Karl-Franzens-University Graz
Faculty of Social and Economic Sciences

Should I wait or should I lie? Path dependency and timing in repeated honesty decisions under frames

Christian Schitter, Stefan Palan



Working Paper 2018-05

September 6, 2018

An electronic version of the paper may be downloaded from:

University of Graz: sowi.uni-graz.at/forschung/working-paper-series/
RePEc: ideas.repec.org/s/grz/wpsses.html

Working Paper Series
Faculty of Social and Economic Sciences
Karl-Franzens-University Graz
ISSN 2304-7658
sowi.uni-graz.at/forschung/working-paper-series/
sowi-wp@uni-graz.at

UNIVERSITY OF GRAZ

Should I wait or should I lie? Path dependency and timing in repeated honesty decisions under frames

Christian Schitter[a], Stefan Palan[a,]*

Working Paper 2018-05

September 6, 2018

Abstract

We experimentally investigate time and path dependency in sequences of decisions about whether to be honest under a gain, a lottery and a loss framing. We find only small timing patterns over rounds, but clear evidence for more dishonesty after streaks of unfavorable outcomes. The latter implies a systematic path dependency in repeated honesty decisions. We observe an increase in dishonesty from gain to lottery and from lottery to loss framing. Increased dishonesty generally appears earlier and faster under a loss compared to a gain or a lottery frame. Surprisingly, the most honest participants are also those most clearly exhibiting path dependency in reports, in line with a behavior akin to "moral accounting".

Keywords: honesty, decision sequences, frames, path dependencies

[a] University of Graz, Department of Banking and Finance, Universitaetsstrasse 15, 8010 Graz, AUSTRIA

[*] Corresponding author. Tel.: +43(316)380-7306, E-Mail stefan.palan@uni-graz.at.

Working Paper Series
Faculty of Social and Economic Sciences
Karl-Franzens-University Graz
ISSN 2304-7658
sowi.uni-graz.at/forschung/working-paper-series/
sowi-wp@uni-graz.at

Should I wait or should I lie? Path dependency and timing in repeated honesty decisions under frames ☆

Christian Schitter[a,*], Stefan Palan[a]

[a] *Department of Banking and Finance, University of Graz*

Abstract

We experimentally investigate time and path dependency in sequences of decisions about whether to be honest under a gain, a lottery and a loss framing. We find only small timing patterns over rounds, but clear evidence for more dishonesty after streaks of unfavorable outcomes. The latter implies a systematic path dependency in repeated honesty decisions. We observe an increase in dishonesty from gain to lottery and from lottery to loss framing. Increased dishonesty generally appears earlier and faster under a loss compared to a gain or a lottery frame. Surprisingly, the most honest participants are also those most clearly exhibiting path dependency in reports, in line with a behavior akin to "moral accounting".

Keywords: Honesty, decision sequences, frames, path dependencies

1. Introduction

In 2012, the swimmer Ye Shiwen attracted attention with her unexpectedly good performance at that year's Olympic Games. Many commentators saw an unusual improvement compared to her performance at previous sports events and uttered suspicions about its cause. Doping tests were negative, but still suspicion remained. This led to an outraged debate about the ethicality of suspecting a rule-compliant athlete of unethical conduct due to only circumstantial evidence, or, more poignantly, due to nothing else than delivering an "outstanding performance".[1] Callaway (2012) saw this event as motivation for discussing performance profiling as a way to increase rule-compliance in sports. Performance profiling refers to a technique combining repeated doping tests and an analysis of performance

☆The authors thank Thomas Teubel for support in collecting the data. We are grateful to participants of the Austrian Working Group on Banking and Finance 2016 for helpful comments. Financial support from the University of Graz is gratefully acknowledged.

*Corresponding author

Email addresses: `christian.schitter@edu.uni-graz.at` (Christian Schitter), `stefan.palan@uni-graz.at` (Stefan Palan)

[1] See e.g. https://www.theguardian.com/sport/london-2012-olympics-blog/2012/aug/09/dean-macey-olympic-drugs, retrieved 17-05-2017.

changes over time. It is intended to assess the likelihood of the use of performance-enhancing substances and to raise red flags to trigger focused tests and investigations. The number of outraged comments that followed in *Nature*'s comments section was so large that the journal's online system could not process them.[2] Most of the outrage centered around singling out the specific case of Ye, but also more broadly criticized the method itself for, e.g., not adding much value beyond existing doping tests. Zhong et al. (2012) replied to Callaway (2012) by comparing performance histories of other top athletes with Ye's history. They showed that Ye's performance was in fact not an unusual case. It could therefore be argued that they thus also used performance profiling, but in this case with the intent to clear Ye of the accusation of cheating.

Similar situations to the one described above occur in many other economic contexts and come to mind easily: A salesperson with average track record suddenly reports an outstanding quarterly result. A mediocre student suddenly writes an excellent test. An insurance holder constantly paying fees but not registering claims suddenly files an unusually large claim. A tax payer with a long record of filing no deductions suddenly files high deductions. In light of the high frequency of such situations and of the uproar about Ye's case: Should inconsistent longitudinal observations of performance or reports in general be cause for heightened suspicion about cheating, misreporting and violation of norms? Or would such assumptions be ineffectual and – on top – unfair against the person under scrutiny? Put differently, and in more neutral terms: Are there patterns of moral behavior in repeated decision-making which could aid in the task of estimating the likelihood of an observation having a fraudulent cause? Auditing strategies and monitoring decisions could benefit from knowledge about patterns in such situations.

Considering the examples mentioned earlier, it is well worth noting that situations can differ in how they are perceived: Depending on their expectations, athletes can either perceive winning as a gain (e.g., at the beginning of their career) or losing as a loss (e.g., when they are already highly successful). An insurance holder may similarly perceive the insurance contract as a lottery: To assess whether the contract was "profitable", she may weigh total premia paid against the value of the expected redemption payment in case of a loss. As yet another example, citizens often perceive paying taxes as a loss (depending on how the tax filing process is structured; see Engström et al., 2015). We therefore also ask: Do patterns of moral behavior in repeated decisions change under frames pertaining to the value domain (gain, lottery, loss)?

The present research also bears relevance for experimental honesty experiments, which have been surging in experimental economics in recent years (see e.g. Abeler et al. 2016 for a meta-study). Such experiments often utilize sequences of reports in which a subject can repeatedly decide whether to be honest or dishonest (e.g. Weisel and Shalvi 2015). If we find significant longitudinal patterns, controlling for them could be a strategy to increase the robustness and validity of the results of such experiments.

To summarize, we aim to assess whether there are behavioral patterns in repeated de-

[2]Editor's note at http://www.nature.com/news/why-great-olympic-feats-raise-suspicions-1.11109, retrieved 17-05-2017.

cisions about whether to be or not to be honest. To that end, we gather evidence from sequences of self-reports under a gain, a lottery and a loss frame in an online experiment, where dishonest self-reports increase participants' final payoffs. These data permit us to establish empirically-founded longitudinal patterns of dishonesty under different frames. The current study is the first comprehensive behavioral investigation into this topic.

2. Material and methods

2.1. Experimental design

We employ a between-subjects online variation of the die paradigm pioneered by Fischbacher and Föllmi-Heusi (2013) (FFH paradigm from here on), closely following the design of Effron et al. (2015). Under the FFH paradigm, subjects throw a die in private and self-report the unverified outcome, which is directly tied to subjects' payoffs. Subjects thus have the option of increasing their payoff by reporting outcomes with a payoff higher than the one actually observed. This method does not permit inferring honesty from individual reports, but allows estimating aggregate dishonesty by comparing the realized distribution of reports to the theoretical distribution assuming honesty. The FFH paradigm as an experimental method is therefore well-suited for investigating dishonesty, as it comprehensively preserves the anonymity of dishonest subjects, which is likely to be relevant in the cheating decision (see e.g. Gneezy et al. 2018). Additionally, recent studies find that outcomes of such experiments correlate with rule-compliance in the field (see e.g. Dai et al. 2017 and Cohn and Maréchal 2017).

