MENTAL MONEY LAUNDERING: A MOTIVATED VIOLATION OF FUNGIBILITY Q5

Alex Imas
Booth School of Business, University of Chicago, USA

George Loewenstein
Social and Decision Sciences, Carnegie Mellon University, USA

Carey K. Morewedge
Questrom School of Business, Boston University, USA

Abstract
People exploit flexibility in mental accounting to relax psychological constraints on spending. Four studies demonstrate this in the context of moral behavior. The first study replicates prior findings that people donate more money to charity when they earned it through unethical versus ethical means. However, when the unethically earned money is first "laundered"—the cash is physically exchanged for the *same* amount but from a different arbitrary source—people spend it as if it was earned ethically. This *mental money laundering* represents an extreme violation of fungibility. The second study demonstrates that mental money laundering generalizes to cases in which ethically and unethically earned money are mixed. When gains from ethical and unethical sources were pooled, people spent the entire pooled sum as if it was ethically earned. The last two studies provide mixed support for the prediction that people actively seek out laundering opportunities for unethically earned money, suggesting partial sophistication about these effects. These findings provide new evidence for the ease with which people can rationalize misbehavior, and have implications for consumer choice, corporate behavior, and public policy. (JEL: D01, D03, D64)

1. Introduction

Each year, billions of dollars obtained through unethical means are exchanged for "clean money" through the process of money laundering. The usual motive for money laundering is practical (Van Duyne 2002). Criminals use money laundering to avoid Q6 being caught and penalized for illegal behavior. We suggest an analog in the mind that

The editor in charge of this paper was Imran Rasul.

Acknowledgments: We thank Ainslinn Bohren for helpful comments on the manuscript and Christy Kang and Chang-Yuan Lee for research assistance.

E-mail: alex.imas@chicagobooth.edu (Imas); GL20@andrew.cmu.edu (Loewenstein); morewedge@gmail.com (Morewedge)

Journal of the European Economic Association 2021 0(0):1–25 DOI: 10.1093/jeea/jvaa059

people exploit flexibility in mental accounting (Thaler 1985, 1999) to avoid similar self-imposed sanctions. People often penalize themselves when spending money they obtained in an unethical fashion. Unethically earned money is spent more prosocially, in less purely self-interested ways, than money earned ethically (Gneezy, Imas, and Madarász 2014; Tasimi and Gelman 2017). Virtuous spending can act as a salve for guilt arising from unethical gains (Cryder, Springer, and Morewedge 2012), but it comes at the opportunity cost of less spending on the self. Our findings suggest that people engage in *mental money laundering* to alleviate psychological constraints on the spending of earnings that are associated with undesirable sources.

In demonstrating mental money laundering, we first replicate the prior finding that people are more likely to use unethically earned money (obtained through lying) prosocially than the same amount ethically earned (obtained through truth telling). Then, in the most straightforward demonstration of mental money laundering, we show that when the unethically earned money is physically exchanged for the same amount of cash from an arbitrary source—when the "dirty" money is wagered in a fair lottery that returns the same sum in "clean" bills—the money is spent as if it was ethically earned. Third, we demonstrate the generalizability of this laundering effect, showing that it extends to situations in which earnings from multiple sources are mixed: Pooling money from unethical and ethical sources in a single envelop leads people to spend the entire sum of cash as if it was ethically earned—even if more than half of the pooled amount was earned unethically. Finally, we explore the extent to which people are aware of and exploit mental money laundering ex ante, and find mixed evidence for this type of sophistication.

1.1. Mental Money Laundering

Mental money laundering represents a violation of the *fungibility* principle, which posits that the source of money should not influence how it is used. Research on mental accounting has shown that both the source of money (e.g. winnings from a betting pool vs. wage income) and how it is labeled (e.g. "child benefits") influence how it is spent (Hines and Thaler 1995; Heath and Soll 1996; O'Curry 1997; Kooreman 2000; Epley, Mak, and Idson 2006; Abeler and Marklein 2017).

