walking and talking on a cell phone. *Applied Cognitive Psychology*, *24*(5), 597–607.

Jordan, J., Mullen, E., & Murnighan, J. K. (2011). Striving for the moral self: The effects of recalling past moral actions on future moral behavior. *Personality and Social Psychology Bulletin*, 1-13.

Kanat-Maymon, Y., Benjamin, M., Stavsky, A., Shoshani, A., & Roth, G. (2015). The role of basic need fulfillment in academic dishonesty: A self-determination theory perspective. *Contemporary Educational Psychology*, *43*, 1–9.

Kelman, H. C. (1967). Human use of human subjects: the problem of deception in social psychological experiments. *Psychological bulletin*, *67*(1), 1.

Khodaie, E., Moghadamzadeh, A., & Salehi, K. (2011). Factors affecting the probability of academic cheating school students in tehran. *Procedia-Social and Behavioral Sciences*, *29*, 1587–1595.

Lind, G. (1999). Scoring of the moral judgment test (mjt). *Unpublished manuscript, University of Konstanz, Konstanz. In: Verweij D Hofhuis, K and Soeters J (2007) Moral Judgement within the Armed Forces, Journal of Military Ethics*, *6*(1), 19–40.

Lind, G. (2008). The meaning and measurement of moral judgment competence. *Contemporary philosophical and psychological perspectives on moral development and education*, 185.

Lipson, A., & McGavern, N. (1993). Undergraduate academic dishonesty at mit. *Results of a study of attitudes and behaviour of undergraduates*.

Martin, A., Quinn, K., & Park, J. (2017). *Mcmcpack* (Tech. Rep.).

Mazar, N., Amir, O., & Ariely, D. (2008). The dishonesty of honest people: A theory of self-concept maintenance. *Journal of marketing research*, *45*(6), 633–644.

McCabe, D. L., Butterfield, K. D., & Trevino, L. K. (2006). Academic dishonesty in graduate business programs: Prevalence, causes, and proposed action. *Academy of Management Learning & Education*, *5*(3), 294–305.

McCabe, D. L., & Trevino, L. K. (1997). Individual and contextual influences on academic dishonesty: A multicampus investigation. *Research in higher education*, *38*(3), 379–396.

McDaniel, T., & Starmer, C. (1998). Experimental economics and deception: A comment. *Journal of Economic Psychology*, *19*(3), 403–409.

Piaget, J. (2013). *Child's conception of number: Selected works* (Vol. 2). Routledge.

Ryan, R. M., & Deci, E. L. (2000). Self-determination theory and the facilitation of intrinsic motivation, social development, and well-being. *American psychologist*, *55*(1), 68.

Shaughnessy, J. J., & Zechmeister, E. B. (1985). *Research methods in psychology.* Alfred A. Knopf.

Shu, L. L., Mazar, N., Gino, F., Ariely, D., & Bazerman, M. H. (2012). Signing at the beginning makes ethics salient and decreases dishonest self-reports in comparison to signing at the end. *Proceedings of the National Academy of Sciences*, *109*(38), 15197–15200.

Spiegelhalter, D., Best, T., Gilks, W., & Lunn, D. (2003). *BUGS: Bayesian inference using Gibbs sampling* (Tech. Rep.). England: MRC Biostatistics Unit. (www.mrc-bsu.cam.ac.uk/bugs/)

Spiegelhalter, D. J., Best, N. G., Carlin, B. P., & Linde, V. D. (2002). Bayesian measures of model complexity and fit. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, *64*(4), 583–639.

Thaler, R. (1980). Toward a positive theory of consumer choice. *Journal of Economic Behavior & Organization*, *1*(1), 39–60.

Valentino, K., Cicchetti, D., Toth, S. L., & Rogosch, F. A. (2011). Mother–child play and maltreatment: A longitudinal analysis of emerging social behavior from infancy to toddlerhood. *Developmental psychology*, *47*(5), 1280.

Zeger, S., & Karim, M. (1991). Generalized linear models with random effects: A Gibbs sampling approach. *Journal of the American Statistical Association*, 79–86.

Appendix

Fig. 1: Quiz grading scheme.

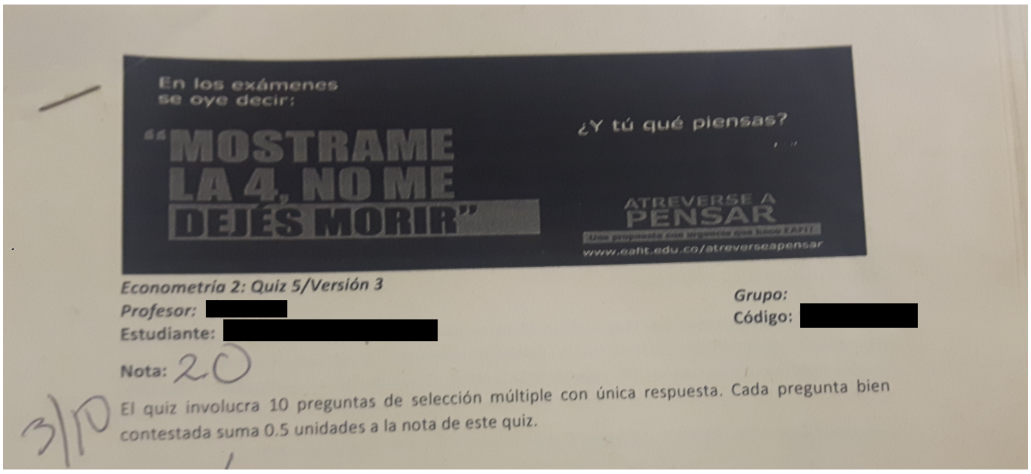


Table 5: Probability of being selected different times in the experiment.

Times	Probability (%)
0	4.9
1	16.8
2	26.4
3	25.1
4	16.1
5	7.4
6	2.4
7	0.6
8	0.1
9	1.3 e-02
10	1.1 e-03
11	6.0 e-05
12	1.4 e-06

Table 6: Marginal effect of attended interventions of representative student: Robustness checks.

	Quantiles			
Variable	2.5%	50%	97.5%	Mean
Longitudinal Logit	2.05	3.25	4.01	3.15
Longitudinal Linear	2.87	3.69	4.56	3.70
Linear	0.00	3.63	6.72	3.63