In contrast to FFH's original study that uses a die as randomization device, we ask subjects participating in our online variant of the FFH paradigm to prepare a coin prior to commencing the experiment, telling them that this coin will be used to determine the variable part of their payment.[3] When starting the experiment, participants first receive the experiment instructions and then go through a series of screens, each of which asks them to toss the coin and report the outcome as Heads or Tails. Heads is always the favorable outcome (associated with a higher final payoff), while Tails is always the unfavorable outcome, as explained in detail later. Obviously, at the individual level, we can neither control whether participants use a coin, nor whether they report the true outcome if they do. Participants therefore have a clear financial incentive to report the favorable outcome, even if they do not actually observe it.

In the instructions, we inform participants that there will be up to 15 rounds of coin tosses, with the actual number of rounds having been determined before the start of the experiment.[4] This generates uncertainty regarding the end point of our sequence, as is the

[3]We chose this method because we believe participants at home to more readily have a coin to hand than a die. Despite the binary outcome of the coin toss, participants can still lie partially (that is, not lie to the full extent) because they are asked to report a series of coin tosses where they can vary the number of honest reports.

[4]We do not specify how this determination was accomplished and, for example, do not suggest that it was the result of a random draw to rule out potential concerns regarding deception.

	Treatment		
	Gain	*Lottery*	*Loss*
Fixed payoff for participating	60	60	60
Endowment at the beginning	0	75	150
Fixed payoff per round	0	-5	0
Payoff per round of reporting HEADS	10	10	0
Payoff per round of reporting TAILS	0	0	-10
Expected payoff after 15 rounds	**135**	**135**	**135**

Table 1: Payoff description by treatment, in GBP-pence. Expected payoffs are values under full honesty.

case in many naturally occurring settings. Nevertheless, we set the number of rounds to 15 for every participant. The uncertainty regarding the duration of the game is thus completely resolved by the time subjects play round 15, and we would expect to observe an end-game effect (if any) at least at that point. After each report, we inform participants about the impact of this report on their payoff. When registering their reports, participants also see how many rounds have elapsed and thus know how many may at most be yet to come.

Participants start the experiment with different endowments in the different treatments. With their coin tosses, they can either add money to their payoff or lose part of their initial endowment. Specifically, in treatment *Gain*, there is no initial endowment and participants are instructed that they will earn 10 GBP-pence[5] for every report of HEADS and nothing for TAILS. In treatment *Lottery*, there is an initial endowment of 75 pence, from which 5 pence are deducted each round as a participation fee. Participants again receive 10 pence for every report of HEADS and nothing for TAILS. In treatment *Loss*, the initial endowment is 1.50 GBP. Participants lose 10 pence for each report of TAILS and neither gain nor lose anything for reporting HEADS. While the amounts may sound small, they exceed the required and commonly expected compensation of participants in our subject pool severalfold, a point we will return to later. The expected payoff is equal across treatments, as is the marginal effect of any particular report (HEADS or TAILS) on the final payoff. Therefore, the results would be the same for all treatments if decisions were immune to framing effects. Table 1 summarizes the treatments.

The description of the task is at all times phrased in a neutral manner, without any reference to honesty or the experimenter's motivation for running the task at hand (in contrast to Effron et al. 2015, who invented a reason for the coin tossing task and urged participants to be honest). Mazar et al. (2008) show that reminders relating to issues of morality have the power to reduce dishonesty, and we therefore believe that such urging could also change other behavioral patterns in the context of our experiment. Appendix A contains our instructions for all treatments, as well as the text of the invitation to the experiment.

We follow the coin tossing task with a questionnaire in which we collect additional

[5] 1 GBP was about 1.32 USD when we ran the experiment at the beginning of August 2016.

controls (e.g., self-reported risk tolerance) for our analysis.[6]

Participants were recruited online[7] via the online platform Prolific.ac (Palan and Schitter, 2018) and received – in addition to the variable payment – a fixed payment of 60 pence for their participation. This fixed payment is in line with recommendations of Prolific.ac for a study with an average duration of 7 minutes like the one here.[8] The payoff from the coin tossing part was paid as a bonus. Its expected value under honesty more than doubles the recommended total payment for this study, while the maximum value more than triples it. We restricted participation to residents of the U.S. who indicated having English as their first language, both to exclude confounding factors from cultural differences and to ensure proficiency in English as far as possible.

We recruited a total of 614 participants in August 2016. For our analysis, we remove 28 (4.6% of total) participants for different reasons,[9] leaving 199, 192 and 195 participants for treatments *Gain*, *Lottery* and *Loss*, respectively. The experiment was run using the survey tool Limesurvey. The final sample consists of subjects from diverse backgrounds: 33.5% indicate being students; 54.8% are full-time and 20.3% part-time employed, while 13.1% indicate being job-seekers and 11.8% report another status (e.g. retired); 41.6% of all subjects are female; mean age is 32.0 (SD 11.2).

2.2. Measurements

We are interested in the incidence of favorable reports. In all experiments, Heads represents the favorable report and we define a binary variable WIN accordingly:

$$\mathrm{WIN}_t^i = \begin{cases} 1, & \text{if participant } i \text{ reports Heads in round } t \\ 0, & \text{otherwise} \end{cases}$$

Since we are interested in path dependency, we need to account for the history preceding a given report. In our case, we define path dependency in terms of the number of consecutive rounds of unfavorable reports prior to a given reporting decision. We call this variable WINLAG and define it as follows:

$$\mathrm{WINLAG}_t^i = t - 1 - \max_{t^*} \left\{ t^* < t \wedge \mathrm{WIN}_{t^*}^i = 1 \right\}, \text{ with } \mathrm{WINLAG}_1^i \equiv 0 \ \forall i$$

On the participant level, we also introduce the following notation: We define $\mathrm{WIN}^i \equiv \sum_{t=1}^{15} \mathrm{WIN}_t^i$ as a participant's total number of favorable outcomes. When multiplied by 10 pence, this equals the total variable payment for the respective participant. Furthermore, let

[6]The questionnaire also contained questions for another experiment. Given that participants are not aware of the contents of the questionnaire before they finish tossing the coin, this does not influence the results of the present study.

[7]Using subjects sourced from dedicated online platforms is by now a common and reliable method for running economic experiments, see e.g. Horton et al. (2011) or Arechar et al. (2018).

[8]The 7 minutes include the questionnaire.

[9]11 subjects did not finish the study, 2 were missing demographic data, and 15 failed an attention check in the questionnaire.

WIN^{Gain}, $\text{WIN}^{Lottery}$ and WIN^{Loss} be the averages of WIN^{i} over all subjects in the respective treatments.

3. Related literature and behavioral predictions

If people were either always honest or always dishonest, all reporting patterns other than 15 times HEADS (the maximum report) would inevitably always stem from an honest person.[10] A broad range of research shows, however, that a majority of people are generally honest, but often defect partially or fully under a range of circumstances (Mazar et al., 2008; Jacobsen et al., 2018). Furthermore, the work of Schindler and Pfattheicher (2016), Grolleau et al. (2016) and Schwartz Cameron et al. (2008) establishes both experimentally and theoretically that people are more likely to lie to avoid a loss than to realize a gain, i.e., that moral behavior exhibits loss aversion. In the existing literature, only Blanco et al. (2015) find no significant loss framing effect on honesty in an FFH-like task. This could, however, be due to the limited statistical power of their setup.[11] Given that situations perceived as involving lotteries do not weigh as strongly as situations perceived as involving losses (Tversky and Kahneman, 1981), it is therefore straightforward to derive our first hypothesis:

H1 $\text{WIN}^{Gain} \leq \text{WIN}^{Lottery} < \text{WIN}^{Loss}$.

