WORKING PAPERS

An offer you can refuse: the effects of transparency with endogenous conflict of interest

Melis Kartal
James Tremevan

March 2016

Working Paper No: 1602

DEPARTMENT OF ECONOMICS

UNIVERSITY OF VIENNA

An offer you *can* refuse: the effects of transparency with endogenous conflict of interest[*]

Melis Kartal[†]
University of Vienna & VCEE

James Tremewan[‡]
University of Vienna & VCEE

March 16, 2016

Abstract

This paper studies the effects of transparency on information transmission and decision-making theoretically and experimentally. We develop a model in which a decision maker seeks the advice of a better-informed adviser. Before giving advice, the adviser may choose to accept a side payment from a third party, where accepting this payment binds the advisor to give a particular recommendation, which may or may not be dishonest. Without transparency, the decision maker learns only the recommendation of the adviser. With transparency, the decision maker learns in addition the decision of the adviser with respect to the side payment. Prior research has shown that transparency is either ineffective or harmful to decisionmakers—because conflicted advisers become more dishonest in their advice. The novelty of our model is that the conflict of interest is endogeneous as the adviser can choose to decline the third-party payment. Our theoretical results predict that transparency is never harmful and may help decision makers. Our experimental results show that transparency improves the accuracy of decision making. However, we also observe that (i) while transparency clearly improves decision making when it is mandatory, the evidence in favor of a voluntary form of transparency is much weaker, and that (ii) the positive effects of transparency decline over time. *JEL Codes: C72, C91, D82, D83.*

[*]We would like to thank Tim Johnson and Dimitri Landa, as well as seminar audiences at Heidelberg University, University of Vienna, WZB Berlin Social Science Center, NYU CESS Annual Experimental Political Science Conference, and the 2015 ESA European Meeting.

[†]Department of Economics, University of Vienna, Email: *melis.kartal@univie.ac.at*.

[‡]Department of Economics, University of Vienna, Email: *james.tremewan@univie.ac.at*.

1 Introduction

In 2013, the Occupational Safety and Health Administration (OSHA) in the U.S. began a public consultation on setting new limits for working with silica dust, which is a major health hazard for construction workers that causes serious lung disease. The OSHA created substantial controversy in the Senate when it requested for the first time that those submitting scientific evidence should disclose their funding sources. A number of senators protested against the request arguing that revealing this type of information would bias the judgement of the agency. In turn, the head of the OSHA defended the request vigorously, claiming that transparency is indispensable so that the information that the agency bases its decision upon meets the highest standard of integrity.[1] How transparency affects advice and whether it improves the accuracy of decision-making in settings such as this, where the expert might be influenced by a third party, is the topic of this paper. Transparency enables the decision maker to learn whether or not the expert accepted a side payment from the third party—without transparency, the decision maker learns only the advice.

Advice is prevalent in a variety of settings, ranging from regulatory agencies, legislatures and judiciaries to medical services and financial markets. In such settings, decision makers often face complex decisions with uncertain outcomes, and therefore seek the advice of an expert in order to increase the likelihood of a successful decision. However, information transmission from the expert to the decision maker may be marred: Even if the expert and the decision maker do not have a conflict of interest *in the absence of a third party*, a third party (e.g. a special interest group or an industry) may sway the expert's advice in its favor by offering him a financial reward.

On the one hand, transparency is assumed to remedy this sort of situation: It protects the decision maker by revealing whether the expert has a conflict of interest that might lead him to give biased advice. On the other hand, one counter-argument against transparency is that disclosing this type of information results in a bias itself: Even if the expert's advice is truthful, the decision maker may dismiss the advice if the expert has accepted funding from an industry or a special interest group.[2] According to the proponents of this idea, the bias against experts funded by third parties is harmful, so "the conflict of interest mania"

[1] Published in *Nature* on March 4, 2014. In the same issue, an editorial piece argued that regulatory agencies must demand conflict of interest statements for the research that they use.

[2] See, for example, Stossel (2005), Weber (2009) and Stossel and Stell (2011).

must be cured.[3] An intimately related debate is whether transparency should be voluntary or mandatory. For example, organizations such as Transparency International advocate mandatory registering of lobbying activity. Most politicians and lobbyists agree that this should be the case; however, in countries where registers exist they often remain voluntary.[4,5]

We address these issues concerning the effects of transparency on strategic information transmission and decision-making theoretically and experimentally. We develop a model in which there are two states of the world (labeled L and R) and two possible policies (labeled l and r). The adviser is an expert who is perfectly informed about the state of the world whereas the decision maker knows only the prior probability of each state. The adviser recommends a policy to the decision maker and, consequently, the decision maker makes a policy choice. The payoff of the decision maker is maximized if the chosen policy matches the state of the world. All else being equal, the adviser and the decision maker have no conflict of interest; i.e., their payoffs are aligned. However, prior to the policy recommendation stage (and after the adviser learns the true state of the world), a third party offers a side payment to the adviser. The third party strictly prefers policy r regardless of the state of the world. If the adviser accepts the offer, then he must recommend policy r.

We consider the following scenarios. In the *transparency* condition, the decision maker is informed whether or not the adviser accepted the third-party payment. In the *non-transparency* condition, the decision of the adviser regarding the payment is not disclosed. We also study a *voluntary-transparency* condition in which transparency is not enforced and the adviser chooses whether or not to disclose his decision regarding the third-party payment.

Prior experimental research on the effects of transparency has produced bleak results: Transparency is either ineffective or has *adverse* effects on decision makers (Cain *et al.*, 2005, 2011; Koch and Schmidt, 2010; Rode, 2010; Loewenstein *et al.*, 2011; Loewenstein *et al.*, 2012; and Ismayilov and Potters, 2014). Our model differs from the previous studies in that the conflict of interest between the adviser and the decision maker is endogenous. On the one hand, the adviser and the decision maker have no conflict of interest if the adviser rejects the third-party payment. Thus, in the transparency condition rejecting the payment is a message from the adviser to the decision maker that his advice is honest. On the other hand, accepting the third-party payment per se is not inherently dishonest. Whether the adviser does accept the third-party payment and how the decision maker interprets the advice of an

[3] Rago, Joseph. "A Cure for 'Conflict of Interest' Mania." Wall Street Journal, 26 June 2015.

[4] http://www.transparencyinternational.eu/wp-content/uploads/2015/04/Lobbying_web.pdf

[5] http://www.oecd.org/gov/ethics/Lobbying-Brochure.pdf

adviser who accepted the third-party payment are determined in equilibrium.

Our theoretical results are as follows.[6] In the non-transparency condition, the adviser always accepts the payment from the third party and recommends r. As a result, the adviser's recommendation is uninformative. We denote this equilibrium "the corrupt equilibrium." In the transparency condition, there are two equilibria that we are particularly interested in. The first one is the corrupt equilibrium, in which behavior is the same as in the equilibrium of the non-transparency condition. The second equilibrium involves honest adviser behavior: The adviser always rejects the third-party payment and makes a truthful recommendation, which the decision maker follows. We denote this "the honest equilibrium." Sustaining the honest equilibrium requires the type of bias which opponents of transparency argue will be the result of disclosure: If the adviser accepts the third-party payment and recommends policy r, this prompts the decision maker to believe that the adviser is dishonest and choose policy l. This decision maker bias emerges *only off-the-equilibrium path* as the adviser rejects the payment in the honest equilibrium. Moreover, this bias is *necessary* to improve decision making with transparency because there is a significant agency problem. Both honest and corrupt equilibria also exist with voluntary transparency.

Our model does not generate very sharp predictions regarding the effects of transparency on information transmission due to a multiplicity of equilibria in the transparency and voluntary-transparency conditions—this multiplicity persists even if we apply suitable versions of the commonly used equilibrium refinements, such as *the Intuitive Criterion* and *D1*. To gain further insights regarding the effects of transparency, we designed and ran an experiment on the basis of our theoretical model, implementing each of the three conditions discussed above.

Overall, transparency clearly improves decision making relative to non-transparency; however, the evidence regarding the effect of voluntary transparency is much weaker. We find that transparency improves the accuracy of decisions made in state L, the state in which the adviser gives a "dishonest" recommendation in the corrupt equilibrium. While transparency improves the accuracy of decisions in state L, it has no impact on the accuracy in state R. Thus, we conclude that transparency improves decision making.

The mechanism through which decision making is improved in the transparency condition is consistent with our theory. Many decision makers and advisers view rejecting the

[6] In our theoretical analysis, we focus on parameter values such that there is a nontrivial "agency problem"—i.e., the adviser can be swayed by the third party. Otherwise, there is no difference between equilibrium outcomes in the transparency and non-transparency conditions.

payment from the third party as a way to enhance the adviser's credibility. When the state is R, more advisers reject the payment and recommend r in the transparency treatment than in the non-transparency treatment. When the state is L, many advisers reject the payment and recommend the correct policy even in the non-transparency treatment—this can be explained by lying-aversion—however, even more advisers do so with transparency. Thus, advisers' willingness to reject the payment in the transparency treatment stems not only from lying-aversion but also from strategic signalling of honesty. Signalling honesty by refusing the third-party payment is potentially beneficial because if an adviser accepts the payment and recommends policy r in the transparency treatment, a sizeable proportion of decision makers find it suspicious and choose policy l. One caveat is that although the fraction of decision makers who mistrust advisers that accept the payment is sizeable, it is also far from being a majority: Advisers learn that the negative bias among decision makers against advisers who accept the third-party payment is not too prevalent. As a result, positive effects of transparency do not seem to be resilient and decline over time.

Our study sheds light on the effects of transparency on the adviser and the decision maker behavior in an environment where the adviser might be swayed by a third-party who favors the implementation of a specific policy. In particular, ours is the first study which shows that transparency can help decision makers in both theory and practice. Transparency is becoming more and more important because the share of private enterprise in the funding of research has been rising steeply. According to the National Science Foundation (NSF) in the US, the share of industry and government roughly tracked each other until late 80's. However, industry has since considerably outpaced government in terms of research funding.[7] This has bestowed corporations with an immense influence and ability to shape policy-making as well as public opinion via the research institutions and experts that they fund.[8] In response to this, regulatory agencies, academic journals, NGOs and government agencies have started demanding more transparency. However, this demand for transparency is not without its

[7] http://www.nsf.gov/statistics/seind14/index.cfm/chapter-4/c4s1.htm

[8] For example, the energy industry and the tobacco industry have used experts extensively in their endeavour to supress information regarding the hazards of their products. Recently, ExxonMobil, the world's biggest oil company, made the headlines for its funding of climate change denial. It was reported that ExxonMobil knew as early as 1981 of climate change, and despite this the firm spent a huge amount of money over the next three decades to promote its denial (*The Guardian*, July 8, 2015). On a related note, 47 percent of US voters believe that climate change is caused by human activities, compared with 97 percent of climate scientists (Yale Project on Climate Change Communication 2013).

backlash, as we discussed. Therefore, transparency is a not only very important but also sensitive issue that calls for a thorough theoretical and empirical evaluation.

The paper proceeds as follows. Section 2 discusses the related literature. Section 3 presents the theoretical model and Section 4 the theoretical results. Sections 5 and 6 describe the experimental setup and the experimental results, respectively. Presenting all three conditions in full generality requires a considerable increase in length and complexity of exposition with little corresponding increase in insight; therefore, we focus on only the transparency and non-transparency conditions throughout Sections 3-6. The main results for the voluntary-transparency condition are presented in Section 7. Section 8 concludes.

2 Related Literature

There is an extensive literature on strategic information transmission from a better-informed sender to a receiver, dating back to the seminal works by Crawford and Sobel (1982) and Sobel (1985). In this literature, the receiver and the sender have payoffs that are misaligned to a certain degree. This misalignment results in a bias in the sender's communication; the sender's equilibrium message to the receiver is noisy and can even be uninformative, depending on the precise structure of payoffs.

In most of the literature on strategic information transmission, the extent to which sender and receiver payoffs are (mis)aligned is exogenous. Our model endogenizes it since the adviser may choose to decline the third-party payment; transparency affects the incentives of the adviser to do so. With transparency, the adviser can reject the payment and signal to the decision maker that their payoffs are aligned. Thus, our transparency condition allows the adviser to engage in "costly signaling" whereas the non-transparency condition is essentially a cheap-talk game.

The extent to which the adviser and the decision maker payoffs are aligned is also endogenous in Durbin and Iyer (2009) and Inderst and Ottaviani (2012) but this is due to strategic third parties. There is a single third party in Durbin and Iyer whereas there are multiple third parties in Inderst and Ottaviani. In these models, third parties set side payments for the adviser (sales commissions, bribes, etc.) in order to maximize their return. Inderst and Ottaviani (2012) also analyzes the effects of transparency in their setting and show that transparency can have adverse effects on the decision maker, unlike in our model. The focus of Inderst and Ottaviani is on a setting different than ours: In their model, third parties produce horizontally-differentiated products and compete for consumers using advisers who are incentivized by sales commissions. We abstract from the competition of third parties and focus instead on the effects of transparency with a single—or disproportionately powerful—third party (e.g., oil and energy industry, tobacco industry, gun rights lobby, etc.) that can affect policy and sway public opinion through political lobbying and funding research institutes and experts.

Also related is the model by Potters and van Winden (1992, 2000) in which there are two players, a policy-maker and a better-informed interest group (i.e., a lobby) that can send the policy-maker a costly message. Their model differs from the cheap talk literature since sending a message entails a cost to the lobbyist—this relates to lobbying costs. Our model can be thought of as a lobbying model (absent lobbying costs) in which there is a third party that can influence the message of the lobbyist to the policy-maker.

As already mentioned, prior experimental research on transparency pointed out that disclosing conflict of interests is either ineffective or harmful to decision makers (Cain *et al.*, 2005, 2011; Koch and Schmidt, 2010; Rode, 2010; Loewenstein *et al.*, 2011; Loewenstein *et al.*, 2012; and Ismayilov and Potters, 2014).[9] The most important distinction of our model from the previous experiments is that the adviser can choose to avoid a conflict of interest. In prior research, advisers did not have a choice regarding their incentives and were always conflicted. The only experiment other than ours that documents positive effects of transparency is the independent study by Sah and Loewenstein (2014), which we became aware of while completing our work. Sah and Loewenstein (2014) does not provide a formal model of advice. In their experimental design, the adviser can accept or reject a conflict of interest. If the adviser accepts the conflict of interest, then the adviser and the decision maker have misaligned payoffs; otherwise, their payoffs are aligned. In the transparency treatment, the decision maker is informed whether or not the adviser accepted the conflict of interest. Our model differs from their design because accepting the third-party fund is

[9] Loewenstein et al. (2011) put forth two explanations for the adverse effects of transparency that they observe. The first one is strategic. Advisers anticipate that receivers will discount their advice once their conflict of interest is revealed. So, they tend to give more biased advice in order to counteract such discounting by decision makers. However, decision makers fail to sufficiently discount biased advice and thus obtain worse outcomes with transparency. The second explanation refers to moral considerations: Transparency may result in moral licensing. According to the moral licensing hypothesis, information disclosure can undermine the willingness of the adviser to engage in moral behavior, which harms the decision maker.

not equivalent to having a conflict of interest—the adviser can accept the fund and still optimally give honest advice. More generally, our model is consistent with the idea that a relationship between the adviser and the third party is not inherently wrong.