One type of source effect involves the ethicality of how money is earned. People spend money earned from ethical and unethical sources in different ways. Windfall gains received from questionable sources (e.g. tobacco companies) are more likely to be spent on utilitarian goods such as school supplies than on hedonic goods like ice cream (Levav and McGraw 2009). Money earned through deception or thievery is more likely to be spent on charitable giving than the equivalent amount earned through honest means—at least by those who are not naturally inclined toward these activities (Gneezy, Imas, and Madarász 2014; Tasimi and Gelman 2017). It has been argued that evoked feelings of guilt drive the observed increase in prosocial behavior (Gneezy et al. 2014; Cryder, Springer, and Morewedge 2012). However, studies that replicate these effects have also shown that the ethicality of how money is sourced does not affect prosociality across domains. Choosing unethical methods of earning

Q7

cash, for example, does not make people more likely to volunteer their *time* (Park and Meyvis 2011). This hints at the possibility that the compensatory behavior is not driven by a general motive to assuage guilt but is specific to the money itself. It appears to be driven by negative associations between the money and its source.

Source effects are believed to emerge through the psychological association between money and different mental accounts (Hines and Thaler 1995; Prelec and Loewenstein 1998). As in conventional accounting, a particular set of earnings and prospective expenses can often be booked to a variety of accounts. Booking funds to a particular mental account affects how it is spent. Money associated with an "entertainment" account is more likely to be used for dining at a restaurant, for example, than money associated with "savings" or "household expense" accounts. To date, mental accounting theories suggest that people use heuristics—that is, rules of thumb—to allocate money to different accounts (Zhang and Sussman 2018). Q8 The specific heuristics proposed include adaptation (Baucells and Hwang 2017), availability (Morewedge, Holtzman, and Epley 2007), and similarity judgments (Henderson and Peterson 1992), which all predict a direct relationship between the source of earnings and the account to which they are booked. However, in some cases, this relationship may be ambiguous. For example, an academic might mentally assign an unusually expensive dinner to regular income, earnings from savings, or consulting fees. Given this flexibility, what factors determine how actual earnings and prospective expenses are assigned between mental accounts?

Studies suggest that people may engage in motivated mental accounting, assigning ambiguously sourced earnings to accounts that facilitate desired purchases (e.g. Cheema and Soman 2006; Sussman and Alter 2012). We propose, analogously, that people value and exploit flexibility in mental accounting to dissociate money from undesirable sources. This source dissociation lifts constraints to spend the money virtuously, allowing people to spend it in more selfish ways. In essence, people engage in motivated mental accounting by taking advantage of the opportunity to obfuscate the source of unethical gains so that they can spend the money unencumbered by moral constraints.

The idea that people exploit malleability in mental accounting to behave selfishly is complementary to prior work that demonstrates the willingness to take advantage of "moral wiggle room" when it is available (Dana, Weber, and Kuang 2007; Exley 2016). People use a diverse set of strategies to get themselves "off the hook", to liberate themselves to behave selfishly while maintaining a positive moral identity (Bénabou, Falk, and Tirole 2018; Bénabou and Tirole 2011). Tactics for exploiting moral wiggle room include (mis)interpreting information, such as persuading oneself that self-interested advice one gives to another person is actually best for them (Gneezy et al. 2020). People may also seek or avoid social interactions (Lazear, Malmendier and Weber 2012; Dana, Cain, and Dawes 2006), such as using a different store entrance to avoid solicitations for charity (Andreoni, Rao, and Trachtman 2017).

We begin by presenting a simple theoretical framework of mental money laundering, which we use to interpret and rationalize the empirical results. The model generates three predictions that (1) that people will be more likely to spend

money prosocially if it is associated with an unethical mental account versus an ethical one; (2) mental money laundering breaks this association and leads to less prosocial spending from unethically-earned gains; and (3) those sophisticated about the laundering mechanism will value and seek out opportunities to do so ex ante. Four experiments—described earlier in this section—test these predictions.

Mental money laundering has implications for the design of compensation schemes, policy interventions, and emerging financial technologies. Firms may respond to the moral qualms of their employees or consumers by bundling earnings or purchases with prosocial elements. For example, a company producing ethically questionable products may be able to pay a smaller wage premium to workers who are otherwise morally opposed to its practices by also engaging in corporate social responsibility (CSR) (e.g. "greenwashing"; Nyilasy, Gangadharbatla, and Paladino 2014). Since earnings are pooled, workers can associate them with the positive actions of the firm rather than the negative actions, which lifts constraints on their behavior, including compunctions about working at the company in the first place. Financial software applications that enable consumers to create and label digital "accounts" may make it easier for people to engage in mental money laundering. Given the ubiquity of pooling income in modern electronic bank accounts (e.g. dividends from "green" and "brown" firms), mental money laundering is likely to be a widespread phenomenon. As discussed further in Section 4, policymakers and firms seeking to leverage spending constraints imposed by source effects should take the malleability of mental accounting into account, distributing funds through methods that hinder mental money laundering (e.g. using benefit-specific credit cards).