Concerning temporal patterns of dishonesty, we know of two psychological effects which postulate an increase in dishonesty over time: First, Gino and Bazerman (2009) show that a gradual erosion of morality (slippery slope) is perceived as being less reprehensible than abrupt changes in honesty. This implies that people may perceive exhibiting a gradual increase of dishonesty to be more compatible with their respective self-images than exhibiting large amounts of dishonesty from the start. Gino and Bazerman (2009) thus predict low dishonesty early on, and an increase of dishonesty over time. Second, Mead et al. (2009) argue that ego-depletion caused by tasks which reduce an individual's self-control increases dishonesty in subsequent tasks. As self-control is assumed to be a limited resource that can be reduced by a variety of factors, seemingly unrelated tasks that deplete self-control can therefore also increase the propensity for dishonesty. Further advancing this argument, Gino et al. (2011) establish that refraining from lying by itself reduces self-control, and for this reason predict that moral failure will increase over time. Given the experimental design in these papers, however, the authors can observe directly neither behavior in longer sequences of honesty decisions, nor how recovery from ego-depletion (through, e.g., positive affect; see Tice et al., 2007) affects dishonesty.

[10] Note that always reporting the favorable outcome is also the only rational strategy whenever subjects derive utility solely from maximizing their own monetary outcomes, since our setup explicitly rules out any monetary punishment for dishonesty.

[11] Blanco et al. (2015) employ a version of the FFH paradigm with a comparably small sample size of only 48 participants per treatment. Additionally, they use 10 uniformly distributed observed states (i.e., less than 5 observations per state on average), further exacerbating the issue of having few participants. This setup results in standard errors too large to allow for clear inference on the effect of interest.

Directly testing these predictions, a meta-study of FFH-like honesty experiments with repeated decisions by Abeler et al. (2016) only finds a very small, albeit significant increase in dishonesty over rounds. Effron et al. (2015) study sequences of honesty decisions under a gain frame and initially diagnose a non-negligible linear increase in dishonesty over rounds. However, using more sophisticated methods, they finally report that the increase is non-linear and is in fact due to a cheating-at-the-end effect, in which dishonesty is significantly more prevalent when people are aware that a specific decision constitutes their last chance to cheat. They explain the latter effect through regret aversion. Contrary to these results, Touré-Tillery and Fishbach (2012) find dishonest behavior in sequences of real effort tasks to be particularly high in the middle of a sequence, which they argue may be due to the fear of increased scrutiny at the beginning and end of sequences. However, given the specific tasks and setup of Touré-Tillery and Fishbach (2012),[12] we rely on the literature more closely following the FFH paradigm to postulate:

H2a The proportion of dishonest reports increases over time/rounds.
H2b There is a cheating-at-the-end-effect. Dishonest reports are significantly more prevalent in the decision perceived to offer the last chance for cheating.

The literature so far has little to say about the interplay of framing and ego-depletion. Pocheptsova et al. (2009) find that ego-depletion promotes reference-dependent choices induced by framing, e.g., that there is a higher susceptibility to loss-aversion among depleted than among non-depleted subjects. We therefore assume that ego-depletion will result in a faster and earlier increase of dishonesty under the loss than under the gain frame. We also hypothesize that refraining from lying is generally harder in a loss frame than in a gain frame, such that self-regulatory resources are depleted more quickly in the former than in the latter. Taking together these two arguments, we postulate:

H2c The increase over time is more pronounced and happens faster in the *Loss* frame than in the *Lottery* and *Gain* frames.

Another part of the literature reports on variations of what we will call "moral accounting" models (e.g., moral self-licensing and conscience accounting). Here, moral action can balance out immoral action, thereby making moral failure more likely when there is an opportunity for doing good that can be rationalized to offset said immoral behavior. This mechanism holds for moral actions preceding an immoral act (see Merritt et al. 2010 for an overview), for an opportunity to do good after an immoral act when the existence of this opportunity is known in advance (Gneezy et al., 2014), and even for refraining from behaving immorally as a way of building credit for behaving immorally later on (Effron et al., 2012). For sequences of honesty decisions, these findings imply path dependency: People should

[12]Touré-Tillery and Fishbach (2012) present 5 experiments in which participants could lower their workload (e.g., how long texts are that they had to proofread) or present themselves more favorably (e.g., reporting to closely follow a religious ritual) by being dishonest in repeated reports. Contrary to our experiment, subjects could not improve their income by being dishonest, which could be one reason for different findings compared to the FFH paradigm.

feel more entitled to morally default after a series of honest reports because of the moral credit they have built up by being honest. Defaulting, however, uses up this accumulated credit, and thus they would return to being honest after the default. Specifically, for our situation we posit:

H3a A higher WINLAG implies a higher probability of a favorable reported outcome.
H3b Overall levels of dishonesty at time $t+1$ are independent of WINLAG_t^i when following a favorable report $\text{WIN}_t^i = 1$.

Note that Effron et al. (2015), as a side result in their appendix, find no conclusive evidence regarding oscillation of reports, that is, correlation between consecutive reports under a gain frame. At first glance, this result seems to counter at least our hypothesis H3a. However, in our notation, they only look at results for WINLAG=0 and WINLAG=1, while we also account for higher WINLAGs.

Concerning framing, it is unclear what the source of the perceived moral credit in moral accounting models is, and therefore also how framing might affect it. We assume that refraining from dishonesty is a cost which creates moral credit when incurred. In this case we find it reasonable to assume that realizing a perceived loss by being honest is more costly for subjects than foregoing a perceived gain by being honest. Consequently, moral failure should happen earlier under a loss than under a gain and a lottery frame:

H3c The proportion of dishonesty following positive WINLAG increases earlier under the *Loss* frame than under the *Lottery* and *Gain* frames.

In total, our contribution to the literature is the following: We are the first to comprehensively discuss path and time dependency in sequences of honesty decisions and whether identified patterns change under frames, using a variant of the canonical FFH paradigm. While some of our analyses serve as a replication study of previous results under gain framing and in one-shot decisions, our contribution expands upon these results, particularly with respect to the question of path dependencies.

4. Results

4.1. General differences between treatments

Overall, we find average WIN^i to significantly exceed the expected value of 7.5 favorable outcomes in all treatments, as reported in Table 2. This indicates dishonesty in all treatments, in line with most of the existing literature. While the deviation from 7.5 may seem small in absolute terms, this impression is misleading: To obtain the deviations displayed in Table 2, we require estimated proportions of dishonesty of 17.5%/24.9%/30.9% in the *Gain*/*Lottery*/*Loss* treatments, respectively. Thus, of the 50% of cases where a participant is expected to have observed Tails in treatment *Gain*, in 17.5% of the cases the participant reports Heads. This of course is a non-negligible amount of dishonesty, and the picture is

		Treatment			
		Gain	*Lottery*	*Loss*	Total
Participants		199	192	195	586
Average WIN_t^i		8.81	9.36	9.82	9.33
t-test for difference from 7.5, p-value		<0.001	<0.001	<0.001	<0.001
Observed % WIN^i		58.76%	62.43%	65.44%	62.18%
Binomial test for difference from 50%, p-value		<0.001	<0.001	<0.001	<0.001
Estimated % dishonesty		17.52%	24.86%	30.87%	24.37%
Wilcoxon rank-sum test, p-value	*Gain*		0.037	<0.001	
	Lottery			0.037	

Table 2: Statistics of treatment and total results. Wilcoxon rank-sum tests are Holm (1979)-adjusted for multiple testing.

stronger in the *Lottery* and *Loss* frames.[13]

A Holm (1979)-adjusted, pairwise Wilcoxon rank-sum test comparing treatments diagnoses a highly significant difference between the *Gain* and *Loss* treatments ($W = 14210$, $p < 0.001$) as well as significant differences between the *Gain* and *Lottery* ($W = 16498$, $p = 0.037$) and the *Lottery* and *Loss* ($W = 21069$, $p = 0.037$) treatments in Table 2.

The distributions of reports at the participant level shown in Figure 1 confirm the findings regarding the differences between frames. The histograms exhibit a mostly smooth shift in all categories towards higher outcomes when going from *Gain* to *Lottery* and finally to *Loss*.

Overall, these results corroborate previous studies on framing and honesty in one-shot decisions in a repeated setting:

Result 1 We find that $\text{WIN}^{Gain} < \text{WIN}^{Lottery} < \text{WIN}^{Loss}$. This supports H1.