Finally, our model and experimental results relate to studies on lying aversion. Previous experiments have shown that people may find lying morally costly and avoid it (see, for example, Gneezy, 2005; Sanchez-Pages and Vorsatz, 2007; Hurkens and Kartik, 2009; Gibson *et al.*, 2013). Our results are consistent with previous findings: Many advisers choose to tell the truth even if there is a material incentive to lie.

3 The Model

We develop a model that involves two active players, a decision maker (D) and an adviser (A), and an inactive third-party, namely the special interest group (SIG).[10] There are two possible states of the world. Nature draws the state $S \in \{L, R\}$ such that

$$S = \begin{cases} L, & \text{with probability } p \\ R, & \text{otherwise.} \end{cases}$$

The prior probability p is common knowledge. The adviser (A) learns the true state of the world whereas the decision maker (D) knows only the prior p. After learning the state, A recommends a policy $s \in \{l, r\}$ to D, who then chooses a policy. D prefers the policy to match the state of the world. After learning the state and prior to making a recommendation, A decides whether or not to accept a payment from the special interest group (SIG).[11] SIG strictly prefers policy r regardless of the state of the world, and if A accepts the payment, then he must recommend r.[12] If A rejects the payment, then he decides whether to recommend l or r. We consider the following scenarios.

[10] We abstract from the strategic behaviour of the SIG and assume that it always offers a side payment to the adviser and that the side payment is an *exogeneously* determined amount. In this study, we focus on the decision of the adviser regarding the side payment with and without transparency and how the decision-maker interprets the advice from an adviser who accepted the payment.

[11] Therefore, accepting the payment is not inherently dishonest: The timing of the model allows the adviser to accept the side payment only in state R and remain honest.

[12] We relax this assumption and discuss its implications in Section 4.5. In that section, we also relax the assumption that the third-party payment does not affect the adviser's information and discuss its implications.

(i) In the "transparency condition", D learns whether or not A accepted the payment before choosing the policy.
(ii) In the "non-transparency condition", D has no information about the decision of A regarding the payment.
(iii) In the "voluntary-transparency condition", transparency is not enforced and A chooses whether or not to disclose his decision regarding the third-party payment.

As stated before, presenting all three conditions in full generality requires a considerable increase in length and complexity of exposition with little corresponding increase in insight; therefore, we focus on a comparative analysis of transparency and non-transparency conditions throughout Sections 3-6; Section 7 presents the main results for the voluntary-transparency condition.

3.1 Payoffs

D obtains a payoff of $\bar{\pi}$ if the chosen policy matches the state of the world, and a payoff of $\underline{\pi}$ otherwise, where $\underline{\pi} < \bar{\pi}$. A receives

(i) $\alpha > 0$ if D chooses the policy that A recommends,
(ii) $\gamma > 0$ if A recommends the "better" policy for D and
(iii) $\beta(S) \geq 0$ if A accepts the payment in state $S \in \{L, R\}$,

where $\beta(L) > \gamma$. We assume that $\alpha > 0$ because being followed is good for the adviser's reputation and enables future business prospects, and that $\gamma > 0$ due to reasons such as lying aversion or maintaining a reputation for honesty.

We allow for state-dependent side payments. Note that $\beta(L)$ and $\beta(R)$ can differ even if SIG is not informed about the true state. If $\beta(L) > \beta(R)$, then SIG offers A a fixed payment equal to $\beta(R)$ and commits to supplementing it by the amount $\beta(L) - \beta(R)$ if the state turns out to be L.

Finally, we impose the inequality $\beta(L) > \gamma$ because it rules out the case in which transparency makes no difference relative to the non-transparency condition.[13] For example, if $\gamma > \beta(L)$, then giving honest advice (or having an honest reputation) is very important to the adviser, and therefore, the adviser will never give dishonest advice "in a reasonable equilibrium" with or without transparency.

[13] We prove this claim in the Appendix.

4 Equilibrium Analysis

Let $a \in \{r_A, r_R, l\}$ denote A's action, where $a = r_A$ if A accepts the payment and recommends r; $a = r_R$ if A rejects the payment and recommends r; and $a = l$ if A recommends l (recall that A *must* reject the payment in order to recommend l). Next, let m denote A's "message" to D. In the transparency condition, D observes A's action completely. So, A's message is equivalent to his action; i.e., $m = a$. In the non-transparency condition, A's message is equivalent to his policy recommendation. If $a \in \{r_A, r_R\}$, then $m = r$. If $a = l$, then $m = l$. Hence, the message space is $\{r_A, r_R, l\}$ in the transparency condition and $\{r, l\}$ in the nontransparency condition.

We look for the perfect Bayesian Equilibria (PBE) of the games described above. Let $\Pr_A(m|S)$ denote the probability that A chooses message m in state S. Also, let $\mu(S|m)$ denote the belief D attaches to state S conditional on m, and let $\Pr_D(s|m)$ denote the probability that D chooses policy s after observing message m. A PBE entails the following:

(i) D's strategy maximizes his payoff given m and $\mu(S|m)$;

(ii) $\mu(S|m)$ is formed using A's strategy $a(S)$ by applying Bayes' rule whenever possible;

(iii) given $\mu(S|m)$ and D's strategy, A's strategy maximizes his payoff.

We maintain the assumption that $p < 1/2$ throughout our theoretical analysis and we also implement it in our experimental design. This gives rise to a more interesting contrast between the equilibrium predictions for the transparency and non-transparency conditions than the case where $p \geq 1/2$ (see Footnote 19 in Section 4.2 for details.).[14]

Assumption 1 $p < 1/2$.

4.1 Baseline Model: The Game Without SIG

We first study the model without SIG, as a benchmark for subsequent analysis. This analysis reflects in a simple way the need for reasonable equilibrium selection criteria. Since there is no SIG, $a \in \{l, r\}$ and $a = m$. In this case, there are two possible equilibria. The first equilibrium is fully informative: A always recommends the correct policy and D follows the recommendation. The second one is uninformative: A always recommends r regardless of the realized state, and D always follows. The second equilibrium is not only pareto-dominated but also dependent on the following "unreasonable" out-of-equilibrium belief. If

[14] The formal equilibrium analysis with $p \geq 1/2$ is available upon request.

A recommends l, then D interprets this as "strong evidence" for state R, which deters A from recommending l in state L.[15] But why should D believe that A is more likely to recommend l in state R than in state L? This is not *intuitive* given that A obtains an additional payoff $\gamma > 0$ (for honesty, reputation, etc.) if he recommends l in state L, not in state R. In order to eliminate equilibria (with and without SIG) that rely on unreasonable out-of-equilibrium beliefs, we impose a modification of the Intuitive Criterion by Cho and Kreps (1987).

Definition 1 *Message m is* ***equilibrium-dominated in state*** *$S \in \{L, R\}$ if the equilibrium payoff of A in S is greater than the highest possible payoff of A from m in S.*

Given this definition, we can present a version of the Intuitive Criterion suited to our setting with two states.

Definition 2 *An equilibrium* ***fails to satisfy the Intuitive Criterion (IC)*** *if there exists an out-of-equilibrium message m such that:*

(i) *m is equilibrium-dominated in state S; and*

(ii) *m is a profitable deviation for A in state $S' \neq S$ if D best-responds to m according to the belief $\mu(S|m) = 0$.*

The uninformative equilibrium does not satisfy the Intuitive Criterion (IC) because recommending l is equilibrium dominated in R, and recommending l in state L is a profitable deviation for A if D best-responds according to the belief $\mu(R|l) = 0$.[16] As a result, the only PBE that satisfies the IC in this game is fully informative: A always provides truthful advice and D follows the advice.

4.2 The Game With SIG and Non-transparency

Next, we analyze the game with SIG assuming that D cannot observe the decision of A regarding the side payment. We show that there exists a generically unique equilibrium, which we denote as the "corrupt equilibrium." In the corrupt equilibrium, A always chooses

[15] More formally, either $\mu(R|l) > 0.5$ and $\Pr_D(l|l) = 0$ or $\mu(R|l) = 0.5$ and $\Pr_D(l|l) < 5/6$ holds in the uninformative equilibrium. Both deter A from recommending l in state L.

[16] To see why l is equilibrium dominated in R, note that the equilibrium payoff of A in state R is $\alpha + \gamma$ whereas the highest possible payoff from choosing l in state R is α. Since $\gamma > 0$, l is equilibrium dominated in R.

r_A (i.e., accepts the payment and recommends r) and D always follows A's advice and chooses r.[17]

Proposition 3 *There exists a (generically) unique equilibrium in the non-transparency condition. On the equilibrium path, A always accepts the payment and recommends r, and D follows the advice with probability one.*

Proof. In the Appendix. ■

It is easy to see how the corrupt equilibrium arises.[18] Since A's message is always r in the corrupt equilibrium, it is uninformative and the decision maker must rely on his prior to choose a policy. Then, it is in the best interest of D to choose r—state R is more likely than state L by Assumption 1. Given this, A attains the highest possible payoff in both states if he chooses r_A and has no incentive to deviate. Thus, the belief $\mu(L|r) = p < 0.5$ is consistent with A's strategy.[19]

4.3 The Game With SIG and Transparency

We now analyze the game assuming that D observes whether or not A accepted the payment from the SIG. We focus on equilibria that satisfy the IC. For brevity of exposition, we assume that $\alpha + \gamma > \beta(L)$ in the main text—this is the case we implement in the experiment. However, in the Appendix we characterize fully the set of equilibria that satisfy the IC without resorting to this assumption. If $\alpha + \gamma > \beta(L)$, then there exist at most three equilibria. The first is the corrupt equilibrium in which A always chooses r_A and D always chooses r. The second is what we denote the "honest equilibrium." In the honest equilibrium, A always rejects the payment and provides honest advice, and D follows A's advice. Such

[17] What we mean by generic uniqueness is that the equilibrium path play and payoffs are unique. However, there are various belief systems that support the corrupt equilibrium because A never chooses l in equilibrium.

[18] This is where the assumption that $p < 1/2$ matters; a corrupt equilibrium exists and is the unique equilibrium *only if* $p < 1/2$. As a result, the comparison between the transparency and non-transparency conditions is much more clean and behaviorally more interesting under that assumption.

[19] This equilibrium may seem strange in that D takes advice from A even though this has no informational advantage. However, taking advice has no cost and allowing for a tiny probability that A never lies (indeed some advisers do not lie in our experiment) makes taking advice strictly optimal.

an equilibrium exists if $\alpha \geq \beta(R)$ in addition to our assumption that $\alpha + \gamma > \beta(L)$.[20] Depending on parameter values, there may also be a mixed-strategy equilibrium such that A always chooses r_A in state L (i.e., accepts the payment and gives false advice), and randomizes between r_R and r_A in state R; and D randomizes between l and r if A chooses r_A and chooses r if A chooses r_R.

Proposition 4 *Assume that $\alpha + \gamma > \beta(L)$. There are (at most) three equilibria that satisfy the IC:*

1) *Corrupt Equilibrium:* $\Pr_A(r_A|L) = \Pr_A(r_A|R) = 1$; $\Pr_D(r|r_A) = 1$; $\mu(R|r_A) = 1 - p$.[21]

2) *Honest Equilibrium (if $\alpha \geq \beta(R)$):* $\Pr_A(l|L) = 1$; $\Pr_A(r_R|R) = 1$; $\Pr_D(r|r_R) = 1$; $\Pr_D(l|l) = 1$; $\mu(R|r_R) = 1$, $\mu(R|l) = 0$ and $\mu(R|r_A) \leq 0.5$; $\Pr_D(r|r_A) = 0$.

3) *Mixed-strategy Equilibrium (if $\beta(L) \geq \beta(R) + \gamma$):* $\Pr_A(r_A|L) = 1$; $\Pr_A(r_A|R) = \frac{p}{1-p}$ and $\Pr_A(r_R|R) = \frac{1-2p}{1-p}$; $\Pr_D(r|r_R) = 1$; $\mu(R|r_R) = 1$; $\Pr_D(r|r_A) = 1 - \beta(R)/\alpha$; $\mu(R|r_A) = 0.5$.[22]

Proof. In the Appendix. ■

Our "normative" measure of interest is the accuracy of decision making rather than the efficiency of the aggregate expected payoff because the former has wider social welfare implications. In that sense, the honest equilibrium is superior to both the corrupt equilibrium and the mixed-strategy equilibrium whereas the corrupt equilibrium and the mixed-strategy equilibrium are ex-ante identical.[23] Thus, transparency increases the (ex-ante) expected accuracy relative to non-transparency.

The mechanism of the corrupt equilibrium is identical to that in the nontransparency condition. As for the honest equilibrium, first note that A must reject the payment in both states in order to sustain truthful information revelation because there is a significant "agency problem" (due to our assumption that $\beta(L) > \gamma$). If A always gave honest advice and yet

[20] Although the strategy of SIG is beyond the scope of this paper, it is unlikely that a strategic SIG would set $\beta(S)$ so high that $\alpha < \beta(R)$ or $\alpha + \gamma < \beta(L)$.

[21] Just as in the corrupt equilibrium with nontransparency, there are various out-of-equilibrium beliefs that support the equilibrium.

[22] In the knife-edge case with $\beta(R) + \gamma = \beta(L)$, there is also a continuum of mixed-strategy equilibria in which A randomizes between r_A and l in state L and between r_A and r_R in state R. See the details in the Appendix.

[23] The ex-ante expected accuracy is 100% in the honest equilibrium whereas it equals $100(1-p)\%$ in both the corrupt equilibrium and the mixed-strategy equilibrium.

accepted the payment in state R, then D would be better off following A's recommendation. But this would give A an incentive to deviate and choose r_A in state L.[24]

We now provide a more detailed discussion of the honest equilibrium focusing on the simple case with $\beta(R) = \beta(L)$. If $\beta(R) = \beta(L)$, then $\beta(R) + \gamma \leq \beta(L)$ doesn't hold, and there are only two equilibria that satisfy the IC: the corrupt equilibrium and the honest equilibrium (assuming that both $\alpha \geq \beta(R)$ and $\alpha + \gamma > \beta(L)$ hold). The honest equilibrium requires an off-the-equilibrium path belief $\mu(L|r_A)$ such that $\mu(L|r_A) \geq 0.5$: If the adviser were to choose r_A, then this would prompt the decision maker to suspect the honesty of the adviser, believe that $\mu(L|r_A) \geq 0.5$ and choose policy l—this deters the adviser from accepting the payment in equilibrium.[25] The belief that $\mu(L|r_A) \geq 0.5$ is not ruled out by the Intuitive Criterion because r_A is not an equilibrium-dominated action in state L. However, there is a caveat. Given that $\beta(R) = \beta(L)$ and $\gamma > 0$, the belief $\mu(L|r_A) \geq 0.5$ is *unreasonable according to D1* (Cho and Kreps, 1987), which is a more stringent criterion than IC. We introduce our definition of D1, next.