The paper proceeds as follows. Section 2 outlines the theoretical framework and derives predictions. Section 3 presents results from four experiments that demonstrate mental money laundering and examine people's sophistication about the mechanism. Section 4 discusses implications and concludes the paper.

2. Conceptual Framework

Our framework partly builds on the model of altruism under moral constraints outlined in Gneezy, Imas, and Madarász (2014) while incorporating features of mental accounting. This allows us to more directly connect our proposed mechanism to the literature on dynamic moral decision-making. We present the general framework here and derive predictions for each of our empirical settings in the sections that follow.

An individual makes a sequence of allocation decisions between herself and others in periods $t = 1,\ 2,\ 3$ and consumes the final allocation in $T = 4$. Let $c, d \in \mathbb{R}^+$ correspond to her own consumption and the consumption of others, respectively. In $t = 3$ the decision-maker (henceforth, DM) maximizes utility by choosing an allocation $\{c, d\}$ from her current wealth W_O. Her utility from this allocation can be expressed as

$$u(c, d, g^t),$$

where u is a continuous, twice-differentiable function, and $g^t \in [0, 1]$ is a state variable corresponding to the fraction of wealth mentally linked to an unethical source through mental accounting. Psychologically, this variable captures the guilt generated by holding money that is associated with norm violations.

Wealth can be earned from multiple sources, which influences which mental account it is "booked" to (Hines and Thaler 1995). Here, we consider two mental accounts, one associated with ethical sources E and one associated with unethical sources UE. Earnings that are associated with the former are termed W_E^t and those associated with the latter are termed W_{UE}^t. Overall wealth can be represented as a sum of earnings associated with the ethical and unethical accounts, $W_O^t = W_E^t + W_{UE}^t$. Let $g^t = W_{UE}^t / (W_{UE}^t + W_E^t)$ capture the fraction of wealth assigned to the unethical mental account. We assume that money obtained through ethical (unethical) means is initially booked to account E (UE).

As discussed in the preceding section, individuals have been shown to have a different marginal propensity to consume from different mental accounts, which is a violation of fungibility. To capture the lower propensity to fund own consumption through unethically obtained earnings, we allow the marginal utility of own consumption to be a function of g^t, such that the greater the proportion of wealth is associated with unethical behavior, the lower the marginal utility of own consumption, $u_{cg} < 0$. We follow the literature in assuming that guilt is a negative emotion that has a direct effect on utility, namely, that holding "dirty" cash that is associated with an unethical source decreases utility, $u_g < 0$.

To formalize mental money laundering, consider a DM who arrives at $t = 2$ with wealth W_O^1, where $g^1 > 0$, that is, some of her earnings were attained unethically and booked to account UE. Earnings from the first period are then exchanged or mixed with wealth from a different source to yield an overall wealth of W_O^2. This sum may be the smaller, the same, or, as in the case of mixing, larger than the original amount. This new source is not linked to unethical practices, such that if cash from the source represented the DM's overall wealth, $g = 0$.[1]

The DM makes a decision about how to allocate funds between her own and others' consumption in the next period, from W_O^3, where $W_O^3 = W_O^2$. Which account does the DM associate these earnings with; that is, what is the g^3 that applies to the pooled funds? There are three possibilities: (1) associate the earnings with the original source, such that $g^3 = g^1$; (2) associate the earnings with the second source, such that $g^3 = 0$; or (3) the association of earnings with accounts is some convex combination of the two sources, such that

$$g^1 > g^3 > 0.$$

Motivated mental accounting predicts that, given multiple possibilities of mental accounting operations, the DM will choose the operation that maximizes her utility. It

1. The psychological mechanism through which the DM "launders" earnings, which changes the mental account they are associated with, is explored empirically in the sections that follow. Here, we consider the implications of laundering for behavior.

is straightforward to show that utility will be maximized if mental accounting follows the second possibility, $g^3 = 0$, because carrying dirty money has a direct negative effect on utility, $u_g < 0$; motivated mental accounting predicts that the DM will associate her entire wealth with the ethical source.