4.2. Time dependency

Figure 2 displays the estimated dishonesty per round and treatment. Dishonesty increases over time in all treatments when comparing start and end points, but with large variation in between. A test for trends in proportions delivers marginal significance for *Gain* at the 10%-level ($p = 0.093$), significance for *Lottery* ($p = 0.003$) and no significance for *Loss* ($p = 0.138$). While estimated dishonesty in the *Loss* treatment is mostly above the other two treatments, there is thus no evidence that it increases. The proportion of favorable reports in *Loss* is high throughout and has the highest peak in dishonesty of all rounds and treatments in round 13.

[13]We estimate dishonesty for binomial outcomes as in Moshagen and Hilbig (2017): Let p be the true probability of a subject observing the favorable outcome in private (in our case $p = 0.5$) and, consequently, let $1 - p$ be the true probability of a subject observing the unfavorable outcome. Let q be the fraction of favorable reports. Under dishonesty, $q > p$. The estimated potential dishonesty d can then be inferred based on the assumption that no subject will report the unfavorable outcome when observing the favorable outcome (i.e., no one will lie to his or her own disadvantage). In this case, the following equation holds for the percentage of honest reports $1-d$ that is unfavorable: $(1-d)(1-p) = 1-q$, as a fraction p of all honest reports will be favorable by chance, which forms part of q. It follows that $d = \frac{q-p}{1-p}$, or in our case $d = \frac{q-0.5}{0.5}$. Standard errors are adjusted accordingly in all presented results.

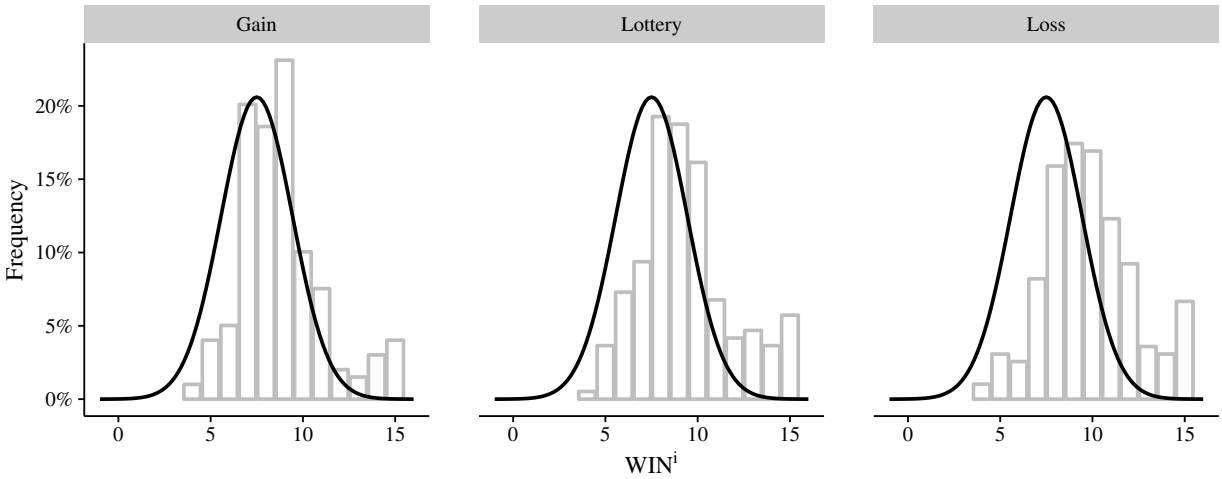

Figure 1: Distribution of accumulated outcomes on participant level per treatment. Bell curve (dark black) indicates expected distribution under full honesty.

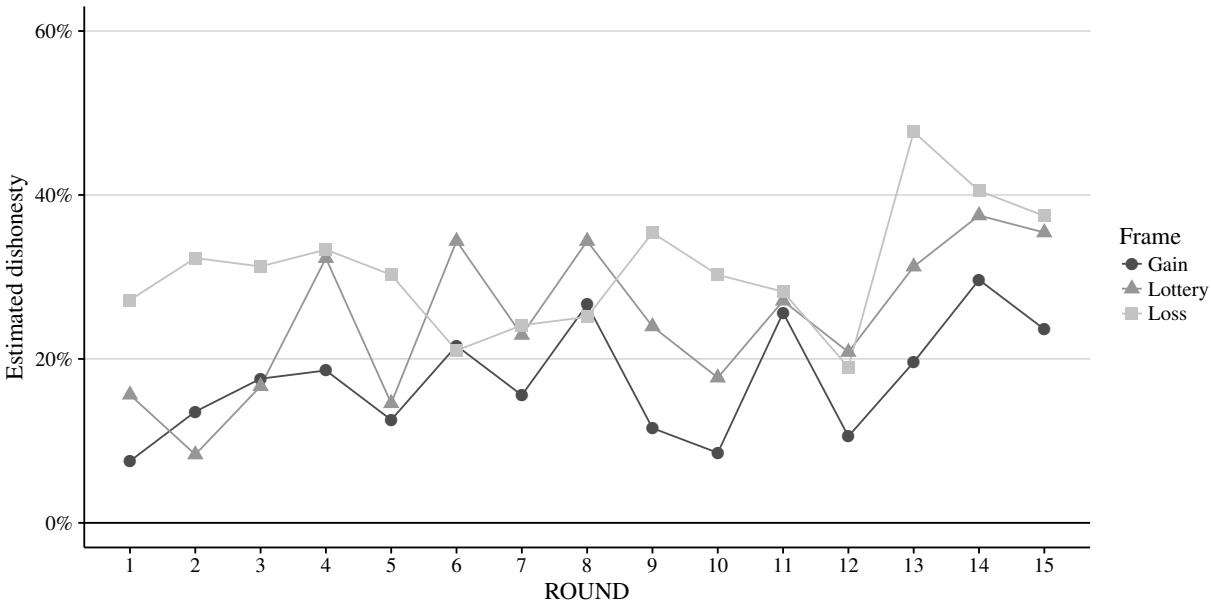

Figure 2: Percentage of estimated dishonesty per round and by treatment.

The logistic regression analyses in Table 3 support these initial observations. Contrary to the direct evidence in trends, however, we do not find an interaction effect between timing and treatments here.

Our findings are in line with Effron et al. (2015), who also identify a round effect under a gain frame, but find this observation to be driven by a strong cheating-at-the-end-effect. The latter effect might also be the single cause for why our logistic regression exhibits an increase over time. Replicating Effron et al. (2015)'s reverse-Helmert coded logistic regression in

		Coeff.	Std. Error	Wald χ^2	Pr> χ^2
Full model	Intercept	0.21***	0.05	3.98	0.00
(n=8790)	Lottery	0.15**	0.07	2.25	0.02
	Loss	0.28***	0.07	4.14	0.00
	ROUND	0.02***	0.00	4.37	0.00
Full model	Intercept	0.24***	0.07	3.65	0.00
incl. interactions	Lottery	0.06	0.09	0.63	0.53
(n=8790)	Loss	0.29***	0.10	3.05	0.00
	ROUND	0.01**	0.01	2.06	0.04
	ROUND x Lottery	0.01	0.01	1.18	0.24
	ROUND x Loss	-0.00	0.01	-0.11	0.91

$^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$

Table 3: Logistic regression with dependent variable WIN_t^i. Standard errors clustered at the participant level.

Table 4, we also find significantly higher cheating levels only in the last three rounds. Given that our participants – contrary to Effron et al. (2015) – do not know for certain which will be the last round prior to round 15, earlier cheating-at-the-end seems reasonable. Similarly to Effron et al. (2015), we therefore conclude that this increase in the last three rounds is the driver of the significance of the ROUND variable in the regressions in Table 3. Taken together, these observations lead us to report the following result:

Result 2 The proportion of dishonest reports increases slightly over rounds (lending tentative support to H2a), with no significant interaction between rounds and frames (not supporting H2c). However, given the small effect size and a disproportionate increase in the final three rounds, the observed effect is driven not by a relevant timing component but rather by a cheating-at-the-end-effect (supporting H2b).