Definition 5 *Let $D_S(m)$ be the set of best responses by D that make A "strictly" prefer m over the equilibrium message in state S, and let $D^0_S(m)$ be the set of best responses by D that make A "weakly" prefer m over the equilibrium message in state S. An equilibrium satisfies D1 if $\mu(S|m) = 0$ holds for every out-of-equilibrium message m such that $D^0_S(m) \subseteq D_{S'}(m) \neq \varnothing$.*

If $\beta(R) = \beta(L)$, then the honest equilibrium does not satisfy D1 because $\beta(R) = \beta(L)$ and $\gamma > 0$ imply that $D^0_L(r_A) \subseteq D_R(r_A)$, and therefore $\mu(L|r_A) = 0$ is a requirement of D1. But the belief $\mu(R|r_A) = 1$ destroys the honest equilibrium. Intuitively, the adviser choice r_A must be a "strong" signal of state R according to D1 because—fixing the strategy of D—the payoff of A from choosing r_A is *always* higher in state R than in state L. The latter follows

[24]The agency problem disappears if $\beta(L) < \gamma$, and is weak if $\beta(L) = \gamma$. In either case, the payoff from honesty (or reputation for honesty) is quite high, which is why transparency does not make a difference relative to the non-transparency condition. We prove this claim in the Appendix after proving Proposition 4.

[25]Thus, sustaining honest equilibrium requires the type of bias which opponents of transparency argue will be the result of disclosure: If the adviser accepts the third-party payment, this prompts the decision maker to distrust the adviser. However, the bias emerges only off-the-equilibrium path (since the adviser rejects the payment in the honest equilibrium) and is a necessary condition so that transparency improves decision making because there is a significant agency problem (i.e., $\beta(L) > \gamma$).

because (i) the side payment is identical in both states; and (ii) if A chooses r_A, then he is telling the truth in state R but not in state L.

More generally, if $\beta(R) + \gamma > \beta(L)$, then the honest equilibrium does not satisfy D1. If, however, $\beta(R) + \gamma \leq \beta(L)$, then the honest equilibrium satisfies both IC and D1. In our treatments, we used parameter values such that $\beta(R) + \gamma \leq \beta(L)$ holds.

Proposition 6 *Assume that $\alpha + \gamma > \beta(L)$. If $\beta(R) + \gamma \leq \beta(L)$, then all three equilibria stated in Proposition 4 exist and satisfy D1.*

Proof. In the Appendix. ■

4.4 Extensions

We close the theory section with a discussion of richer variations of the model. It may be argued that in certain fields of expertise, the adviser may use third-party payment in order to acquire more accurate information. However, allowing for this would not change our main results unless we also relax the assumption that the adviser who accepts the third party funding must recommend r. Our results are robust to jointly assuming that accepting the payment increases the accuracy of the adviser's information and an adviser who accepts the third party payment can choose between recommending l or r with a positive—but not too high—probability.

Assume, as an example, that the adviser who accepts the third party payment is perfectly informed about the state whereas an adviser who does not accept it gets a signal $\sigma \in \{L, R\}$ about the state such that $\Pr(\sigma|S) = 0.8$ if $\sigma = S$. Further, assume that $p = 0.4$, $\alpha = 6$, $\beta(L) = 5$, $\beta(R) = 2$, $\gamma = 1$ (these are the parameter values that we use in our experimental design). If the adviser who accepts the side payment is free to choose the policy recommendation with a probability that is weakly lower than 0.38, then our main findings go through. For example, an honest equilibrium that satisfies the Intuitive Criterion and D1 exists and is the most informative equilibrium—even though the adviser is perfectly informed about the state only if he accepts the side payment. If, however, the aforementioned probability is higher than 0.38, then an honest equilibrium that satisfies the Intuitive Criterion and D1 does not exist in the transparency condition.

How plausible is it to assume that an adviser who accepts the third party payment will have to recommend the policy that the third party favors (with a sufficiently high probability)? We believe that it is a very reasonable assumption in many circumstances. As an example, tobacco companies, which have been repeatedly sued both for fraud—hiding from the public what they knew about their product—and in order to recover health costs associated with smoking, employed dozens of experts in order to testify on their behalf. Kenneth Ludmerer, a distinguished professor of history and medicine at Washington University in St. Louis testified as an expert on medical history on behalf of the tobacco industry over a period of 15 years. From the testimony of Ludmerer in 2002:

> *Question: Doctor, is it your opinion that cigarette smoking contributes to the development of lung cancer in human beings?*
>
> *Answer: I have no opinion on that.*

Ludmerer's testimony implies that tobacco companies cannot be held liable for any wrongdoing since he, the expert, has no opinion on whether cigarette smoking contributes to the development of lung cancer—as recently as in 2002! Ludmerer was paid more than $550,000 by the tobacco industry (Delafontaine, 2015). In our opinion, it is highly unlikely that the industry would pay this amount to an expert who could give an affirmative answer to the question above. On the contrary, Robert Proctor and Louis Kyriakoudes, two (of only three) experts who testified against the tobacco industry were subject to harassment by the industry (Delafontaine, 2015).[26]

5 Experimental Design and Hypotheses

The theory does not generate very sharp predictions regarding the effects of transparency on information transmission and decision-making due to equilibrium multiplicity. This multiplicity cannot be resolved using refinements such as the IC or D1, as already discussed. To gain further insights regarding the effects of transparency, we designed and ran an experiment which implemented our theoretical model. Our aim is to empirically answer the following research questions:

1) Does transparency help decision makers?

2) Do advisers behave differently if decision makers observe whether or not they accepted the side payment? In particular, are advisers less likely to accept the payment in order to "signal" their honesty?

[26] http://www.thenation.com/article/big-tobacco-and-historians/

3) Do decision makers take into account or ignore the decision of advisers regarding the side payment when they choose the policy?

4) Does the accuracy of decisions depend on whether transparency is mandatory or voluntary?

We ran three treatments implementing each of the transparency, non-transparency, and voluntary-transparency conditions. Subjects participated in only one of the three treatments. The experiment consisted of 40 rounds. We used stranger matching; i.e., subjects were randomly rematched with a new counterpart each period. At the beginning of the experiment, each subject was randomly assigned to be a receiver (i.e., decision maker) or a sender (i.e., an adviser). Subjects remained in the same role for the first 20 rounds. After 20 rounds were over, the roles were switched and subjects remained in their new role until the end of the experiment. Subjects were not informed at the beginning of the experiment that roles would be switched after the first 20 rounds. We applied role-switching because it facilitates learning; subjects better comprehend the decision problem of their counterpart and, thus, the overall game if they play in both roles.

Before the start of the actual experiment and the assignment of roles, subjects went through a tutorial and answered control questions in order to enhance their understanding of the game. Once the tutorial was over, the actual experiment began. Decision making in each round of the experiment was as described in the theory section. We used the strategy method with decision makers, giving us more observations at each information set and analyze better whether subjects use equilibrium strategies. We used the strategy method only with decision makers because the game would be more difficult to explain if we used the strategy method with both advisers and decision makers. We used neutral language and there was no mention of a special interest group. It was only stated that the sender could choose to accept an extra payment, in which case the sender had to recommend r. Instructions can be found in the Appendix.

Recall from the discussion in the previous section that the honest equilibrium satisfies both IC and D1 if $\beta(R) + \gamma \leq \beta(L)$. Parameter values we used in the experiment are as follows: $p = 0.4$, $\alpha = 6$, $\beta(L) = 5$, $\beta(R) = 2$, $\gamma = 1$, $\bar{\pi} = 10$, $\underline{\pi} = 5$. Given these values, there exist three equilibria that satisfy IC and D1 as described in Proposition 4.[27]

[27] We have also run sessions in which all the parameter values were the same except that $\beta(R) = \beta(L) = 4$. Therefore, $\beta(R) + \gamma \leq \beta(L)$ is not satisfied and the honest equilibrium is not robust to D1. We found that if $\beta(R) = \beta(L)$, transparency did not have a statistically significant impact on

The experiment was conducted at the experimental laboratory of Vienna Center for Experimental Economics (VCEE) at the University of Vienna. Subjects were recruited from the general student population in Vienna via e-mail solicitations. We ran three sessions for each treatment. In each session, there were 24 subjects. At the beginning of each session, each subject was randomly assigned to a matching group of 12 subjects. There was no interaction between subjects assigned to different matching groups. Therefore, we have six independent observations for each treatment.

All sessions were conducted using a computer program written in Z-Tree (Fischbacher, 2007). Earnings were denoted in "points" which were exchanged at the rate of 1 euro for 3 points. Each subject was paid for two randomly drawn rounds in each role. Sessions lasted around 75 minutes and the average payoff per subject was approximately 16 euros (including a fee for completing the questionnaire at the end of the experiment).

Table 1: Decision Maker Payoffs by State and Equilibrium type

Equilibrium Type	**State L**	**State R**
Corrupt Equilibrium	5	10
Honest Equilibrium	10	10
Mixed-Strategy Equilibrium	6.67	8.89

Finally, we present the two experimental hypotheses that we derive from our theoretical results given the parameters used in the design. The first hypothesis is the accuracy of decision-making in each state. Note that even though the hypothesis refers to decision maker payoffs, there is a one-to-one mapping from decision maker payoff to accuracy.

Hypothesis 1 (a) *Transparency makes decision makers better off in state L (via the honest equilibrium and the mixed-strategy equilibrium).* **(b)** *Transparency makes decision makers worse off in state R (via the mixed-strategy equilibrium).*

We derived Hypothesis 1 taking into account *all* possible equilibria given the realized state. Part (a) is due to the following. L is the state in which A has a financial incentive to give a "dishonest" recommendation—A accepts the third-party payment and recommends r in state L in the corrupt equilibrium (of either treatment). As a result, D makes the wrong decision

decision-making, and it had either zero or little impact on other dimensions. Therefore, we decided to move forward with treatments where $\beta(R) + \gamma \leq \beta(L)$ holds so that the honest equilibrium is robust to both IC and D1.

in state L and obtains a payoff of 5 in the corrupt equilibrium. The corrupt equilibrium is the unique equilibrium of the non-transparency treatment. However, the transparency treatment has other equilibria (namely, the honest equilibrium and the mixed-strategy equilibrium) which make D better off in state L than the corrupt equilibrium (see Table 1 for decision maker payoffs by state and equilibrium type). Part (b) follows because, while D always makes the correct decision in state R in the corrupt equilibrium and the honest equilibrium, this is not true in the mixed-strategy equilibrium of the transparency treatment; D sometimes makes the wrong decision in state R in the mixed-strategy equilibrium.

The existence of the honest equilibrium and the mixed-strategy equilibrium in the transparency and voluntary-transparency treatments generates the following hypotheses about adviser behavior.

Hypothesis 2 (a) *Transparency reduces the fraction of advisers who accept the payment in both state L and state R.* **(b)** *Transparency increases the fraction of advisers who reject the payment and recommend the correct policy in both L and R.*

6 Results

6.1 Does Transparency Help Decision makers?

In this section, we analyze the accuracy of decisions by treatment. We start with state L. Hypothesis 1 (a) states that transparency makes the decision maker better off in state L. As explained before, this is because the transparency treatment has the honest equilibrium and the mixed-strategy equilibrium both of which are better than the corrupt equilibrium in state L for the decision maker.

Table 2: Percentage of Correct Decisions by State and Treatment

	Transparency	**Non-transparency**	**Mann-Whitney p–value**
state L	46.13%	29.69%	0.008 (one-sided test)
state R	75.61%	79.27%	0.169 (one-sided test)

Table 2 displays the proportion of correct choices by state and treatment. According to Table 2, 46.1% of decisions in state L is correct in the transparency treatment. The figure falls to 29.7% in the nontransparency treatment. The difference across the two treatments is statistically significant according to a Mann-Whitney test (one-sided p-value is 0.008). We conclude that transparency improves the accuracy of decision makers in state L, in line with Hypothesis 1 (a).

The decision maker accuracy in state L in the non-transparency treatment is much larger than the theoretical prediction of zero mainly due to two reasons. As the next section shows, there are many advisers who recommend the correct policy in state L rejecting the payment even in the nontransparency treatment, likely due to lying-aversion. Decision makers who follow such advisers make the correct decision in state L.[28]

Next, we study the accuracy of decisions in state R. Recall that Hypothesis 1 (b) predicts that the decision maker is worse-off in state R with transparency but this is only because of the mixed strategy equilibrium (see Table 1). Table 2 shows that 75.6% of decision makers choose policy r if the state is R in the transparency treatment. The percentage in the nontransparency treatment is 79.3%. This difference between the two treatments is not statistically significant—p-value from a one-sided Mann-Whitney test is 0.169. Thus, there is no statistical support for Hypothesis 1 (b).

Overall, our answer to Question 1 in Section 5 is affirmative as transparency makes decision makers better off in state L and has no effect in state R.

Table 3: Random Effects Probit Estimations of Accuracy in the Transparency Treatment

Coefficients	**State L**	**State R**
Period # (complete data)	-0.0098	0.0101
	(0.0074)	(0.0069)
Period # (data in the first half)	-0.0222	0.0003
	(0.0223)	(0.0105)
Period # (data in the second half)	-0.0309**	.0214
	(0.0140)	(0.0195)

Notes: (1) Standard errors are clustered by matching group. (2) *** p<0.01, ** p<0.05, * p<0.1

Figure 1 displays the proportion of correct choices in state L by period and treatment. The figure shows that there is no discernible time trend in either treatment. We obtain the same conclusion from a random effects panel probit regression of the probability a decision maker chooses l in state L with the round number as the independent variable. We present

[28] In the nontransparency treatment, there are also some decision makers who choose policy l even if the adviser recommends r, and that happens to be the correct decision if the state realization is L (see Section 6.3).