Having outlined the general framework, we now proceed to describe the empirical setup and to derive specific predictions from the model regarding how mental money laundering affects behavior in each setting.

3. Experimental Studies

3.1. Study 1: Mental Money Laundering

The first experiment examines the effects of mental money laundering on prosocial spending. Participants were paired, and one was assigned to the role of sender and the other the role of receiver. The experiment had three stages.

The first stage of the experiment followed a procedure adapted from Gneezy et al. (2014) and Gneezy (2005). In this two-player "deception game", one player, the sender, had private information, and the other, the receiver, made a choice based on a message conveyed by the sender. After the sender sent a message, the payoffs for both players depended on the choice the receiver made. This type of situation can be modeled using a cheap talk sender–receiver game (Crawford and Sobel 1982).

The game, which was used in all three experiments, had two possible monetary outcomes. Only the sender knew about the monetary outcomes of each option. The receiver had no information regarding the incentives paid by each option. Participants played the deception game in one of two conditions. In the lie incentives condition, incentives were not aligned so that the sender earned more money from one payoff option and the receiver earned more money from the other. In the truth incentives condition, incentives were aligned so that both the sender and receiver earned more money from one option than the other. After deciding on the message, the sender was paid in cash according to the choice made by the receiver.

In the second stage, some senders had the cash from the first stage routed through a lottery that had a high probability of returning the same amount that was staked in one of three experimental lottery conditions. Senders randomly assigned to the laundered lottery condition had their money entered into a lottery that had a high probability of returning the same amount, but as physically different bills. It exchanged potentially "dirty" money for "clean" money. Senders randomly assigned to the unlaundered lottery condition had their money entered into a similar lottery that had a high probability of the participant keeping the *same* physical bills that were staked. Senders randomly assigned to the no lottery condition did not have their money staked in a lottery.

In the third stage, all participants were given the option of donating part or all of their earnings to a charity.

Under the lie incentives condition, the exchange of cash in the laundered lottery condition was meant to increase ambiguity in the source of the earnings. The original

source of the money was unethical behavior, but the proximate source of the bills was the lottery. There was no ambiguity in the other two conditions: the source of the bills was unethical behavior. We predicted that in the laundered lottery condition, the increased scope for flexibility in mental accounting would allow participants to dissociate earnings from the unethical source and associate them with the lottery. Lifting the psychological constraints on spending should free the participants to give less to charity compared to those in the other two conditions.

Method. Participants ($N = 1{,}348$) were recruited in even-numbered groups (minimum of 4, maximum of 8) and assigned to laboratory cubicles on opposite sides of a room.[2]

The study was run in three waves. The first wave was the initial investigation, using a student population randomly assigned to four conditions: lie incentives with no lottery, lie incentives with laundered lottery, truth incentives with no lottery, and truth incentives with laundered lottery. The second wave, in addition to increasing our confidence in the (already significant) findings from the first study, introduced an unlaundered lottery condition to test whether the effects observed in the first study were due to the lottery per se or the fact that the lottery exchanged the staked amount for money from a different source. This second wave used a non-student sample from the community randomly assigned to three conditions, all under lie incentives: no lottery, laundered lottery, and unlaundered lottery. The third wave sought to replicate the pattern of findings using a student population randomly assigned to five conditions: lie incentives with all three lottery conditions and truth incentives with laundered lottery and unlaundered lottery. The first wave was conducted using paper-and-pencil instructions. The second and third waves were implemented on Qualtrics, with the instructions displayed on a computer monitor. A table depicting the phases of the experiment, in each condition, appears below.