4.3. Path dependency

We begin our investigation of path dependency by noting that participants can easily influence WINLAG. If a subject is willing to always lie, she will always have a WINLAG of 0. Similarly, if a subject occasionally reports honestly to decrease potential suspicion of being dishonest, the effect would be that WINLAG remains small, most likely 0 or 1. In general, we would hence expect those subjects with high WINLAG to be the more honest subjects. It is therefore particularly interesting to study Panel (a) in Figure 3. The Figure documents that the empirically estimated dishonesty generally increases in WINLAG. This suggests an erosion of honesty following unfavorable prior outcomes, implying path dependency in moral behavior. While dishonesty levels develop similarly over WINLAG in the *Gain* and *Lottery* treatments, the *Loss* treatment differs. Starting from only a marginally higher level than in the other treatments in case of WINLAG= 0, the proportion of dishonest reports increases more sharply (and significantly) for WINLAG= 1 in *Loss* than in *Gain* and

	Coeff.	Std. Error	Wald χ^2	Pr> χ^2
Intercept	0.36***	0.04	9.54	0.00
Lottery	0.15***	0.05	2.88	0.00
Loss	0.29***	0.05	5.29	0.00
ROUND=2	0.01	0.06	0.24	0.81
ROUND=3	0.03	0.03	0.89	0.37
ROUND=4	0.05*	0.02	1.95	0.05
ROUND=5	-0.01	0.02	-0.46	0.64
ROUND=6	0.02	0.02	1.09	0.28
ROUND=7	-0.00	0.01	-0.18	0.86
ROUND=8	0.02*	0.01	1.67	0.09
ROUND=9	0.00	0.01	0.27	0.79
ROUND=10	-0.01	0.01	-0.89	0.38
ROUND=11	0.01	0.01	1.14	0.25
ROUND=12	-0.01	0.01	-1.40	0.16
ROUND=13	0.02**	0.01	2.54	0.01
ROUND=14	0.02***	0.01	3.09	0.00
ROUND=15	0.01**	0.01	1.97	0.05

$^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$

Table 4: Logistic regression with dependent variable Win_t^i. ROUND factors are reverse-Helmert coded.

Lottery.[14] However, for WINLAG> 1 we see no significant changes anymore.[15] Experiencing the situation as situated in the loss domain seems to induce more participants to become dishonest for lower WINLAGs. This is also evident in Table 5: There are more participants with at least one WINLAG> 2 in *Gain* (38.2%) than in *Lottery* (36.5%) and in *Loss* (24.6%). This means more participants endure longer WINLAGs in both *Gain* and *Lottery* than in *Loss*.

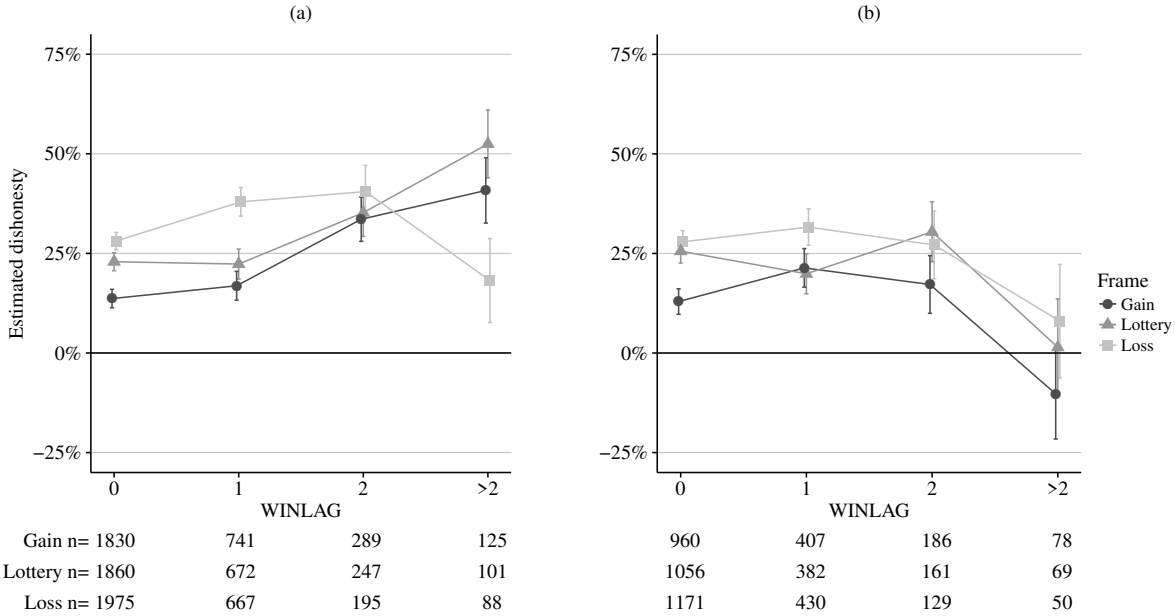

Figure 3: (a) Estimated dishonesty for WIN_t^i by WINLAG and by treatment; (b) WIN_{t+1}^i for observations with $\text{WIN}_t^i = 1$ by WINLAG_t^i and treatment. Table below chart lists number of observations. Error bars indicate standard errors.

The logistic regressions in Table 6 very well reflect the results from the graph. In treatments *Gain* and *Lottery*, the increase of dishonesty over WINLAG is strong and highly significant for WINLAG= 2 and WINLAG> 2 compared to WINLAG= 0. Given the different behavior in *Loss*, only WINLAG= 1 yields a significant increase in dishonesty over WINLAG= 0. Contrary to the other treatments, dishonesty falls for WINLAG> 2 in treatment *Loss*, but not significantly. This lends further support to the observation that participants in the *Loss* treatment are more willing to be dishonest in order to prevent longer WINLAGs from occurring. The full regression is highly robust to adding a range of controls, like sex, age, employment status, student status or risk tolerance (see Appendix B), as well as to different exclusion rules based on statistical ex-post indicators of dishonesty (see Appendix C).

[14]This is in line with the finding of Effron et al. (2015) that there is no significant oscillation between round n and $n-1$ under a gain frame.

[15]Note that the results for WINLAG$\geq$ 2 need to be interpreted with caution, since the number of observations drops substantially, as reflected in much larger standard errors.

		WINLAG=0		WINLAG=1		WINLAG=2		WINLAG>2	
		Obs.	*Exp.*	*Obs.*	*Exp.*	*Obs.*	*Exps*	*Obs.*	*Exp.*
Fraction of 8204 obs.	*Gain*	58.5%	50.0%	26.6%	26.8%	10.4%	12.5%	4.5%	10.7%
	Lottery	62.1%	50.0%	25.0%	26.8%	9.2%	12.5%	3.8%	10.7%
	Loss	65.2%	50.0%	24.4%	26.8%	7.1%	12.5%	3.2%	10.7%
	Total	61.9%	50.0%	25.4%	26.8%	8.9%	12.5%	3.8%	10.7%
Fraction of 586 subjects	*Gain*	100.0%	100.0%	95.5%	100.0%	82.4%	94.0%	38.2%	64.8%
	Lottery	100.0%	100.0%	94.3%	100.0%	72.4%	94.0%	36.5%	64.8%
	Loss	100.0%	100.0%	93.3%	100.0%	65.1%	94.0%	24.6%	64.8%
	Total	100.0%	100.0%	94.4%	100.0%	73.4%	94.0%	33.1%	64.8%

Table 5: Observed *(Obs.)* and expected *(Exp.)* percentages by treatment and overall (1) of reports with respective WINLAG (upper panel) and (2) of participants with at least one respective WINLAG (lower panel). Excluding $t = 1$, where WINLAG is always 0 by definition. Expected percentages derived by generating all possible binary sequences of length 14 and counting respective incidences.