Figure 1: Decision maker Accuracy in State L

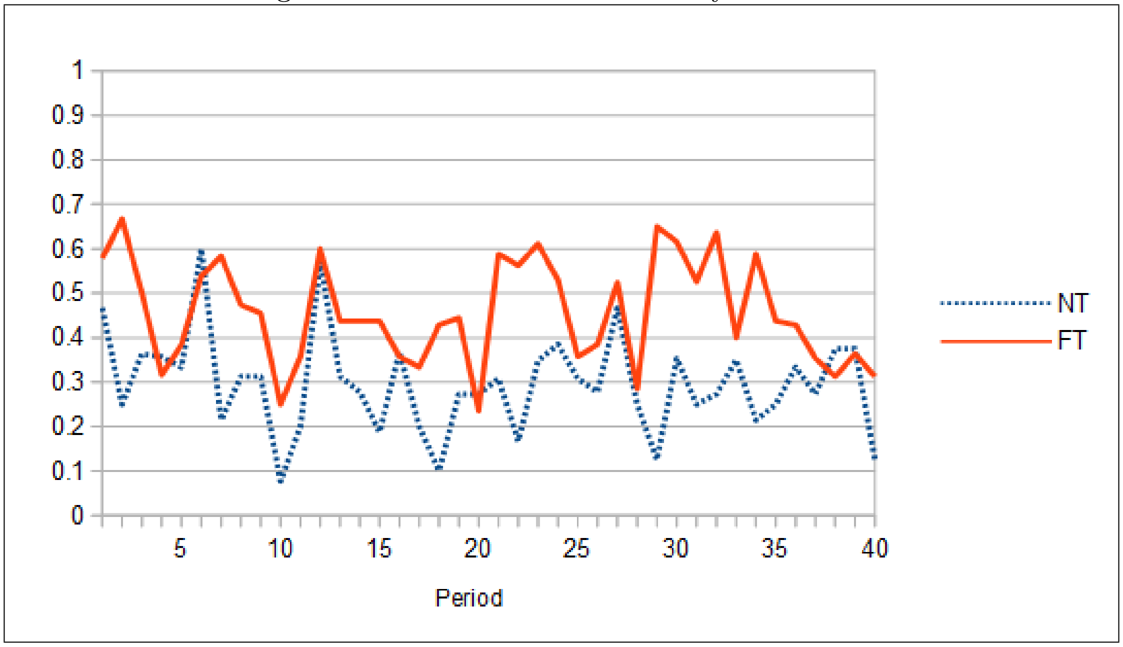
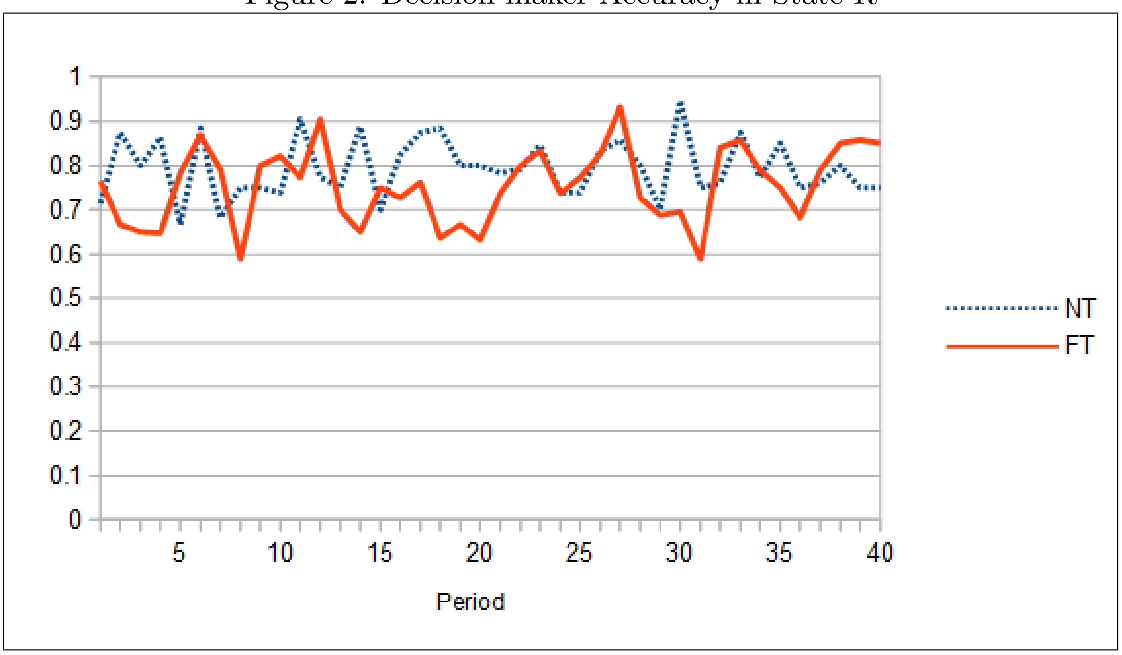


Figure 2: Decision maker Accuracy in State R



the regression result for the transparency treatment in Table 3. The coefficient of the trend variable is insignificant. We also run the regressions using the data before role switching and after role switching. The coefficient of the trend variable is significant only in the second half, which suggests that accuracy of decisions made in state L declines over time in the second half of the transparency treatment. Figure 2 displays the proportion of correct choices in state R by period and treatment. Just like in state L, there is no discernible time trend in either treatment. We obtain the same conclusion from a random effects panel probit regression of the probability a decision maker chooses r in state R as a function of the round number. The regression result for the transparency treatment is given in Table 3. The coefficient of the trend variable is insignificant. The coefficient of the trend variable is also insignificant when we run separate regressions using the data before role switching and after role switching.[29]

Next, we have a detailed look at subjects' behavior in order to answer Questions 2 and 3.

6.2 Adviser Behavior

We start by analyzing the percentage of adviser mistakes in the data. Given our equilibrium specification, there are two mistakes that A can make: (i) rejecting the payment and recommending r in state L; and (ii) recommending l in state R.[30] The percentage of such mistakes is quite low and do not differ significantly across treatments, as Table 4 shows.

Table 4: Advisor Mistakes by Treatment and State

	Transparency	**Non-transparency**	**Mann-Whitney p–value**
State L	2.3% (14 out of 620)	1.4% (8 out of 586)	0.459 (two-sided test)
State R	3.3% (27 out of 820)	5.7% (49 out of 854)	0.198 (two-sided test)

We now investigate whether transparency affects A's behavior. Part (a) of Hypothesis 2 predicts that transparency reduces the fraction of advisers who accept the side payment

[29] Regressions for the non-transparency treatment show that there is no time trend in neither the first half, nor the second half (nor the complete data) in either state.

[30] We do not interpret rejecting the payment and recommending r in state R in the non-transparency treatment as a mistake, as this may be due to amoral conviction. Nevertheless, the proportion of advisors who reject the payment in state R in the non-transparency treatment is very low, as we report in the main text and Table 7.

and part (b) predicts that transparency increases the fraction of advisers who reject the payment and recommend the correct policy in both states.

We start by analyzing A's behavior in state L. Table 5 shows that transparency reduces the fraction of advisers who accept the payment in state L as predicted: 76.3% of advisers accept the payment in state L in the nontransparency treatment and transparency reduces this figure to 67.4%. This difference across the two treatments is statistically significant according to a Mann-Whitney (one-sided p-value $= 0.055$). A closer look at the data reveals that role switching had a clear impact on behavior and that transparency effect is stronger in the first half. Table 6 shows that, in the first half, 77.7% of advisers accept the payment in state L in the nontransparency treatment and transparency reduces this figure to 66.3%. This difference is statistically significant at 5% level according to a Mann-Whitney test (one-sided p-value $= 0.0125$). In the second half, transparency still reduces the fraction of advisers who accept the payment in state L (74.7% in the nontransparency treatment versus 68.4% with transparency); however, the effect is not statistically significant.

Table 5: Percentage of Advisors who accept the payment

	Transparency	**Non-transparency**	**Mann-Whitney p–value**
State L	67.42%	76.28%	0.055 (one-sided test)
State R	81.59%	91.57%	0.215 (one-sided test)

Recall that advisers make mistakes in both treatments, as reported in Table 4. As a result, in both treatments, the fraction of advisers who recommend the correct policy is slightly lower than the fraction of advisers who reject the payment. Combining Tables 4 and 5, we can see that the fraction of advisers who reject the payment and recommend l in state L is 30.3% in the transparency treatment and 22.3% in the nontransparency treatment.[31] This difference is statistically significant according to a Mann-Whitney test at 10% level (one-sided p-value $= 0.075$). If we consider the first half and the second half of the data separately, we again find that the effect of transparency is stronger before role switching. Table 7 shows that in the first half, 21.68% of the advisers reject the payment and recommend l in the nontransparency treatment, and transparency increases this to 31.3%. The difference is

[31] The fraction of advisors who recommend l in state L is equal to 1 minus the fraction that accepts the payment in state L (67.42% from Table 5) minus the fraction that rejects the payment but recommends the wrong policy in state L (2.3% from Table 4). Thus, $30.28\% = 1 - 67.42\% - 2.3\%$. Similarly, $22.32\% = 1 - 76.28\% - 1.4\%$ in the nontransparency treatment.

statistically significant (one-sided p-value $= 0.027$ according to a Mann-Whitney test). Once again, transparency has no statistically significant impact in the second half.

Table 6: Percentage of Advisors who accept the payment (before Role Switching)

	Transparency	**Non-transparency**	**Mann-Whitney p–value**
State L	66.32%	77.97%	0.0125 (one-sided test)
State R	77.70%	91.71%	0.189 (one-sided test)

We now analyze adviser behavior in state R. Table 5 shows that transparency reduces the proportion of advisers who accept the payment in state R, as we predicted: 91.6% of advisers accept the payment in state R in the nontransparency treatment and transparency reduces this to 81.6%. This difference is not statistically significant according to a Mann-Whitney test.

Table 7: Percentage of Advisor Behavior Consistent with Honest Equilibrium (before Role Switching)

	Transparency	**Non-transparency**	**Mann-Whitney p–value**
State L	31.3%	21.68%	0.027 (one-sided test)
State R	19.48%	2.76%	0.017 (one-sided test)

However, the effect of transparency in state R is statistically significant if we compare the fraction of advisers who choose r_R—i.e., the fraction of advisers who reject the payment *and* recommend the correct policy in state R.[32] The fraction of advisers who choose r_R in state R can be derived from Tables 3 and 4. On average, 2.6% of advisers choose r_R in state R without transparency. This increases to 15% with transparency. The difference across the two treatments is significant according to a Mann Whitney test (one-sided p-value $= 0.036$). The effect is even stronger in the first half of the experiment, as Table 7 shows. On average, 19.5% of advisers reject the payment and recommend r in state R with transparency whereas only 2.8% do so without transparency (one-sided p-value $= 0.017$).

From these findings, we conclude that transparency reduces the percentage of advisers who accept the payment—in line with Hypothesis 2 part (a)—and increases the percentage

[32] Recall that adviserssometimes do mistakes; an adviser who rejects the payment in state R sometimes fails to recommend the correct policy.

Figure 3: Percentage of Advisors who recommend l in state L

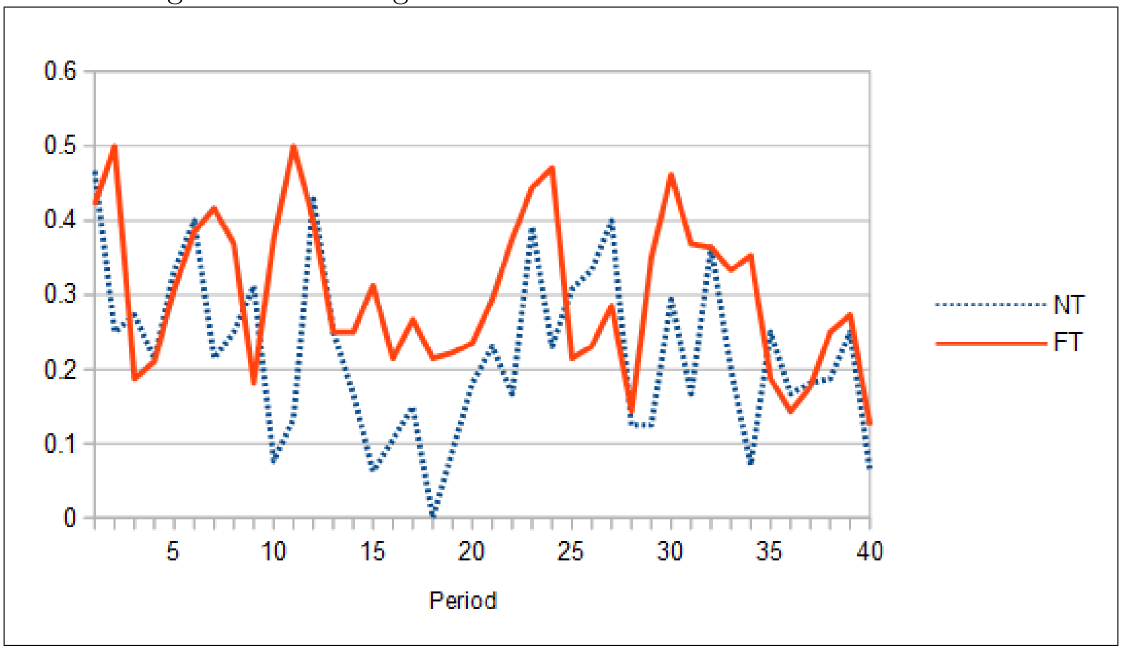
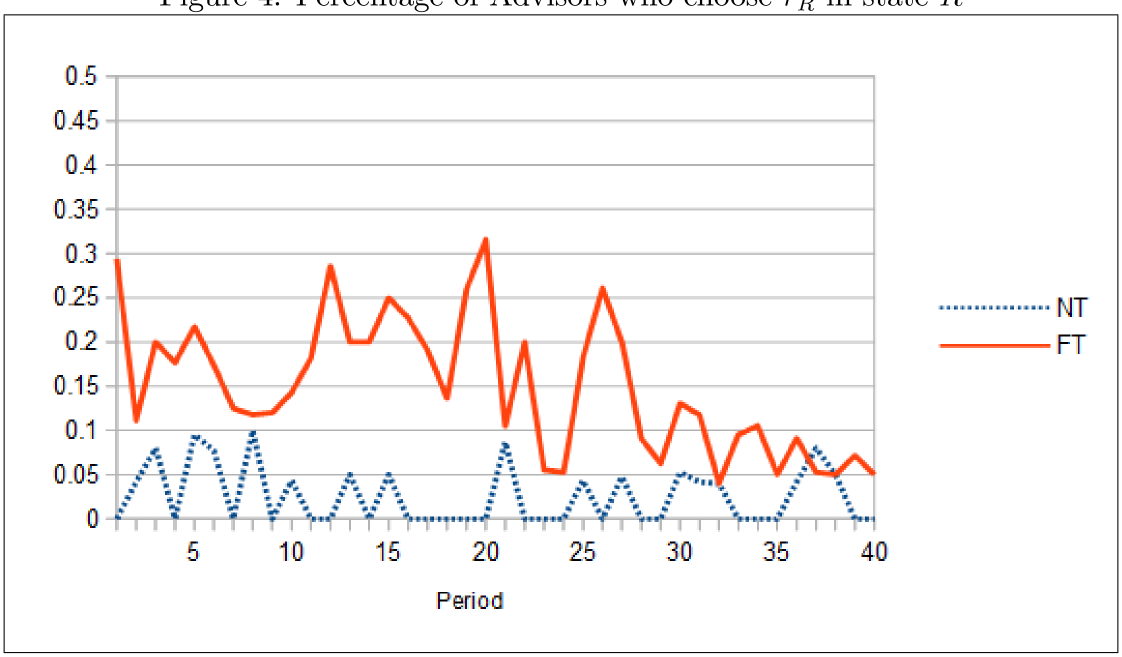


Figure 4: Percentage of Advisors who choose r_R in state R



who rejects the payment and recommends the correct policy in both states—in line with part (b).