In the first phase of the study, each participant on one side of the room was assigned the role of the sender and paired anonymously with a participant on the other side of the room, who was assigned to the role of receiver. Each pair then played a version of the deception game with ten possible messages (Gneezy et al. 2014).[3] The Q9
sender was randomly assigned a private number between 0 and 9 and was asked to choose one of 10 possible messages to send the receiver. The message was "Choosing __ will earn you more money than any other number", with the blank corresponding to a number from 0 to 9. The sender was informed that if the receiver chose a number that corresponded to the sender's private number, both players would be paid according

2. Experimental materials for all three experiments can be found in the Online Appendix and in the Open Science Framework depository at: https://osf.io/sq6vm/?view_only=88c7f93d26ce4f16b37eefd3c3a4a4f1

3. An early version of the game (Gneezy 2005) included only two possible messages, which introduced the possibility of sophisticated deception in the form of sending a truthful message with the intention to deceive (Sutter 2009). To preclude this concern, we follow Erat and Gneezy (2012) in expanding the message space to ten possible messages, with one message corresponding to the truth and nine messages corresponding to a lie.

to payment option Y, and if the receiver chose any other number, both players would be paid according to option X. Only the sender knew about the monetary outcomes of each option; the receiver had no information.

In the lie incentives condition, option Y paid the sender \$10 and the receiver \$20. Option X paid the sender \$20 and the receiver \$5. Hence, the sender earned more money and the receiver less money if the sender lied (i.e. sent a number other than his or her private number) and the receiver followed the suggestion implicit in the message (or chose any number other than the sender's private number). In the truth incentive condition, option Y earned both the sender and the receiver \$20. Option X earned both players \$5. If the receiver did follow the sender's advice, then a sender in the lie incentives condition had to lie to earn \$20 while a sender in the truth incentives condition had to tell the truth to earn the same amount. After the sender chose a message to send and the receiver made a choice, both players were paid their earnings in cash (in \$1 bills).

In the second phase, senders were randomly assigned to either the no lottery, laundered lottery, or unlaundered lottery condition. In the no lottery condition, the senders did not participate in a lottery and simply proceeded to the third phase. In the lottery conditions, senders chose two numbers between 1 and 12: a success number and a failure number. The outcome of the lottery was decided using a simple random number generator in the form of a digital 12-sided die that produced a number between 1 and 12. The die was rolled once for the entire group of senders participating in the experiment. If the number generated matched the sender's chosen success number, her earnings would double ($p = 0.083$). If it matched her failure number, she would lose her earnings ($p = 0.083$). If it matched neither number, her earnings would be returned to her ($p = 0.83$). In the laundered lottery condition, participants deposited their earnings into one box to enter the lottery, and the earnings from the lottery were given to the participants from a second box that contained a different physical source of money. Given the setup, most participants received the same amount they placed into the first box, but the physical money (i.e. the \$1 bills) returned to them was different.

To make sure that the observed effects are due to changing the source of the money rather than simply exposing money to risk, participants in the unlaundered lottery conditions were entered into the same lottery but it did not exchange the physical money. They did not place their earnings into a box when entering the lottery; the funds remained with them while the lottery was resolved. If the outcome of the lottery matched the participant's chosen success number, she was given the positive difference. If the outcome matched her failure number, she lost her earnings. Most important, if the outcome matched neither number, she simply kept her earnings. The key difference between the laundered and unlaundered lotteries was that most participants in the unlaundered lottery conditions would ultimately keep the money they directly earned by either lying (i.e. dirty money) or telling the truth, whereas those in the laundered lottery conditions had their earnings replaced with the same amount of money, but from a different source—the box of clean money associated with the lottery.

In the third phase of the study, after receiving their earnings in cash, all senders privately read a passage asking how much of their payment they would like donate to Save the Children, a charity chosen because of its non-partisan nature, ensuring that all,

or almost all, participants could be expected to view it as a virtuous and desirable target for funding. Senders then placed the amount they chose to donate into an envelope and sealed it. Care was taken such that neither the experimenter nor other participants could observe how much each sender donated. During this time, receivers were given an unrelated task to work on so they would not disrupt the decisions of the senders.

Predictions. We now derive predictions for the experimental setting using the framework outlined in Section 2. For simplicity, we assume that the DM's utility from this allocation can be expressed as

$$u(c, d, g) = c^{\beta(g)} d^{\alpha},$$

where c is her own consumption, d is the allocation to charity, $\beta + \alpha \leq 1$, and $\beta, \alpha \geq 0$. The DM maximizes utility subject to the budget constraint $W_O = c + d$. To capture $u_{cg} < 0$, we allow β to be a function of g, such that the greater the proportion of wealth associated with unethical behavior, the lower the marginal utility of own consumption, that is, the derivative $\beta_g < 0$ and hence the cross partial derivative $u_{cg} < 0$. This captures the psychological restriction on spending generated by the association of money with the unethical mental account: if the earnings were ethically obtained, the DM would prefer to donate less of it to charity. We now derive our first prediction on how prosocial behavior is affected by the proportion of wealth associated with this unethical mental account.