		Coeff.	Std. Error	Wald χ^2	Pr> χ^2
Gain	Intercept	0.19**	0.08	2.31	0.02
(n=2985)	WINLAG=1	0.06	0.09	0.67	0.50
	WINLAG=2	0.41***	0.14	2.86	0.00
	WINLAG>2	0.57**	0.23	2.52	0.01
	ROUND	0.01	0.01	1.48	0.14
Lottery	Intercept	0.28***	0.08	3.39	0.00
(n=2880)	WINLAG=1	-0.02	0.11	-0.17	0.87
	WINLAG=2	0.27*	0.14	1.94	0.05
	WINLAG>2	0.65***	0.22	3.00	0.00
	ROUND	0.02***	0.01	3.13	0.00
Loss	Intercept	0.48***	0.08	5.97	0.00
(n=2925)	WINLAG=1	0.22**	0.11	2.00	0.05
	WINLAG=2	0.27	0.17	1.63	0.10
	WINLAG>2	-0.24	0.21	-1.13	0.26
	ROUND	0.01*	0.01	1.79	0.07
Full model	Intercept	0.16***	0.06	2.63	0.01
(n=8790)	Lottery	0.16**	0.07	2.24	0.03
	Loss	0.30***	0.07	4.14	0.00
	WINLAG=1	0.08	0.06	1.39	0.16
	WINLAG=2	0.33***	0.09	3.77	0.00
	WINLAG>2	0.36***	0.13	2.71	0.01
	ROUND	0.02***	0.00	3.71	0.00

$^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$

Table 6: Logistic regression with dependent variable WIN_t^i, by treatment and overall. Standard errors clustered at the participant level.

While there are comparably few observations with WINLAG>2 overall (absolute numbers in Figure 3 below graphs and relative numbers in Table 5), the fraction of participants exhibiting such a WINLAG at least once in the experiment is 33.1%. This number is lower than the expected 64.8% under full honesty, but still substantial. In line with the observed shift towards higher reports in the *Lottery* and *Loss* frames, the number of subjects with at least one WINLAG>2 also decreases under these frames: From 38.2% in *Gain* to 36.5% in *Lottery* and 24.6% in *Loss*. We conclude:

Result 3 The estimated proportion of dishonest reports increases for WINLAG$=2$ and WINLAG>2 in the *Gain* and *Lottery* treatments, lending partial support to H3a. Dishonesty in the *Loss* treatment increases earlier than in the other frames (namely for WINLAG$=1$) and is not significantly different from WINLAG$=0$ for WINLAG>1, supporting H3c. Overall, we find strong patterns of path dependency, which differ between the *Gain* and *Lottery* frames on the one hand and the *Loss* frame on the other hand.

4.4. Does moral failure stick?

If path dependent behavior were really driven by a moral accounting-like effect as conjectured in section 3, we would expect that after building up moral credit, dishonesty would occur as shown in the previous subsection: The longer we observe unfavorable outcomes, the more likely is moral failure in the subsequent report. On the other hand, in truly accounting-like behavior, a dishonest participant should return to honest reporting and should start building up moral credit again after having used previously accumulated moral credit by defaulting. This would imply a lower likelihood of dishonesty after one (or several) dishonest reports.

To investigate whether such patterns are visible in our data, we look at the reported outcome immediately following reports with $\text{WIN}_t^i = 1$. In other words, we restrict our analysis to observations of $\text{WIN}_{t+1}^i | \text{WIN}_t^i = 1$. If moral failure was sticky (in line with the ego-depletion argument, but in contrast to the moral accounting argument in section 3), we would expect the proportion of favorable reports to be higher immediately following a reported favorable outcome. Panel (b) in Figure 3 suggests that this is generally not the case. After reporting Heads, the average probability of a favorable report is constantly high in the cases where WINLAG $\in \{0, 1, 2\}$. Surprisingly, it is almost at truthful outcome levels for WINLAG > 2 under all frames (a binomial test for difference from 0% results in p$\gg 0.1$ for all frames). The logistic regressions in Table 7 confirm these findings. This is consistent with the notion that those participants who allow a WINLAG>2 to occur are in general the more honest participants. Nevertheless, they still react to WINLAG, if only later than the others. We interpret this finding as evidence for moral accounting at least in this group, as after moral default these participants return to honest levels again. In total, we find:

Result 4 WINs immediately following favorable outcomes with WINLAG $\in \{0, 1, 2\}$ display no clear pattern. For all treatments, there is a sharp decrease in dishonesty for WINs immediately following favorable outcomes with WINLAG>2 to statistically honest levels.

		Coeff.	Std. Error	Wald χ^2	Pr> χ^2
Gain	Intercept	0.12	0.12	0.99	0.32
(n=1631)	WINLAG=1	0.17	0.15	1.12	0.26
	WINLAG=2	0.07	0.19	0.39	0.70
	WINLAG>2	-0.50**	0.26	-1.97	0.05
	ROUND	0.02*	0.01	1.84	0.07
Lottery	Intercept	0.33***	0.12	2.83	0.00
(n=1668)	WINLAG=1	-0.13	0.15	-0.87	0.38
	WINLAG=2	0.11	0.19	0.57	0.57
	WINLAG>2	-0.54**	0.26	-2.12	0.03
	ROUND	0.03***	0.01	2.60	0.01
Loss	Intercept	0.42***	0.11	3.81	0.00
(n=1780)	WINLAG=1	0.07	0.14	0.52	0.60
	WINLAG=2	-0.03	0.21	-0.17	0.87
	WINLAG>2	-0.46	0.32	-1.44	0.15
	ROUND	0.02**	0.01	2.35	0.02
Full model	Intercept	0.13	0.09	1.52	0.13
(n=5079)	Lottery	0.19*	0.10	1.88	0.06
	Loss	0.28***	0.10	2.87	0.00
	WINLAG=1	0.04	0.09	0.46	0.65
	WINLAG=2	0.05	0.11	0.47	0.64
	WINLAG>2	-0.51***	0.16	-3.23	0.00
	ROUND	0.02***	0.01	3.86	0.00

$^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$

Table 7: Logistic regression with dependent variable WIN^i_{t+1} if $\text{WIN}^i_t = 1$, by treatment and overall. Standard errors clustered at the participant level.

This lends support to H3b only for a subgroup of participants who are more willing to be honest, and therefore endure longer streaks of unfavorable outcomes in general.

5. Discussion and conclusion

In our experiment, we find clear behavioral evidence for temporal patterns in sequences of honesty decisions. Furthermore, we observe that some of these patterns are affected by frames: In line with the existing literature, we observe that dishonesty is lowest when the situation is framed as a gain, higher when it is framed as a lottery and highest when it is framed as a loss. Timing-wise, we find a statistically significant, yet in terms of outcomes relatively small increase of dishonesty over time under the *Gain* and *Lottery* frames, but not under the *Loss* frame. The observed timing effects seem, however, to be driven mainly by a cheating-at-the-end effect that is visible under all frames, where a strong increase in dishonesty only towards the very end of the sequence leads to the diagnosis of an increase

over time. While our results therefore do not exhibit strong patterns of honesty erosion over time, we find a clear increase in dishonesty after longer streaks of unfavorable reports, particularly under the *Gain* and *Lottery* frames. A specific pattern is visible for very honest participants (those who accept longer streaks of unfavorable outcomes), who in the round after a positive outcome report honestly, as far as statistics can tell.

While the suitability of these factors and the reference point for favorable and unfavorable outcomes have to be verified for different situations and circumstances, we believe that our findings should diminish the feeling that such indicators are "unfair" per se. They instead show that inferences about the likelihood of dishonesty based on such longitudinal observations may be well-founded (in a statistical sense). We are thus confident that our study can provide valuable new arguments for a nuanced discussion of the use of statistical instruments in the pursuit of uncovering dishonest behavior. While the detection of an uncommon pattern should not immediately be cause for charges of lying or cheating, it may constitute justifiable grounds for increased scrutiny of the case in question.

References

Abeler, J., Raymond, C., Nosenzo, D., 2016. Preferences for Truth-Telling. CESifo Working Paper 6087.

Arechar, A. A., Gächter, S., Molleman, L., 2018. Conducting interactive experiments online. Experimental Economics 21 (1), 99–131.

Blanco, C., Ezquerra, L., Rodriguez-lara, I., 2015. Incentives to cheat under loss aversion. Working Paper.

Callaway, E., 2012. Why great Olympic feats raise suspicions. Nature News.

Cohn, A., Maréchal, M. A., 2017. Laboratory Measure of Cheating Predicts Misbehavior at School. The Economic Journal. Forthcoming.

Dai, Z., Galeotti, F., Villeval, M. C., 2017. Cheating in the Lab Predicts Fraud in the Field. An Experiment in Public Transportations. Management Science. Forthcoming.