Thus, our answer to Question 2 is also affirmative: Transparency does affect advisers' behavior and makes them more likely to reject the payment and follow the honest equilibrium strategy. One caveat is that the effects of transparency on adviser behavior weaken over time, as already discussed. Figures 3 and 4 display the evolution of adviser behavior consistent with the honest equilibrium in state L and state R, respectively. Both figures reveal a downward trend in the transparency treatment.

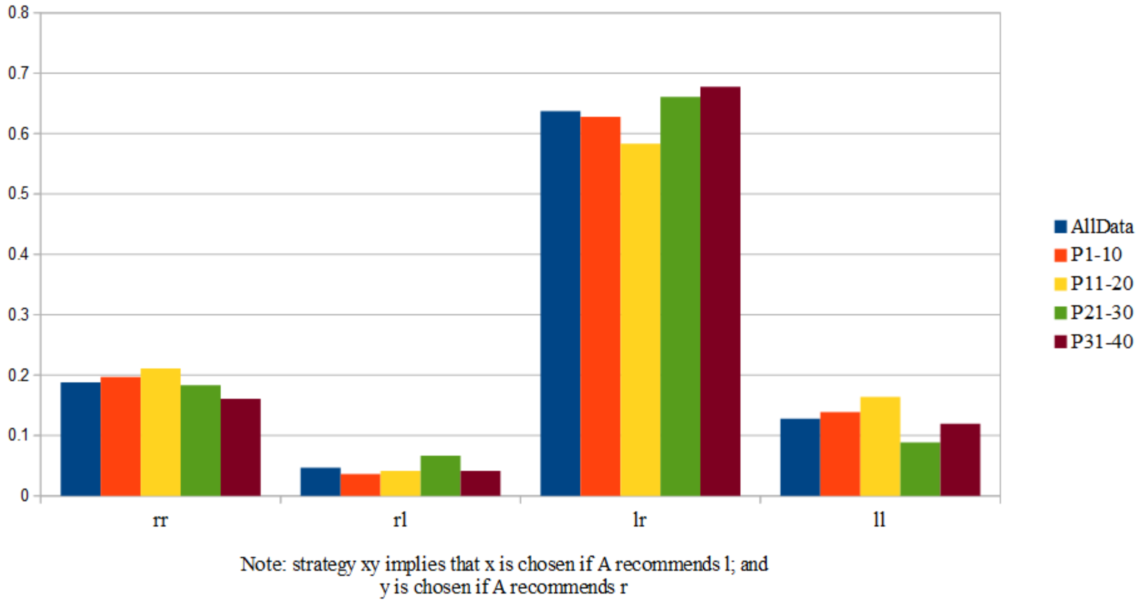
What are the plausible explanations for this observation? As we will discuss in the next section in more detail, a large fraction of decision makers choose to follow the recommendation of the adviser *even if* the adviser accepts the payment from the SIG. This is likely to reduce over time the proportion of advisers who behave in accordance with the honest equilibrium requirements. We will elaborate on this idea in the next section. Moreover, role switching may have had a negative influence on the decision makers of the first half of the experiment, who became advisers in the second half. For instance, decision makers in the first half of the transparency treatment were recommended the correct policy less than 1/3 of the time in state L, as Table 7 shows. This type of behavior may have resulted in a resentment among decision makers and facilitated the pervasiveness of the corrupt equilibrium behavior in the second half when the decision makers in the first half became advisers to their previous advisers.

6.3 Decision maker Behavior

As explained in Section 5, we used the strategy method with decision makers in order to elicit their full strategy. This method enables us to have more observations from decision makers at every information set and analyze better whether decision makers use equilibrium strategies.

We start with the non-transparency treatment. The corrupt equilibrium is the unique equilibrium in this treatment. To be more precise, the equilibrium is unique on the equilibrium path but there are different off-the-equilibrium path beliefs that support the same equilibrium, none of which can be eliminated using IC or D1. Thus, out of four possible pure strategies, two are consistent with the corrupt equilibrium: (i) always choose r; and (ii) choose r if $m = r$ and l if $m = l$. The strategy "choose r if $m = r$ and l if $m = l$" is by far the most popular strategy—63.8% of the decision makers use it. Figure 5 shows the

Figure 5: Decision maker Strategies in the Nontransparency Treatment



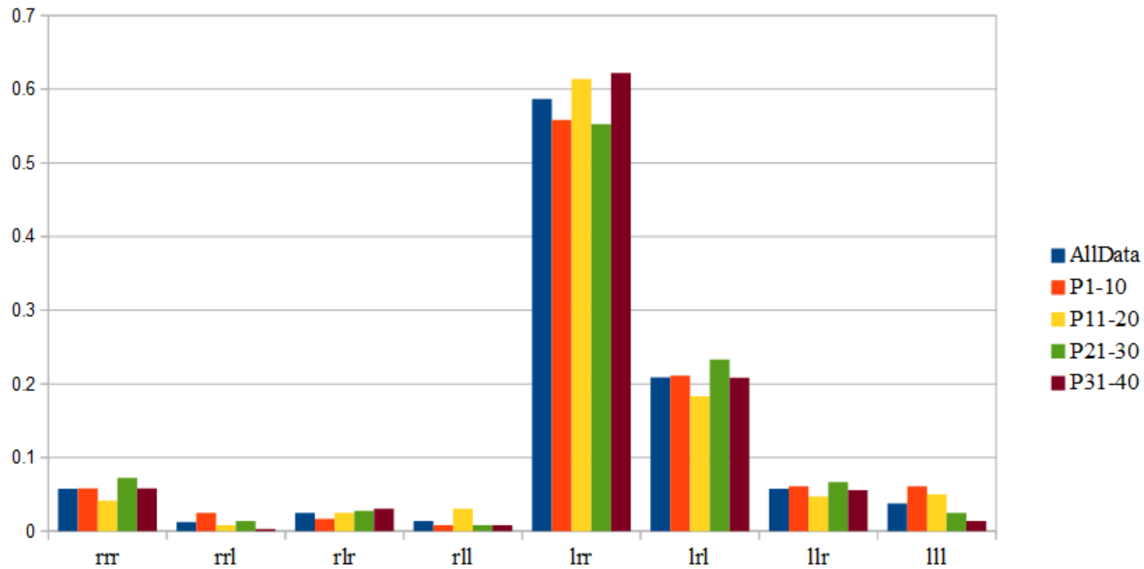
distribution of decision maker strategies in the transparency treatment. In total, 82.6% of the decision makers use strategies that are consistent with the corrupt equilibrium in the nontransparency treatment.

We next find the best response of the decision maker in the nontransparency treatment given the empirical distribution of the adviser recommendations. This is equivalent to finding the empirical posterior probability of state R given that (i) the adviser recommends r; and (ii) the adviser recommends l. We find that the empirical best response of the decision maker is "choose r if $m = r$ and l if $m = l$," this is indeed the most frequently observed strategy in the data.

Figure 6 shows the distribution of decision maker strategies in the transparency treatment. Of the eight possible decision maker strategies, "choose r if $m \in \{r_A, r_R\}$ and l if $m = l$" is the most popular one—about 59% of decision makers use this strategy. This strategy is consistent with the corrupt equilibrium. More generally, there are four decision maker strategies that are consistent with the corrupt equilibrium in the transparency treatment: (i) always r; (ii) choose r if $m \in \{r_A, r_R\}$ and l if $m = l$; (iii) choose r if $m \in \{r_A, l\}$ and l if $m = r_R$; and (iv) choose r if $m = r_A$ and l if $m \in \{l, r_R\}$.[33] In total, 72.7% of the decision makers use strategies that are consistent with the corrupt equilibrium.

[33] The strategies except for the second one may seem unreasonable; however, there is no standard refinement that could rule out such strategies.

Figure 6: Decision maker Strategies in the Transparency Treatment



Note: strategy xyz implies that x is chosen if A recommends l;
y is chosen if A rejects payment and recommends r; and
z is chosen if A accepts payment and recommends r

There is only one strategy that is consistent with the honest equilibrium: choose r if $m = r_R$ and l if $m \in \{r_A, l\}$. As we discussed before, the honest equilibrium requires a bias against those advisers who accept the payment. Indeed, we observe that a non-negligible fraction of decision makers are suspicious of advisers who accepted the side payment in the transparency treatments. 21% of decision makers use the honest equilibrium strategy; it is the second most popular strategy after the corrupt equilibrium strategy "choose r if $m \in \{r_A, r_R\}$ and l if $m = l$."

We next find which strategy is the optimal decision maker strategy in the transparency treatment given the empirical distribution of adviser behavior. To that aim, we compute the empirical posterior probability of state R given that (i) A accepts the payment; (ii) A rejects the payment and recommends r; and (iii) A recommends l. It turns out that the optimal decision maker strategy is the corrupt equilibrium strategy, "choose r if $m \in \{r_A, r_R\}$ and l if $m = l$."

This finding might seemingly imply that it is suboptimal for decision makers to use the honest equilibrium strategy but there is a caveat. The bias against advisers who accept the payment and the awareness of the advisers with respect to such a bias are likely the reasons why decision makers are better off with transparency. Arguably, the behavior of

Table 8: Random Effects Probit Estimations of Accepting SIG Payment

	Coefficients	
Constant and Independent Variables	**Model 1**	**Model 2**
FA	0.109 (0.0114)***	
NFA	-0.0708 (0.0372)*	
FR	-0.0971 (0.0307)***	
$Accept_{i,t-1}$	0.156 (0.121)	0.0481 (0.264)
$[P(F\|A)]$		1.195 (0.386)***
$[P(F\|R)]$		-0.735 (0.298)**
t	0.00837 (0.00879)	0.0140 (0.00912)
Constant	0.384 (0.151)**	0.260 (0.584)

Notes: (1) The dependent variable is the adviser's binary choice between accepting the payment (= 1) and rejecting (= 0). (2) The independent variables in Models 1 and 2 are defined in the main text. (3) Robust standard errors are given in parentheses. (4) * (**; ***) indicates significance at the 10% (5%; 1%) level. (5) Errors are clustered at the session level.

advisers is endogenous and shaped by the strategy of decision makers over the course of the experiment. Put differently, had the decision makers "always followed" the advisers, then we might have obtained a very different empirical distribution of adviser choices and perhaps transparency would not have benefited decision makers.

We now follow up on this line of reasoning and do a regression analysis in order to see whether (i) advisers who were previously (not) followed by the decision maker *due to* rejecting (accepting) the side payment are less likely to accept the payment; and (ii) advisers who were previously (not) followed by the decision maker *despite* accepting (rejecting) the side payment are more likely to accept the payment. We run a random effects panel probit regression of the probability A accepts the payment in the transparency treatment at time t (denoted by $Accept_t$) as a function of (i) the number of times A accepted the payment and was followed until t (denoted by FA); (ii) the number of times A accepted the payment and was *not* followed until t (denoted by NFA); (iii) the number of times A rejected the payment and was followed until t (denoted by FR); (iv) the lagged dependent variable; and (v) the round number (t). Thus, the panel model is given by

$$Accept_{i,t} = \beta_0 + \beta_1 FA_{i,t-1} + \beta_2 NFA_{i,t-1} + \beta_3 FR_{i,t-1} + \beta_4 Accept_{i,t-1} + \beta_5 t + \varepsilon_{i,t} + \gamma_i > 0.$$ [34]

We predict that $\beta_1 > 0$, $\beta_2 < 0$, and that $\beta_3 < 0$. The results of this regression are given in Table 8 under the column titled Model 1. The coefficient signs are as we predicted and the coefficients are significant. Model 2 is a variation on the same theme and involves the explanatory variables (i) the fraction of times the adviser was followed conditional on accepting the payment until t (denoted by $[P(F|A)]$); (ii) the fraction of times he was followed conditional on rejecting the payment until t (denoted by $Pr(F|R)$); (iii) the lagged dependent variable; and (iv) the round number (t). Thus, the panel model is given by

$$Accept_{i,t} = \beta_0 + \beta_1 \left[P(F|A)\right]_{i,t-1} + \beta_2 \left[P(F|R)\right]_{i,t-1} + \beta_3 Accept_{i,t-1} + \beta_4 t + \varepsilon_{i,t} + \gamma_i.$$

We predict that $\beta_1 > 0$ and $\beta_2 < 0$. The results of this regression are given in Table 8 under the column titled Model 2. The coefficient signs are as we predicted and the coefficients are significant.

These regression results support our notion that adviser behavior is endogenous and is shaped by the behavior of decision makers over the course of the experiment. They also suggest that the empirical percentage of decision makers who are biased against advisers that accept the third-party payment is not sufficiently high, and therefore, the percentage of adviser behavior consistent with the honest equilibrium declines over time.

7 Voluntary-transparency Condition

Finally, we provide a brief discussion of the theoretical and experimental results in the voluntary-transparency condition. Equilibria exist both with and without disclosure in the voluntary-transparency condition.[35] Importantly, the main results of interest are identical to the transparency condition: The corrupt equilibrium exists as well as the honest equilibrium (provided that $\alpha \geq \beta(R)$ and $\alpha + \gamma \geq \beta(L)$ as in the transparency condition). In fact, there

[34] To be more precise, the right hand side specifies the underlying latent propensity that $Accept_{i,t} = 1$.

[35] Due to the substantial increase in complexity of the voluntary-transparency condition, and the fact that the important intuition regarding equilibrium behaviour is apparent from the analysis of the transparency condition, we omit a formal exposition of the voluntary-transparency condition. The equilibrium analysis and theoretical results are available upon request.

are various corrupt equilibria and various honest equilibria because equilibrium multiplicity in this condition is more severe than in the transparency condition. However, the main features of a corrupt equilibrium and honest equilibrium are exactly the same as before. On the equilibrium path of every corrupt equilibrium, A accepts the payment in both states and is followed by D. On the equilibrium path of every honest equilibrium, A rejects the payment in both states, recommends the better policy for D and is followed.[36]

Next, we present the data analysis. In a nutshell, while we find clear evidence that transparency improves decision making relative to non-transparency, the evidence regarding the effect of voluntary transparency is much weaker. This seems to be because the adviser behavior with voluntary transparency falls somewhere in between the behavior with transparency and the behavior with non-transparency. Put differently, the positive effects of transparency on the adviser behavior are diluted if the decision to reveal is up to the adviser.

We now present a more detailed overview of the data. To distinguish between the three treatments, we denote the transparency treatment by T-treatment, the voluntary transparency treatment by VT-treatment and the nontransparency treatment by NT-treatment. We start by analyzing the accuracy of decisions in VT-treatment and compare it with the other treatments.