PREDICTION 1. *Prosocial allocations, d, will increase with g. The DM will allocate greater resources to others as more wealth becomes associated with the unethical mental account,* W_{UE}.

Proof. Solving for the first-order conditions, the optimal allocations of own and others' consumption are $c^* = W_O\beta/(\alpha + \beta)$ and $d^* = \frac{W_O\alpha}{\alpha+\beta}$, respectively. Since $\beta_g < 0$ by assumption, we have $d^*_g = -\beta_g W_O\alpha/(\alpha + \beta)^2 > 0$. □

Under this prediction, the DM will be more likely to donate to charity if, given the same amount of wealth W_O, a greater proportion is associated with the unethical mental account.

The second prediction concerns the effect of laundering on prosocial behavior. The lottery exchanges unethical gains W_O^1 for the same amount but from a different source; because earning money through a lottery is not unethical per se, the level of guilt associated with the laundered sum W_O^2 is smaller than the level associated with W_O^1, $g^1 > g^3$. In turn,

PREDICTION 2. *Mental laundering will reduce prosocial spending.*

Proof. Because $g^1 > g^3$, it follows trivially from Prediction 1 that d^* from W_O^1 will be higher than d^* from W_O^2. □

TABLE 1. Experimental design and treatments (in rows) by waves.

Incentive	Stage 1	Stage 2	Stage 3
Truth[1]	Sender–receiver game	No lottery	Charitable donation
Truth[3]	Sender–receiver game	Unlaundered lottery	Charitable donation
Truth[13]	Sender–receiver game	Laundered lottery	Charitable donation
Lie[123]	Sender–receiver game	No lottery	Charitable donation
Lie[23]	Sender–receiver game	Unlaundered lottery	Charitable donation
Lie[123]	Sender–receiver game	Laundered lottery	Charitable donation

Notes: Each row is a treatment. Waves in which each treatment was included are indicated by superscripts.

Together, these predictions imply a negative and significant effect of laundering on prosocial spending in the lie incentives conditions but no such effect in the truth incentives conditions.

Results

Messages Followed and Sent. In total, 76% of receivers followed the sender's messages, which is comparable to rates observed in previous studies: 78% in Gneezy (2005) and 75% in Gneezy et al. (2014). Removing pairs in which receivers did not follow senders' messages left 1,024 participants.

Of those senders, 61% sent a false message in the lie incentives conditions and 6% sent a false message in the truth incentives conditions. The rate of lying in the lie incentives conditions is comparable to rates observed in previous studies: 52% in Gneezy (2005) and between 60% and 75% in Gneezy et al. (2014, Table 1), depending on the condition. The rate of lying in the truth incentives conditions is lower than rates observed in previous studies (e.g. a surprisingly high 24% in Gneezy et al. 2014). Note that the incentives of both sender and receiver are aligned in the truth incentives condition, so we should expect the number of liars in the truth incentives conditions to be low if participants understand the incentives. A key difference between our experiments and those in Gneezy et al. (2014) is that our studies took place in the lab rather than a classroom, and each sender was matched with a receiver in the same session.[4] We made these changes to maximize trust and comprehension of the experimental instructions, which may account for the greater fraction of participants who, sensibly, told the truth in the truth incentives condition.

Prosocial Spending. In our main analysis, we focus on senders earning $20 across lie and truth incentive conditions ($N = 343$) after all potential losses and gains were tabulated (i.e. excluding participants in lottery conditions who lost or doubled their game earnings). First, we examine donation behavior of those who lied in the lie incentives conditions. We predicted that lying senders whose money was routed

4. In Gneezy et al. (2014), each sender had a 1 in 20 chance of being matched with a receiver in a different classroom, and hence had a 1 in 20 chance of being paid according to the outcomes of the game.