Dohmen, T., Falk, A., Huffman, D., Sunde, U., Schupp, J., Wagner, G., 2011. Individual risk attitudes: Measurement, determinants, and behavioral consequences. Journal of the European Economic Association 9 (3), 522–550.

Effron, D., Miller, D., Monin, B., 2012. Inventing racist roads not taken: the licensing effect of immoral counterfactual behaviors. Journal of Personality and Social Psychology 103 (6), 916–932.

Effron, D. A., Bryan, C. J., Murnighan, J. K., 2015. Cheating at the End to Avoid Regret. Journal of Personality and Social Psychology 109 (3), 395–414.

Engström, P., Nordblom, K., Ohlsson, H., Persson, A., 2015. Tax compliance and loss aversion. American Economic Journal: Economic Policy 7 (4), 132–164.

Fischbacher, U., Föllmi-Heusi, F., 2013. Lies in disguise-an experimental study on cheating. Journal of the European Economic Association 11 (3), 525–547.

Giatti, L., Camelo, L. d. V., Rodrigues, J. F. d. C., Barreto, S. M., 2012. Reliability of the MacArthur scale of subjective social status - Brazilian Longitudinal Study of Adult Health (ELSA-Brasil). BMC Public Health 12 (1096).

Gino, F., Bazerman, M. H., 2009. When misconduct goes unnoticed: The acceptability of gradual erosion in others' unethical behavior. Journal of Experimental Social Psychology 45 (4), 708–719.

Gino, F., Schweitzer, M. E., Mead, N. L., Ariely, D., 2011. Unable to resist temptation: How self-control depletion promotes unethical behavior. Organizational Behavior and Human Decision Processes 115 (2), 191–203.

Gneezy, U., Imas, A., Madarász, K., 2014. Conscience Accounting: Emotion Dynamics and Social Behavior. Management Science 60 (11), 2645–2658.

Gneezy, U., Kajackaite, A., Sobel, J., 2018. Lying Aversion and the Size of the Lie. The American Economic Review 108 (2), 419–453.

Grolleau, G., Kocher, M. G., Sutan, A., 2016. Cheating and Loss Aversion: Do People Cheat More to Avoid a Loss? Management Science 62 (12), 3428–3438.
Holm, S., 1979. A Simple Sequentially Rejective Multiple Test Procedure. Scandinavian Journal of Statistics 6 (2), 65–70.
Horton, J. J., Rand, D. G., Zeckhauser, R. J., 2011. The Online Laboratory Conducting Experiments in a Real Labor Market. Experimental Economics 14 (3), 399–425.
Jacobsen, C., Fosgaard, T. R., Pascual-Ezama, D., 2018. Why Do We Lie? A Practical Guide To the Dishonesty Literature. Journal of Economic Surveys 32 (2), 357 – 387.
Mazar, N., Amir, O., Ariely, D., 2008. The Dishonesty of Honest People: A Theory of Self-Concept Maintenance. Journal of Marketing Research 45 (6), 633–644.
Mead, N. L., Baumeister, R. F., Gino, F., Schweitzer, M. E., Ariely, D., 2009. Too Tired to Tell the Truth: Self-Control Resource Depletion and Dishonesty. Journal of Experimental Social Psychology 45 (3), 594–597.
Merritt, A. C., Effron, D. A., Monin, B., 2010. Moral Self-Licensing: When Being Good Frees Us to Be Bad. Social and Personality Psychology Compass 4 (5), 344–357.
Moshagen, M., Hilbig, B. E., 2017. The statistical analysis of cheating paradigms. Behavior Research Methods 49 (2), 724–732.
Palan, S., Schitter, C., mar 2018. Prolific.acA subject pool for online experiments. Journal of Behavioral and Experimental Finance 17, 22–27.
Pocheptsova, A., Amir, O., Dhar, R., Baumeister, R. F., 2009. Deciding Without Resources: Resource Depletion and Choice in Context. Journal of Marketing Research 46 (3), 344–355.
Schindler, S., Pfattheicher, S., 2016. The frame of the game: Loss-framing increases dishonest behavior. Journal of Experimental Social Psychology 69, 172–177.
Schwartz Cameron, J., Miller, D. T., Monin, B., 2008. Deservingness and unethical behavior in loss and gain frames. Working Paper.
Tice, D. M., Baumeister, R. F., Shmueli, D., Muraven, M., 2007. Restoring the self: Positive affect helps improve self-regulation following ego depletion. Journal of Experimental Social Psychology 43, 379–384.
Touré-Tillery, M., Fishbach, A., 2012. The end justifies the means, but only in the middle. Journal of Experimental Psychology: General 141 (3), 570–583.
Tversky, A., Kahneman, D., 1981. The Framing of Decisions and the Psychology of choice. Science 211 (4481), 453–458.
Weisel, O., Shalvi, S., aug 2015. The collaborative roots of corruption. Proceedings of the National Academy of Sciences of the United States of America 112 (34), 10651–6.
Zhong, W., Wu, H., Li, L., 2012. Olympics: Some facts about Ye Shiwen's swim. Nature Correspondence 488, 459.

Appendix A. Experimental instructions

Appendix A.1. Invitation to experiment

Title of invitation:

Financial Decision Making (bonus reward up to 1.50)

Study description in invitation:

To participate in this study, you need to have a coin on hand. Please make sure you have a coin ready before beginning. In this study you can earn bonus payments of up to 1.50!

Appendix A.2. Gain Frame instructions

PLEASE READ THESE INSTRUCTIONS CAREFULLY!

You will get a fixed payment of 0.60 for completing all sections of this survey.

In Part 1, you take part in a lottery and will get an extra payment depending on several rounds of coin flips: You will be asked to flip a coin once a round and report the outcome. Each round in which you report heads, 0.10 will be added to this extra payment.

We do not tell you in advance exactly how many rounds of coin flipping there will be, but the maximum possible is 15 rounds. The exact number of rounds was determined before the survey started and is unaffected by your reported outcomes.

Please click Next to start with the coin flips.

Appendix A.3. Lottery Frame instructions

PLEASE READ THESE INSTRUCTIONS CAREFULLY!

You will get a fixed payment of 0.60 for completing all sections of this survey.

In Part 1, you take part in a lottery and will get an extra payment depending on several rounds of coin flips: You will be asked to flip a coin once a round and report the outcome. The starting value of the extra payment is 0.75. Every round, 0.05 will be deducted from this extra payment as a participation fee. However, each round in which you report heads, 0.10 will be added to this extra payment.

We do not tell you in advance exactly how many rounds of coin flipping there will be, but the maximum possible is 15 rounds. The exact number of rounds was determined before the survey started and is unaffected by your reported outcomes.

Please click Next to start with the coin flips.

Appendix A.4. Loss Frame instructions

PLEASE READ THESE INSTRUCTIONS CAREFULLY!

You will get a fixed payment of 0.60 for completing all sections of this survey.

In Part 1, you take part in a lottery and will get an extra payment depending on several rounds of coin flips: You will be asked to flip a coin once a round and report the outcome. The starting value of the extra payment is 1.50. Each round in which you report tails, 0.10 will be deducted from this extra payment.

We do not tell you in advance exactly how many rounds of coin flipping there will be, but the maximum possible is 15 rounds. The exact number of rounds was determined before the survey started and is unaffected by your reported outcomes.

Please click Next to start with the coin flips.

Appendix B. Regression including controls

Besides the main experiment, we collected additional information in a questionnaire and received some additional participant demographics on some of the participants from the platform Prolific.ac. From these, we use selected demographic and questionnaire answers as controls for our key regression analysis. As Prolific.ac does not have all demographic information on all their participants, we excluded those with missing answers in this analysis. Regarding demographics, we use gender in dummy variable FEMALE; AGE in years; whether a participant is NO_STUDENT; and employment status using dummy variables

FULL_TIME, PART_TIME, UNEMPLOYED and OTHER_EMPLOYMENT. We also include SOCIAL_STATUS, which is self-reported on a scale of 1-10 via the MacArthur ladder,[16] and RISK_TOLERANCE, self-reported on an 11-point Likert scale (from 0 or "not at all willing to take risks" to 10 or "very willing to take risks") following Dohmen et al. (2011). We also asked participants to report how much money they earned in the coin tossing part to see whether participants with higher scores pay more attention to their payoff. EARNING_ESTIMATE is the absolute difference between their estimate and their actual payoff. Finally, we received information, from Prolific.ac, on the number of studies each participant had already taken part in before our experiment, which we included in the variable EXPERIENCE to see whether participant experience also explains some of the effects we observe.