Table 9: Percentage of Correct Decisions by State and Treatment

	Transparency	**Non-transparency**	**Voluntary Transparency**
State L	46.13%	29.69%	39.55%
State R	75.61%	79.27%	75.73%

Table 9 extends Table 2 by adding the data from the VT-treatment. According to the table, accuracy in state L is 9.9% higher in the VT-treatment than in the NT-treatment. The difference is statistically significant according to a one-sided Mann-Whitney test but only at the 10% level. T-treatment is not statistically different from VT-treatment in terms of accuracy in state L. However, T-treatment generates a higher accuracy than NT-treatment in state L, as we explained in Section 6.1 (p-value according to a one-sided Mann-Whitney

[36] Corrupt and honest equilibria differ in the way advisers choose (not) to disclose their decision. For example, A may choose not to reveal his/her decision regarding the side payment in state L but this is still part of an honest equilibrium if A rejects the payment and recommends l in state L. This is theoretically equivalent to revealing that A rejected the side payment because A can recommend l only if it is rejected.

test is 0.008). In state R, there is no statistically significant difference between any two treatments.

We now describe adviser behavior in the VT-treatment. We start with the decision to reveal. We find that a majority of advisers choose to reveal their decision regarding the third-party payment: 68.6% chose to reveal in state L and 81.4% chose to do so in state R. These numbers have a slight upward trend over time. It seems that many advisers chose to reveal their decision regarding the payment because decision makers have a bias against advisors who do not reveal. A vast majority (about 86%) of advisers who choose not to reveal accept the payment and recommend r. As reported in the previous section, more than 82% of decision makers followed a recommendation of r from the adviser in the NT-treatment. The figure falls down to 56.1% if the adviser chooses not to reveal and recommends r in the VT-treatment. As a result, many advisers choose to reveal *even if* they will accept the third-party payment.

Table 10: Percentage of Advisors who accept the payment (before Role Switching)

	Transparency	**Non-transparency**	**Voluntary Transparency (% if A chose to reveal)**
State L	66.32%	77.97%	75.27% (69.83%)
State R	77.70%	91.71%	79.18% (77.75%)

Next, we report our findings on the adviser choice with respect to the third-party payment. Just as in the T-treatment, the percentage of advisers who accepted the payment increased and the percentage of advisers who use the honest equilibrium strategy decreased from the first half to the second half in the VT-treatment. Therefore, we focus on the first half of the data as in Section 6.2. Table 10 extends Table 6 by adding the data from the VT-treatment and presents the percentage of advisers who accepted the payment in all three treatments. We first compare the outcome in state L across treatments. The percentage of advisers who accepted the payment is not statistically different across the VT- and NT-treatments. The same is also true when we compare T-treatment with VT-treatment. However, fewer advisers accept the payment in the T-treatment than in the NT-treatment in state L as we discussed in Section 6.2 (p-value = 0.0125). As for state R, we find no statistical evidence that any two treatments differ in terms of the percentage of advisers who accepted the payment in state R.

What about the percentage of adviser behavior consistent with the honest equilibrium? Table 11 extends Table 7 by adding the data from the VT-treatment. The percentage of adviser behavior consistent with the honest equilibrium in state R is higher in the VT-treatment than in the NT-treatment (however, this is significant at the 10% level) and does not differ across the T- and VT-Treatments. As we discussed in Section 6.2, the percentage of adviser behavior consistent with the honest equilibrium in state R is higher in the T-treatment than in the NT-treatment, which is significant at the 5% level. As for state L, the percentage of adviser behavior consistent with the honest equilibrium does not differ across VT- and NT-treatments and across VT- and T-treatments. However, the percentage in the T-treatment is higher than in the NT-treatment (at the 5% level as we discussed in Section 6.2).

Table 11: Percentage of Advisor Behavior Consistent with Honest Equilibrium (before Role Switching)

	Transparency	**Non-transparency**	**Voluntary Transparency (% if A chose to reveal)**
State L	31.3%	21.68%	22.97% (29.37%)
State R	19.48%	2.76%	18.31% (21.39%)

Overall, these findings suggest that the form of transparency can matter: We do not have sufficient evidence to claim that decision makers are better off in the VT-treatment than in the NT-treatment; however, T-treatment improved decision-making and altered adviser behavior relative to NT-treatment, as we showed in Sections 6.1 and 6.2. This is likely because the adviser behavior in the VT-treatment seems to be a weighted average of the adviser behavior in the T-treatment and the behavior in the NT-treatment. As Tables 10 and 11 show, if the adviser chooses to reveal his decision regarding the third-party payment then the adviser behavior is quite similar to that in the T-treatment. If however the adviser chooses not to reveal, then the adviser behavior is similar to that in the NT-treatment. But a nontrivial percentage of advisers choose not to reveal—especially in state L in which almost 1/3 of advisers do not reveal. As a result, the positive effects of transparency are weakened if the disclosure decision is up to the adviser.

8 Conclusion

Our study sheds light on the effects of transparency on strategic information transmission and decision making in a setting where the advisor might be swayed by a third-party who favors the implementation of a specific policy. Out theoretical results show that transparency is never harmful and may improve the accuracy of decision makers. Experimental results show that decision makers are better off in state L (this is the state in which the adviser has an incentive to lie), and transparency has no effect in state R. With transparency, more advisers reject the side-payment and recommend the correct policy in both states. There are two caveats, though. First, we find that the form of transparency can matter: While transparency clearly improves decision making (relative to non-transparency) the evidence regarding the effect of voluntary transparency is weaker. Second, positive effects of transparency decline over time because if the adviser accepts the side-payment in the transparency treatment, many decision makers ignore it and follow the adviser.

Our analysis has relevant implications for the debate on transparency. Our paper documents positive effects of transparency whereas prior research produced bleak results. Prior research modeled conflict of interest as being exogenous. Our results imply that the debate on transparency should take into account whether the conflict of interest between the adviser and the decision maker should be modeled as being exogenous or endogenous. We believe that the latter is the appropriate approach as experts have in most cases the agency to accept or reject side-payments, gifts and bribes.

8.1 References

Banks, Jeffrey S., and Joel Sobel. (1987). "Equilibrium Selection in Signaling Games," *Econometrica*, 55(3), 647-661.

Cain, Daylian M., Loewenstein, George, and Don A. Moore. (2005). "The Dirt on Coming Clean: Perverse Effects of Disclosing Conflicts of Interest," *Journal of Legal Studies*, 34(1), 1–25.

Cain, Daylian M., Loewenstein, George, and Don A. Moore. (2011). "When Sunlight Fails to Disinfect: Understanding the Perverse Effects of Disclosing Conflicts of Interest," *Journal of Consumer Research*, 37(5), 836-857.

Cho, In-Koo, and David M. Kreps (1987). "Signaling Games and Stable Equilibria," *Quarterly Journal of Economics*, 102(2), 179-222.

Crawford, Vincent P., and Joel Sobel. (1982). "Strategic Information Transmission," *Econometrica*, 50, 1431-1451.

Delafontaine, Ramses. (2015). *Historians as expert judicial witnesses in Tobacco litigation: a controversial legal practice*. Springer.

Durbin, Erik, and Ganesh Iyer. (2009). "Corruptible Advice," *American Economic Journal: Microeconomics*, 1(2), 220-242.

Fischbacher, Urs. (2007). "z-Tree: Zurich toolbox for ready-made economic experiments," *Experimental Economics,* 10(2), 171-178.

Gibson, Rajna, Carmen Tanner, and Alexander F. Wagner. (2013). "Preferences for Truthfulness: Heterogeneity among and within Individuals: Dataset," *American Economic Review*, 103(1), 532–548.

Gneezy, Uri. (2005). "Deception: The Role of Consequences," *American Economic Review*, 95(1), 384–94.

Hurkens, Sjaak, and Navin Kartik. (2009). "Would I Lie to You? On Social Preferences and Lying Aversion," *Experimental Economics*, 12(2), 180–92.

Inderst, Roman, and Marco Ottaviani. (2009). "Misselling through Agents," *American Economic Review*, 99(3), 883-908.

Inderst, Roman, and Marco Ottaviani. (2012). "Competition through commissions and kickbacks," *American Economic Review*, 102(2), 780-809.

Ismayilov, Huseyn, and Jan Potters. (2013). "Disclosing adviser's interests neither hurts nor helps," *Journal of Economic Behavior and Organization*, 93, 314-320.

Loewenstein, George, Cain, Daylian M., and Sah, Sunita. (2011). "The limits of transparency: Pitfalls and potential of disclosing conflicts of interest," *American Economic Review*, 101(3), 423-428.

Loewenstein, George, Sunita Sah, and Daylian M. Cain. (2012) "The unintended consequences of conflict of interest disclosure," *Journal of the American Medical Association*, 307(7), 669-670.

Koch, Christopher, and Carsten Schmidt. (2010). "Disclosing conflicts of interest–Do experience and reputation matter?," *Accounting, Organizations and Society*, 35(1), 95-107.

Potters, Jan, and Frans Van Winden. (1992). "Lobbying and asymmetric information," *Public Choice*, 74(3), 269-292.

Potters, Jan, and Frans Van Winden. (2000). "Professionals and students in a lobbying experiment: Professional rules of conduct and subject surrogacy," *Journal of Economic Behavior & Organization*, 43(4), 499-522.

Proctor, Robert. (2011). *Golden holocaust: origins of the cigarette catastrophe and the case for abolition.* University of California Press.

Rode, Julian. (2010). "Truth and trust in communication: experiments on the effect of a competitive context," *Games and Economic Behavior*, 68(1), 325–338

Sah, Sunita, and George Loewenstein. (2013). "Nothing to Declare: Mandatory and Voluntary Disclosure Leads Advisers to Avoid Conflicts of Interest," *Psychological Science*, 25(2), 575-584.

Sánchez-Pagés, Santiago, and Marc Vorsatz. (2007). "An Experimental Study of Truth-Telling in a Sender-Receiver Game," *Games and Economic Behavior*, 61(1), 86–112.

Sobel, Joel. (1985). "A Theory of Credibility," *Review of Economic Studies*, 52(4), 557-573.

Stossel, Thomas P. (2005). "Regulating Academic–Industrial Research Relationships — Solving Problems or Stifling Progress?" *The New England Journal of Medicine*, 353(10), 1060-1065.

Stossel, Thomas P., and Lance K. Stell. (2011). "Time to 'walk the walk' about Industry Ties to enhance Health," *Nature Medicine*, 17, 437–438.

Sutter, Matthias. (2009). "Deception through Telling the Truth?! Experimental Evidence from Individuals and Teams," *Economic Journal*, 119, 47–60.

Weber, Michael A. (2009). "Academic Physicians Confront a Hostile World: The Creation of ACRE," *The Journal of Clinical Hypertension* 11(10), 533–536.

Yale Project on Climate Change Communication. 2013. "Scientific and Public Perspectives on Climate Change." http://environment.yale.edu/climate-communication/files/ClimateNote_Consensus_Gap_May2013_FINAL6.pdf

ONLINE APPENDIX

Proof of Proposition 3. To see why only a corrupt equilibrium exists, first note that at least one message chosen by A in equilibrium results in D choosing r with probability one; i.e., there exists $m \in \{l, r\}$ where $\Pr_A(m|S) > 0$ for some $S \in \{L, R\}$ such that $Pr_D(r|m) = 1$. Suppose not. Then, $Pr_D(r|m) < 1$ for any $m \in \{l, r\}$ where $\Pr_A(m|S) > 0$. This can be true only if l and r are both chosen in equilibrium with positive probability and $\mu(R|m) \leq 0.5$ for $m \in \{l, r\}$.[37] But $\mu(R|r) \leq 0.5$ and $\mu(R|l) \leq 0.5$ together result in an inconsistency. Since

$$\mu(R|m) = \frac{\Pr_A(m|R)(1-p)}{\Pr_A(m|R)(1-p) + \Pr_A(m|L)p} \leq 0.5,$$

it follows that

$$\Pr_A(m|L) \geq \frac{1-p}{p} \Pr_A(m|R)$$

for $m \in \{r, l\}$. But, $\Pr_A(r|S) + \Pr_A(l|S) = 1$ for $S \in \{L, R\}$. Then,

$$\Pr_A(r|L) + \Pr_A(l|L) = 1 \geq \frac{1-p}{p}(\Pr_A(r|R) + \Pr_A(l|R)) > 1,$$

since $p < 0.5$, a contradiction. Hence, there exists $m \in \{l, r\}$ such that $Pr_A(m|S) > 0$ for some $S \in \{L, R\}$, and $Pr_D(r|m) = 1$. But this message can only be r. Suppose towards a contradiction that it is l; i.e., $Pr_A(l|S) > 0$ for some $S \in \{L, R\}$ and $\Pr_D(r|l) = 1$. The latter requires $\mu(R|l) \geq 0.5$. But if the state is R, then A strictly prefers choosing r_A (thus, $m = r$) rather than l, securing a minimum payoff of $\beta(R) + \gamma$ versus 0. Therefore, $\Pr_A(l|R) = 0$ which contradicts $\mu(R|l) \geq 0.5$. Thus, $\Pr_D(r|r) = 1$. Given that $\Pr_D(r|r) = 1$ and $\beta(L) > \gamma$, D always chooses r_A in equilibrium. Hence, we obtain the desired result.[38]

Proposition 4. *See the statement in the main text for the case where $\alpha + \gamma > \beta(L)$. Assume that $\alpha + \gamma < \beta(L)$. There are (at most) three equilibria that satisfy the IC:*

1) *Corrupt Equilibrium:* $\Pr_A(r_A|L) = \Pr_A(r_A|R) = 1$; $\Pr_D(r|r_A) = 1$; $\mu(R|r_A) = 1 - p$.

2) *Perverse Honest Equilibrium (if $\alpha \geq \beta(R)$):* $\Pr_A(r_A|L) = 1$; $\Pr_A(r_R|R) = 1$; $\Pr_D(r|r_R) = 1$; $\Pr_D(l|r_A) = 1$; $\mu(R|r_R) = 1$ and $\mu(L|r_A) = 0$.

[37] Otherwise, $\Pr_A(m|L) = \Pr_A(m|R) = 1$ for either $m = l$ or $m = r$. This, combined with the fact that $p < 0.5$, implies that D must choose r with probability one, a contradiction.

[38] Off-the-equilibrium path belief $\mu(R|l)$ is not uniquely determined; therefore, $\Pr_D(r|l) = 1$ if $\mu(R|l) > 0.5$, $\Pr_D(r|l) \in [0, 1]$ if $\mu(R|l) = 0.5$ and $\Pr_D(r|l) = 0$ if $\mu(R|l) < 0.5$.

3) *Mixed-strategy Equilibrium (if $\alpha \geq \beta(R)$):* $\Pr_A(r_A|L) = 1$; $\Pr_A(r_A|R) = \frac{p}{1-p}$ and $\Pr_A(r_R|R) = \frac{1-2p}{1-p}$; $\Pr_D(r|r_R) = 1$; $\mu(R|r_R) = 1$; $\Pr_D(r|r_A) = 1 - \beta(R)/\alpha$; $\mu(R|r_A) = 0.5$.

Finally, assume that $\alpha + \gamma = \beta(L)$. Then there always exists a corrupt equilibrium; there exists an honest equilibrium, a perverse honest equilibrium and the mixed-strategy equilibrium if $\alpha \geq \beta(R)$; finally, if $\alpha = \beta(R)$, there also exists a continuum of mixed-strategy equilibria in which A randomizes between r_A and l in state L and between r_A and r_R in state R.

Proof. The proof uses a number of lemmas.

Lemma 7 *For any $m \in \{l, r_A, r_R\}$ chosen by the adviser in equilibrium,*

$$\Pr_D(l|m) = \begin{cases} 1 & \text{if } \Pr_A(m|L) > \frac{1-p}{p}\Pr_A(m|R), \\ \in [0,1] & \text{if } \Pr_A(m|L) = \frac{1-p}{p}\Pr_A(m|R), \\ 0 & \text{if } \Pr_A(m|L) < \frac{1-p}{p}\Pr_A(m|R), \end{cases}$$

Proof. The decision maker strictly prefers choosing l if

$$\mu(L|m)(\bar{\pi} - \underline{\pi}) > (1 - \mu(L|m))(\bar{\pi} - \underline{\pi})$$

for $m \in \{l, r_A, r_R\}$. That is, l is strictly preferred if the adviser chooses m and $\mu(L|m) > 0.5$. In equilibrium, $\mu(L|m)$ must accord with Bayes' law; i.e.

$$\mu(L|m) = \frac{p\Pr_A(m|L)}{p\Pr_A(m|L) + (1-p)\Pr_A(m|R)}.$$

This implies that the decision maker chooses l with probability one if

$$\Pr_A(m|L) > \frac{1-p}{p}\Pr_A(m|R).$$

The rest of the proof follows similar lines and is therefore omitted.

Lemma 8 *At least one action chosen by the adviser in equilibrium is followed by r with probability one, i.e., there exists an $m \in \{l, r_A, r_R\}$ with $\Pr_A(m|S) > 0$ for some $S \in \{L, R\}$ in equilibrium such that $\Pr_D(r|m) = 1$.*

Proof. Suppose that $\Pr_D(r|m) < 1$ for every m with $\Pr_A(m|S) > 0$, $S \in \{L, R\}$ in equilibrium. Then by Lemma 7, $\Pr_A(m|L) \geq \frac{1-p}{p}\Pr_A(m|R)\ \forall m \in \{l, r_A, r_R\}$ (if m is not played in equilibrium then the inequality is satisfied trivially). This implies that

$$\begin{aligned}
1 &= \Pr_A(l|L) + \Pr_A(r_A|L) + \Pr_A(r_R|L) \\
&\geq \tfrac{1-p}{p}(\Pr_A(l|R) + \Pr_A(r_A|R) + \Pr_A(r_R|R)) \\
&= \tfrac{1-p}{p} \\
&> 1,
\end{aligned}$$

as $p < 1/2$, a contradiction.

Lemma 9 *In equilibrium, at least one of r_A and r_R is played, then followed by r with probability one.*

Proof. By Lemma 8, there exists an $m \in \{l, r_A, r_R\}$ with $\Pr_A(m|S) > 0$ for some $S \in \{L, R\}$ in equilibrium such that $\Pr_D(r|m) = 1$. Suppose towards a contradiction that it is l in some equilibrium; i.e., $\Pr_A(l|S) > 0$ for some $S \in \{L, R\}$ and $\Pr_D(r|l) = 1$. The adviser's payoff from choosing l in state R is zero, however a payoff of at least $\beta(R) + \gamma$ can be guaranteed by choosing r_A, so $\Pr_A(l|R) = 0$. But if l is being chosen only in state L, then $\mu(L|l) = 1$ and $\Pr_D(r|l) = 0$, a contradiction.

Lemma 10 *In equilibrium, the adviser never chooses l in R. If the adviser chooses l in equilibrium, then it is always followed by l.*

Proof. By the previous Lemma, in state R, the adviser will be able to earn either $\alpha + \gamma$ by deviating to r_R, or $\alpha + \beta(R) + \gamma$ by deviating to r_A, both of which are greater than the highest possible payoff from l (α). Therefore if the adviser chooses l in equilibrium, he chooses it only in state L, so that $\mu(L|l) = 1$, and $\Pr_D(l|l) = 1$.

Lemma 11 *Suppose that $\alpha + \gamma > \beta(L)$. If $\Pr_A(r_A|S) > 0$ for some $S \in \{L, R\}$ in an equilibrium robust to IC, then (i) $\Pr_A(r_A|L) = 1$ and $\Pr_A(r_A|R) > 0$ must hold if $\beta(R) + \gamma \neq \beta(L)$; (ii) $\Pr_A(r_A|L) = \Pr_A(r_A|R) = 1$ must hold if $\beta(R) + \gamma > \beta(L)$; (iii) $\Pr_A(r_A|R) > 0$ and $\Pr_A(r_A|L) > 0$ if $\beta(R) + \gamma = \beta(L)$ (to be more precise, in addition to the type of equilibria spelled out in (i) and (ii) there exists a continuum of mixed-strategy equilibria in which A randomizes between r_A and l in state L and between r_A and r_R in state R).*

Proof. Suppose that r_A is chosen in equilibrium; i.e., $\Pr_A(r_A|S) > 0$ for some $S \in \{L, R\}$.

(1) First, we show that this implies that $\Pr_A(r_A|L) > 0$ must hold. Assume towards a contradiction that r_A is chosen in equilibrium but only in state R; i.e., $\Pr_A(r_A|R) > 0$ and

$\Pr_A(r_A|L) = 0$. Using Bayes' law, $\mu(R|r_A) = 1$ in equilibrium and thus, $\Pr_D(r|r_A) = 1$. But then A deviates to r_A in state L since $\beta(L) > \gamma$, a contradiction. Note that this result actually holds regardless of the assumption that $\alpha + \gamma > \beta(L)$.

(2) Next, we show that if A chooses r_A in equilibrium in some S, then it must be chosen with positive probability in both states. Suppose not. By (1) above, $\Pr_A(r_A|L) > 0$. Then, by hypothesis, $\Pr_A(r_A|R) = 0$ must hold; i.e., A chooses r_A in equilibrium only in state L. But this implies that $\mu(L|r_A) = 1$ using Bayes' law and that $\Pr_D(l|r_A) = 1$. Thus, by Lemma 9, r_R is played in equilibrium and $\Pr_D(r|r_R) = 1$. As a result, the equilibrium payoff in state R is at least $\alpha + \gamma$ and recommending l must be equilibrium dominated in state R. By IC, $\mu(L|l) = 1$ and thus, A can deviate to recommending l in state L and obtain at least $\alpha + \gamma > \beta(L)$, a contradiction. As a result, $\Pr_A(r_A|R) > 0$ must hold if $\Pr_A(r_A|L) > 0$ in equilibrium. This in turn implies that if r_A is chosen in equilibrium, then it must be chosen with positive probability in both states.

(3) We now show that $\Pr_A(r_A|L) = 1$ and $\Pr_A(r_A|R) > 0$ must hold if r_A is chosen in equilibrium and $\beta(R) + \gamma \neq \beta(L)$. Suppose not so that $\Pr_A(r_A|L) \in (0, 1)$ (as implied by (2) above). This means that A randomizes between l and r_A in state L. Note that A cannot randomize between r_R and r_A in state L. Because if A randomizes between r_R and r_A in state L, then by Lemma 9, $\Pr_D(r|r_R) = 1$ in equilibrium—otherwise, A always chooses r_A. It follows that recommending l is equilibrium dominated in R. But then, recommending l is a profitable deviation in state L since the payoff from r_R in state L is only α. Thus, $\Pr_A(r_A|L) \in (0, 1)$ implies that A randomizes between l and r_A in state L. In equilibrium, $\Pr_A(r_R|R) > 0$ must hold, otherwise $\Pr_D(r|r_A) = 1$ by Lemma 9, which is a contradiction. Put differently, A must randomize between r_A and r_R in state R. We next prove that it cannot be the case that A randomizes between r_A and l in state L *and* between r_A and r_R in state R—unless $\beta(R) + \gamma = \beta(L)$, which we discuss below. To see why, note that $\Pr_D(r|r_A)$ is given by

$$\Pr_D(r|r_A)\alpha + \beta(L) = \alpha + \gamma$$

since A is indifferent between r_A and l in state L and by Lemma 10, $\Pr_D(l|l) = 1$. Thus,

$$\Pr_D(r|r_A) = 1 - \frac{\beta(L) - \gamma}{\alpha}$$

in equilibrium. Since $\Pr_D(r|r_A) < 1$, $\Pr_D(r|r_R) = 1$ must hold by Lemma 9. Thus, A is indifferent between r_A and r_R in state R and between l and r_A in state L only if the knife-edge case

$$\beta(R) + \gamma = \beta(L)$$

holds. As a result, if r_A is chosen in equilibrium, then $\Pr_A(r_A|L) = 1$ and $\Pr_A(r_A|R) > 0$ must hold as long as $\beta(R) + \gamma \neq \beta(L)$. This proves part (i). If however $\beta(R) + \gamma = \beta(L)$ holds, then apart from the equilibria in which $\Pr_A(r_A|L) = 1$ and $\Pr_A(r_A|R) > 0$, there is also a continuum of mixed-strategy equilibria such that A randomizes between r_A and l in state L *and* between r_A and r_R in state R. $\Pr_A(r_A|R)/\Pr_A(r_A|L) = p/(1-p)$ and $\Pr_D(r|r_R) = 1$, $\Pr_D(l|l) = 1$ and $\Pr_D(r|r_A) = 1 - (\beta(L) - \gamma)/\alpha$ hold in such equilibria. This is the case in (iii).

(4) Part (3) implies that if r_A is played in equilibrium and $\beta(R) + \gamma \neq \beta(L)$, then this is the part of the corrupt equilibrium where $\Pr_A(r_A|R) = \Pr_A(r_A|L) = 1$, and it can also be part of a mixed-strategy equilibrium where $\Pr_A(r_A|R) \in (0, 1)$ and $\Pr_A(r_R|R) \in (0, 1)$.[39] However, if $\beta(R) + \gamma > \beta(L)$ and r_A is chosen in equilibrium, then $\Pr_A(r_A|L) = \Pr_A(r_A|R) = 1$ must hold; i.e., a mixed-strategy equilibrium in which $\Pr_A(r_A|L) = 1$ and $\Pr_A(r_A|R) \in (0, 1)$ is not robust to IC. To see why first note that if $\Pr_A(r_A|L) = 1$ and $\Pr_A(r_A|R) \in (0, 1)$, then we must have $\Pr_A(r_R|R) \in (0, 1)$ (see Footnote 39). By Lemma 9, $\Pr_D(r|r_R) = 1$ must hold (otherwise, A chooses r_A with probability one in both states). This means that l is equilibrium dominated in state R and since $\Pr_D(r|r_A)\alpha + \beta(L) < \alpha + \gamma$ (by the indifference condition in state R and $\beta(R) + \gamma > \beta(L)$) recommending l is a profitable deviation in state L. This proves part (ii). Note that the mixed-strategy equilibrium is robust to IC if $\beta(R) + \gamma \leq \beta(L)$: Even if l is equilibrium dominated in state R, recommending l is no longer a profitable deviation in state L given that $\beta(R) + \gamma \leq \beta(L)$.

Finally, we can prove the proposition. First, assume that $\alpha + \gamma > \beta(L)$. Again, we use a number of steps.

(1) Given Lemma 11 and Intuitive Criterion, A must choose in state L either l with probability one or r_A with probability one if $\beta(R) + \gamma \neq \beta(L)$; i.e., either $\Pr_A(l|L) = 1$ or $\Pr_A(r_A|L) = 1$ if $\beta(R) + \gamma \neq \beta(L)$. Suppose not. Then, by Lemma 11 $\Pr_A(r_A|L) = \Pr_A(r_A|R) = 0$ must hold in equilibrium. Moreover, $\Pr_A(r_R|L) > 0$ must hold since $\Pr_A(l|L) < 1$ (by contradictory hypothesis) and $\Pr_A(r_A|L) = 0$. Moreover, $\Pr_D(r|r_R) = 1$ by Lemma 9. Thus, $\Pr_A(r_R|R) = 1$. As a result, $\Pr_A(r_R|L) = 1$ must hold because if $\Pr_A(l|L) > 0$, then

[39] To see why $\Pr_A(r_A|L) = 1$ and $\Pr_A(r_A|R) \in (0, 1)$ imply that $\Pr_A(r_R|R) \in (0, 1)$, note that otherwise $\Pr_D(r|r_A) = 1$ by Lemma 9, a contradiction.

$\mu(L|l) = 1$ and A cannot be indifferent between l and r_R in state L. Yet, IC rules out an equilibrium in which $\Pr_A(r_R|L) = \Pr_A(r_R|R) = 1$. A can deviate to l in state L, which is a profitable deviation given the belief $\mu(L|l) = 1$ (because l is equilibrium-dominated in state R). Hence, if $\beta(R)+\gamma \neq \beta(L)$ either $\Pr_A(r_A|L) = 1$ or $\Pr_A(l|L) = 1$ in any equilibrium that satisfies IC.

(2) If $\Pr_A(l|L) = 1$ in an equilibrium, then $\Pr_A(r_R|R) = 1$ must hold—this constitutes the honest equilibrium and it exists provided that $\alpha \geq \beta(R)$ holds. To see why $\Pr_A(l|L) = 1$ implies that $\Pr_A(r_R|R) = 1$, first note that $\Pr_A(r_A|L) = 0$ implies that $\Pr_A(r_A|R) = 0$ by Lemma 11. Moreover, $\Pr_A(l|R) = 0$ must hold. Since $\Pr_A(r_A|R) = 0$ and $\Pr_A(r_A|L) = 0$, it follows that r_R is played and $\Pr_D(r|r_R) = 1$, by Lemma 9. Given this, choosing l in state R is inferior to choosing r_R. Thus, $\Pr_A(r_R|R) = 1$. The honest equilibrium uses the following out-of-equilibrium belief: If the adviser deviates to r_A, then the decision maker holds the belief that $\mu(L|r_A) \geq 0.5$ and chooses l (with sufficiently high probability). This belief system is not ruled out by the Intuitive Criterion because r_A is not equilibrium dominated in L. Since $\alpha+\gamma > \beta(L)$ by assumption, A does not deviate from l in state L. Also, A does not deviate from r_R in state R provided that $\alpha \geq \beta(R)$.