In total, the identified effects from the paper's main part are not materially affected by adding controls. Specifically, the explanatory value of WINLAG is very consistent even after adding our controls. Only the difference between the Gain and the Loss treatments changes from significant to insignificant, which we explain by reduced degrees of freedom after adding controls.

Interestingly, better knowledge of total payoff seems to increase the likelihood of a high outcome (except for the Loss treatment), consistent with our assumption that participants paying closer attention to payoffs are more likely to be dishonest. Given that this is not informative for our research questions, however, we do not include it in our analysis in the main part, but report it here as a side result.

Appendix C. Regression on datasets excluding likely dishonest participants

To understand honesty behavior of "ordinary" people (as Mazar et al. 2008 put it), we would like to focus on the behavior of participants who are not fully dishonest in our experiment, but rather are conditionally dishonest. We therefore run our key regression on three reduced datasets, in which we exclude likely dishonest participants as identified by a statistical test ex-post. Table C.9 displays the key regression of this paper for all datasets considered. First, "Full" includes the same observations as in the remainder of the paper. Second, "Max" excludes all participants with 15 favorable outcomes. The likelihood for this to happen is about 3 in 100,000, while we observe (over all treatments) about 5,400 in 100,000 (or 5.4%). We therefore find it useful to look at a dataset excluding these (full) liars. Next, we calculate the 95% confidence interval of Heads a participant will throw in our experiment, which is 11 Heads. We observe 25.9% of all participants having a

[16]The MacArthur ladder is a scale commonly used in health research, see e.g. Giatti et al. (2012). Participants are shown a ladder with steps numbered from 1 (at the bottom) to 10 (at the top) and are asked: "Think of a ladder (see image) as representing where people stand in society. At the top of the ladder are the people who are best off – those who have the most money, most education and the best jobs. At the bottom are the people who are worst off – who have the least money, least education and the worst jobs or no job. The higher up you are on this ladder, the closer you are to people at the very top and the lower you are, the closer you are to the bottom. Where would you put yourself on the ladder? Choose the number whose position best represents where you would be on this ladder."

		Coeff.	Std. Error	Wald χ^2	Pr> χ^2
Full model	Intercept	-0.13	0.22	-0.58	0.56
incl. controls	Lottery	0.26***	0.10	2.77	0.01
(n=5505)	Loss	0.11	0.08	1.29	0.20
	WINLAG=1	0.09	0.07	1.22	0.22
	WINLAG=2	0.47***	0.10	4.56	0.00
	WINLAG>2	0.52***	0.16	3.27	0.00
	ROUND	0.01**	0.01	2.44	0.01
	RISK_TOLERANCE	-0.02	0.02	-0.98	0.33
	FEMALE	-0.07	0.07	-1.00	0.32
	AGE	0.01	0.00	1.63	0.10
	PART_TIME	-0.17*	0.09	-1.91	0.06
	UNEMPLOYED	-0.26***	0.10	-2.62	0.01
	OTHER_EMPLOYMENT	-0.07	0.12	-0.58	0.56
	NO_STUDENT	-0.10	0.09	-1.10	0.27
	SOCIAL_STATUS	0.01	0.02	0.59	0.56
	EXPERIENCE	-0.00	0.00	-0.20	0.84
	EARNING_ESTIMATE	0.35***	0.12	2.91	0.00
Full model	Intercept	0.16***	0.06	2.63	0.01
(n=8790)	Lottery	0.16**	0.07	2.24	0.03
	Loss	0.30***	0.07	4.14	0.00
	WINLAG=1	0.08	0.06	1.39	0.16
	WINLAG=2	0.33***	0.09	3.77	0.00
	WINLAG>2	0.36***	0.13	2.71	0.01
	ROUND	0.02***	0.00	3.71	0.00

$^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$

Table B.8: Logistic regression with dependent variable WIN_t^i. Standard errors clustered on participant level. Full model excluding controls added as reference.

greater number than 11 HEADS, when we would expect to see only 5% assuming honesty. The dataset excluding these observations is called "95%". We also use a Runs test at the participant level to investigate whether the series of reported HEADS and TAILS are likely to be random. Excluding those with a 90% probability of not having reported a random sequence results in dataset "Runs Test". Finally, we define the dataset "ExAll" by combining all exclusion rules defined before. This reduces the number of observations compared to the individual exclusion rules even further, as there are 106 participants who violate the 95% threshold (i.e., reporting more than 11 times HEADS), who at the same time have a random sequence with 90% probability, and also 28 participants who have no random sequence with 90% probability, but do not violate the 95% threshold.

Overall, WINLAG is a consistent predictor for higher dishonesty. It is higher the more categories of (likely) dishonest participants are excluded, e.g., 95% and ExAll. ROUND, on the other hand, becomes insignificant for those groups, and so does the difference in frames.

In total, this analysis corroborates our findings from the main part, namely that (1) differences are caused by both a strong shift of distributions to the right and an increase in full claimers combined, (2) effects from ROUND are not strong, and (3) WINLAG is a clear predictor of increased likelihood of dishonesty.

		Coeff.	Std. Error	Wald χ^2	Pr> χ^2
Full model	Intercept	0.16***	0.06	2.63	0.01
(n=8790)	Lottery	0.16**	0.07	2.24	0.03
	Loss	0.30***	0.07	4.14	0.00
	WINLAG=1	0.08	0.06	1.39	0.16
	WINLAG=2	0.33***	0.09	3.77	0.00
	WINLAG>2	0.36***	0.13	2.71	0.01
	ROUND	0.02***	0.00	3.71	0.00
Max	Intercept	0.03	0.06	0.55	0.58
(n=8310)	Lottery	0.14**	0.06	2.16	0.03
	Loss	0.27***	0.07	4.19	0.00
	WINLAG=1	0.23***	0.05	4.21	0.00
	WINLAG=2	0.47***	0.08	5.67	0.00
	WINLAG>2	0.50***	0.13	3.88	0.00
	ROUND	0.02***	0.00	3.40	0.00
95%	Intercept	-0.12**	0.05	-2.14	0.03
(n=6510)	Lottery	0.07	0.05	1.36	0.17
	Loss	0.13**	0.05	2.46	0.01
	WINLAG=1	0.36***	0.06	6.22	0.00
	WINLAG=2	0.73***	0.08	8.90	0.00
	WINLAG>2	0.83***	0.13	6.58	0.00
	ROUND	0.00	0.01	0.46	0.65
Runs Test	Intercept	0.05	0.05	0.88	0.38
(n=7680)	Lottery	0.12*	0.06	1.84	0.07
	Loss	0.28***	0.07	4.19	0.00
	WINLAG=1	0.20***	0.05	3.88	0.00
	WINLAG=2	0.48***	0.08	5.72	0.00
	WINLAG>2	0.58***	0.13	4.38	0.00
	ROUND	0.01***	0.00	2.58	0.01
ExAll	Intercept	-0.08	0.06	-1.37	0.17
(n=6090)	Lottery	0.06	0.05	1.16	0.24
	Loss	0.11**	0.06	2.01	0.04
	WINLAG=1	0.24***	0.05	4.51	0.00
	WINLAG=2	0.70***	0.08	8.50	0.00
	WINLAG>2	0.86***	0.13	6.71	0.00
	ROUND	0.00	0.01	0.39	0.70

$^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$

Table C.9: Logistic regression models with dependent variable WIN_t^i on full and reduced data sets. Full includes all observations as in the main part; Max excludes all participants reporting 15 WINs; 95% excludes those who report above 11 WINs (for which there is a 95% chance to stay below if all reports are honest); Runs Test excludes all who reported a non-random sequence at the 90% significance level; ExAll excludes reports that are above 11 WINs and/or who failed the Runs test at 90% significance. Standard errors clustered on participant level.