(3) If $\Pr_A(r_A|L) = 1$ in equilibrium, then either
(i) $\Pr_A(r_A|R) = 1$; or
(ii) $\Pr_A(r_A|R) \in (0,1)$ and $\Pr_A(r_A|R) + \Pr_A(r_R|R) = 1$—provided that $\beta(R)+\gamma \leq \beta(L)$.
The former is the corrupt equilibrium. Note that IC puts no restriction on the out-of-equilibrium beliefs; $\mu(R|r_R) \in [0,1]$ and $\mu(R|l) \in [0,1]$. The case in (ii) is the mixed-strategy equilibrium. In this equilibrium, $\Pr_D(r|r_A)$ is derived from the indifference condition

$$\Pr_D(r|r_A)\alpha + \beta(R) + \gamma = \alpha + \gamma.$$

Thus, $\Pr_D(r|r_A) = 1 - \beta(R)/\alpha$. Since $\Pr_D(r|r_A) \in (0,1)$, $\mu(R|r_A) = 0.5$ must hold. This, in turn, requires that $\Pr_A(r_A|R) = \frac{p}{1-p}$ and $\Pr_A(r_R|R) = \frac{1-2p}{1-p}$. As explained in part (4) in the proof of Lemma 11, the mixed-strategy equilibrium is robust to IC if $\beta(L) \geq \beta(R)+\gamma$. IC puts no restriction on the out-of-equilibrium belief $\mu(R|l)$—i.e., $\mu(R|l) \in [0,1]$. It follows that off-the-equilibrium path, $\Pr_D(r|l) = 1$ if $\mu(R|l) > 0.5$, $\Pr_D(r|l) \in [0,1]$ if $\mu(R|l) = 0.5$ and $\Pr_D(r|l) = 0$ if $\mu(R|l) < 0.5$.

(4) If $\beta(R)+\gamma = \beta(L)$, there exist (in addition to the above equilibria) a continuum of mixed-strategy equilibria in which A randomizes between r_A and l in state L *and* between r_A and r_R in state R—see part (3) in the proof of Lemma 11.

Next, assume that $\alpha+\gamma<\beta(L)$. Note that in this case, A always chooses r_A in state L, which vastly simplifies the analysis. With $\alpha+\gamma<\beta(L)$, the honest equilibrium disappears. As usual, the corrupt equilibrium always exists. Moreover, if $\alpha<\beta(R)$, then there exists only the corrupt equilibrium because accepting the payment is the dominant strategy in both states. However, if $\alpha\geq\beta(R)$ then there exists a "perverse honest equilibrium" in which A chooses r_A in state L and r_R in state R. Thus, this equilibrium is fully informative, and $\Pr_D(r|r_R)=\Pr_D(l|r_A)=1$. Moreover, if $\alpha\geq\beta(R)$, there is also the mixed strategy equilibrium with $\Pr_A(r_A|R)\in(0,1)$ and $\Pr_A(r_A|R)+\Pr_A(r_R|R)=1$. The mixed-strategy equilibrium is exactly the same as in (3) above: $\Pr_D(r|r_A)=1-\beta(R)/\alpha$, $\Pr_A(r_A|R)=\frac{p}{1-p}$ and $\Pr_A(r_R|R)=\frac{1-2p}{1-p}$.

Finally, assume that $\alpha+\gamma=\beta(L)$. As usual, the corrupt equilibrium always exists. If $\alpha<\beta(R)$, then there exists only the corrupt equilibrium. Since $\alpha<\beta(R)$, A chooses r_A in state R with probability one. By Lemma 9, $\Pr_D(r|r_A)=1$ must hold since $\alpha<\beta(R)$ implies that $\Pr_A(r_R|R)=0$. Given that $\Pr_D(r|r_A)=1$, A will always choose r_A in both states. If, however, $\alpha>\beta(R)$, there are the honest equilibrium, the perverse honest equilibrium and the mixed strategy equilibrium in which $\Pr_A(r_A|L)=1$, $\Pr_A(r_A|R)\in(0,1)$ and $\Pr_A(r_A|R)+\Pr_A(r_R|R)=1$, in addition to the corrupt equilibrium. Finally, if $\alpha=\beta(R)$, there exists, in addition to the list above, a continuum of mixed-strategy equilibria in which A randomizes between r_A and l in state L *and* between r_A and r_R in state R.

We now prove our claim in the main text that transparency does not make a differency if $\beta(L)\leq\gamma$.[40] Transparency does not make a difference in that case because there is always a fully informative equilibrium in both conditions. First assume that $\beta(L)<\gamma$. In that case, the adviser always chooses l in state L in any equilibrium that satisfies the IC. Suppose not. As shown in the proof of Proposition 3, r_A is always chosen in equilibrium, and then followed by r with probability one in the non-transparency condition. As shown in Lemma 9, at least one of r_A and r_R is played, then followed by r with probability one in the transparency condition (Lemma 9 does not require the assumption that $\beta(L)\geq\gamma$). Thus, l is equilibrium dominated in state R in both conditions. But given the belief $\mu(R|l)=0$, the adviser will always recommend l in state L, a contradiction. In state R, r_A is played in equilibrium, and then followed by r with probability one in both conditions. Note that in the transparency condition, r_R is never chosen in an equilibrium that satisfies the IC. This is because r_A

[40] For conciseness, we prove the claim under the assumption that $p<1/2$. The proof under the assumption that $p\geq1/2$ follows similar lines and is available upon request.

is equilibrium dominated in state L (as $\beta(L) < \gamma$). Given the belief $\mu(L|r_A) = 0$, r_A is the best choice in state R. As a result, if $\beta(L) < \gamma$, then the equilibrium that satisfies the IC is unique and fully informative under both conditions. If $\beta(L) = \gamma$, then the fully informative equilibrium described above goes through but there are also other equilibria in both conditions.

Proof of Proposition 6. The mixed strategy equilibrium is not robust to D1 if $\beta(L) < \beta(R) + \gamma$. Since $D_R^0(l) \subseteq D_L(l)$, $\mu(R|l) = 0$ must hold according to D1. Given this belief, recommending l is a profitable deviation in state L. If however $\beta(L) \geq \beta(R) + \gamma$, then even if $D_R^0(l) \subseteq D_L(l)$ holds, it is no longer true that recommending l is a profitable deviation in state L. Thus, the mixed strategy equilibrium is robust to D1 if and only if $\beta(L) \geq \beta(R)+\gamma$. Next, we show that the honest equilibrium is not robust to D1 if $\beta(L) < \beta(R) + \gamma$. To see why, note that the honest equilibrium relies on off-the-equilibrium path belief, $\mu(L|r_A) \geq 0.5$. Since $\beta(L) < \beta(R)+\gamma$ implies that $D_L^0(r_A) \subseteq D_R(r_A)$, $\mu(L|r_A) = 0$ is a requirement of D1, a contradiction. However, D1 does not contradict the belief $\mu(L|r_A) \geq 0.5$ if $\beta(L) \geq \beta(R) + \gamma$ because $D_L^0(r_A) \subseteq D_R(r_A)$ no longer holds. Finally, the corrupt equilibrium is robust to D1 regardless of $\beta(L)$, $\beta(R)$ and γ since A obtains the highest possible payoff in either state in the corrupt equilibrium (given our assumption that $\beta(L) \geq \gamma$) and has no incentive to deviate no matter what off-the-equilibrium path beliefs are.

INSTRUCTIONS

Welcome to this experiment!

During the course of the experiment, you will be asked to make a series of decisions. If you follow the instructions carefully, you can earn a considerable amount of money, depending on your decisions and those made by other participants. The money you earn will be paid to you in cash at the end of the experiment.

Please remain silent during the experiment. If you have any questions, or need assistance of any kind, please raise your hand and an experimenter will come to you.

Roles

In this experiment you will have to make decisons as EITHER:

- a Sender , OR
- a Receiver.

Whether you are a Sender or Receiver will be decided later.

Rounds and Decisions

During the experiment, each Sender will interact 20 times with various Receivers. Each of these interactions is called a Round.

At the beginning of each Round, each Sender will be randomly matched with a Receiver.

In each Round, the Receiver must choose between Options "A" and "B". Which of these Options is better for the Receiver depends on a "Spinner" which randomly selects "L" or "R":

- If L is selected, then Option A is better for the Receiver.
- If R is selected, then Option B is better for the Receiver.

On average, the Spinner selects L 2 out of 5 times, and R 3 out of 5 times.

However, only the Sender will know whether L or R has been selected. After seeing which has been selected, the Sender chooses whether to recommend to the Receiver Option A or B. Before making the recommendation, the Sender must choose whether to accept or refuse an extra Payment. *If the Sender accepts this Payment he/she must recommend Option B. The exact amount of this extra Payment depends on whether the Spinner has selected L or R.*

The Receiver will be informed about whether or not the Sender accepted the Payment.

In summary, each Round proceeds as follows:

- A Spinner randomly selects L or R.
- The Sender decides whether or not to accept the Payment.
- The amount of the Payment depends on whether the Spinner selected L or R.
- The Sender chooses whether to recommend Option A or B. However, if the Sender accepts the payment, he/she *must* recommend B.
- The Receiver sees whether or not the extra payment was accepted (but not the amount of the payment) and which Option was recommended, and then chooses Option A or B.

Points

Depending on the decisions of the Sender and Receiver, each will get a number of Points:

The Sender gets:

- **5 Points** if L is chosen and he/she accepts the extra Payment, OR **2 Points** if R is chosen and he/she accepts the extra Payment, AND
- **6 Points** if the Receiver follows their recommendation (i.e. the Sender recommends A and the Receiver chooses A, or the Sender recommends B and the Receiver chooses B), AND
- **1 Point** if he/she recommends the better Option for the Receiver (i.e. Option A if the Spinner selects L, and Option B if the Spinner selects R).

The Receiver gets:

- **10 Points** if he/she chooses the better Option (i.e. Option A if the Spinner selects L, and Option B if the Spinner selects R), OR
- **5 Points** if he/she chooses the worse Option (i.e. Option B if the Spinner selects L, and Option A if the Spinner selects R).

After the Points have been decided, the Senders and Receivers are randomly rematched and a new Round begins.

At the end of the experiment, two Rounds from these 20 will be randomly selected, and the Points you earned in those rounds will be converted to euros at the rate

$$3 \text{ points} \equiv 1 \text{ EURO}$$

and will be privately paid to you in cash.

When you have read and understood these instructions, type 4567 into the box on your screen and click OK.

FURTHER INSTRUCTIONS

Entering Decisions

Sender

The Sender will see whether the Spinner selects L or R, then make decisions about whether to accept or refuse the payment and which Option to recommend.

Receiver

Instead of actually seeing the Sender's recommendation before making a choice, the Receiver will decide what he/she will choose:

- if the Sender refuses payment and recommends A,
- if the Sender refuses payment and recommends B, AND
- if the Sender accepts the payment and recommends B.

Determining Payoffs

When both the Sender and Receiver have made their decisions, the Sender's decisions will be matched with the choice that the Receiver said he/she would make given the actual Sender's decision.

These are the decisions and choices that will determine how many Points the Sender and Receiver earn in that Round.

Questions and Practice Rounds

Before finding out whether you will be a Sender or Receiver, you will have a chance to make decisions for both a Sender and Receiver and see how the payoffs are determined.

These practice decisions will not affect how much you are paid.

When you are ready to make practice decisions, type 9876 into the box on your screen and click OK.

INSTRUCTIONS

Welcome to this experiment!

During the course of the experiment, you will be asked to make a series of decisions. If you follow the instructions carefully, you can earn a considerable amount of money, depending on your decisions and those made by other participants. The money you earn will be paid to you in cash at the end of the experiment.

Please remain silent during the experiment. If you have any questions, or need assistance of any kind, please raise your hand and an experimenter will come to you.

Roles

In this experiment you will have to make decisions as EITHER:

- a Sender , OR
- a Receiver.

Whether you are a Sender or Receiver will be decided later.

Rounds and Decisions

During the experiment, each Sender will interact 20 times with various Receivers. Each of these interactions is called a Round.

At the beginning of each Round, each Sender will be randomly matched with a Receiver.

In each Round, the Receiver must choose between Options “A” and “B”. Which of these Options is better for the Receiver depends on a “Spinner” which randomly selects “L” or “R”:

- If L is selected, then Option A is better for the Receiver.
- If R is selected, then Option B is better for the Receiver.

On average, the Spinner selects L 2 out of 5 times, and R 3 out of 5 times.

However, only the Sender will know whether L or R has been selected. After seeing which has been selected, the Sender chooses whether to recommend to the Receiver Option A or B. Before making the recommendation, the Sender must choose whether to accept or refuse an extra Payment. *If the Sender accepts this Payment he/she must recommend Option B*. *The exact amount of this extra Payment depends on whether the Spinner has selected L or R.*

The Receiver will not be informed about whether or not the Sender accepted the Payment.

In summary, each Round proceeds as follows:

- A Spinner randomly selects L or R.
- The Sender decides whether or not to accept the Payment.
- The amount of the Payment depends on whether the Spinner selected L or R.
- The Sender chooses whether to recommend Option A or B. However, if the Sender accepts the payment, he/she *must* recommend B.
- The Receiver sees which Option was recommended, and then chooses Option A or B.

Points

Depending on the decisions of the Sender and Receiver, each will get a number of Points:

The Sender gets:

- **5 Points** if L is chosen and he/she accepts the extra Payment, OR **2 Points** if R is chosen and he/she accepts the extra Payment, AND
- **6 Points** if the Receiver follows their recommendation (i.e. the Sender recommends A and the Receiver chooses A, or the Sender recommends B and the Receiver chooses B), AND
- **1 Point** if he/she recommends the better Option for the Receiver (i.e. Option A if the Spinner selects L, and Option B if the Spinner selects R).

The Receiver gets:

- **10 Points** if he/she chooses the better Option (i.e. Option A if the Spinner selects L, and Option B if the Spinner selects R), OR
- **5 Points** if he/she chooses the worse Option (i.e. Option B if the Spinner selects L, and Option A if the Spinner selects R).

After the Points have been decided, the Senders and Receivers are randomly rematched and a new Round begins.

At the end of the experiment, two Rounds from these 20 will be randomly selected, and the Points you earned in those rounds will be converted to euros at the rate

$$3 \text{ points} \equiv 1 \text{ EURO}$$

and will be privately paid to you in cash.

When you have read and understood these instructions, type 4567 into the box on your screen and click OK.

FURTHER INSTRUCTIONS

Entering Decisions

Sender

The Sender will see whether the Spinner selects L or R, then make decisions about whether to accept or refuse the payment and which Option to recommend.

Receiver

Instead of actually seeing the Sender's recommendation before making a choice, the Receiver will decide what he/she will choose:

- if the Sender recommends A, AND
- if the Sender recommends B.

Determining Payoffs

When both the Sender and Receiver have made their decisions, the Sender's decisions will be matched with the choice that the Receiver said he/she would make given the actual Sender's decision.

These are the decisions and choices that will determine how many Points the Sender and Receiver earn in that Round.

Questions and Practice Rounds

Before finding out whether you will be a Sender or Receiver, you will have a chance to make decisions for both a Sender and Receiver and see how the payoffs are determined.

These practice decisions will not affect how much you are paid.

When you are ready to make practice decisions, type 9876 into the box on your screen and click OK.