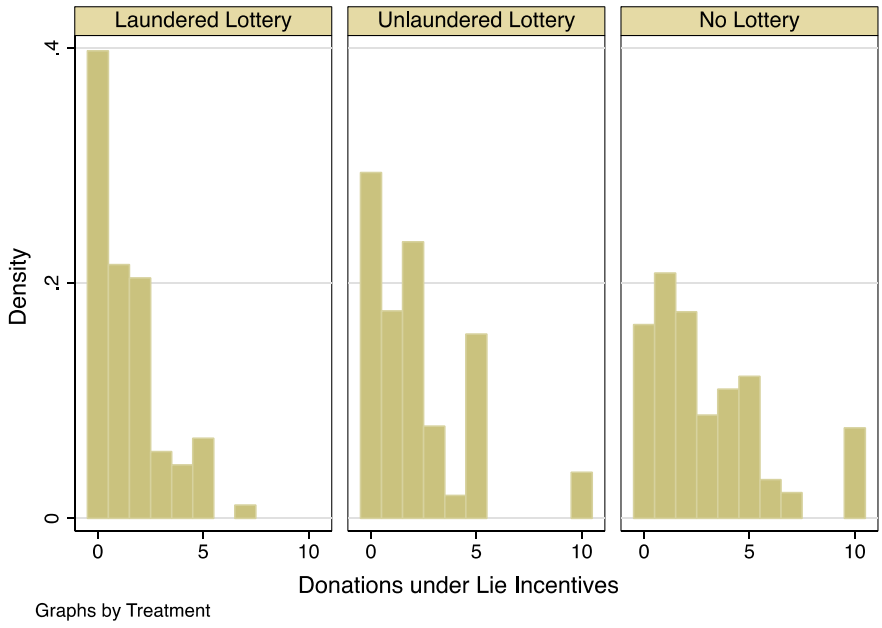

FIGURE 1. Donations under lie incentives by condition.

through the laundered lottery would be more likely to dissociate their earnings from the unethical account and would consequently donate less than lying senders in the no lottery and unlaundered lottery conditions.

Figure 1 shows the donation distributions by treatment for lie incentives. Table 2 presents ordinary least squares (OLS) regressions comparing the amount donated by condition, with dummy variables controlling for the wave of the study.[5] Results were consistent with our predictions: Whereas there is little difference in donation behavior between those in the unlaundered lottery and no lottery conditions ($0.70, 3.5% effect relative to endowment; column 1), participants in the laundered lottery condition donated significantly less than did participants in the no lottery condition ($1.56, 7.8% effect relative to endowment; column 1). Furthermore, participants in the laundered lottery condition donated less than did those in the more procedurally similar unlaundered lottery condition ($0.73, 3.7% effect relative to endowment; column 2). Dissociating dirty money from its source, rather than not exposing it to risk, appears to have driven the observed reduction in prosocial spending. This provides direct support for our Prediction 2. Table A2 in the Online Appendix, which presents summary statistics for each of the three waves separately, shows that the key differences between conditions are evident in all three waves.

5. Here we present results using OLS. Given the truncation of donations at zero, we also analyzed the data using Tobit models. The results, reported in Table A1 in the Online Appendix, are qualitatively unchanged.

TABLE 2. Donations by condition and wave: OLS regressions.

	Dependent variable:				
	Donation amount				
	Lie incentives		*Truth incentives*		*All conditions*
	(1)	(2)	(3)	(4)	(5)
Laundered lottery	−1.56***	−0.73*	0.62	0.39	0.48
	(0.34)	(0.35)	(0.63)	(0.60)	(0.42)
Unlaundered lottery	−0.70		0.23		
	(0.42)		(0.83)		
Lie incentives					1.74***
					(0.37)
Laundered lottery*					−1.79**
Lie incentives					(0.52)
Wave 2	−0.62	−0.09			−0.59
	(0.39)	(0.48)			(0.34)
Wave 3	0.41	0.75	1.03	1.03	0.68*
	(0.42)	(0.51)	(0.61)	(0.67)	(0.29)
Constant	3.11***	1.95***	0.68	0.91	1.02**
	(0.33)	(0.53)	(0.44)	(0.78)	(0.34)
N	230	139	113	88	343
R^2	0.12	0.07	0.08	.03	0.11

Notes: Dependent variable for all models is donation amount. Columns (1) and (2) focus on lie incentives. Column (1) compares unlaundered lottery and laundered lottery conditions to no lottery condition (baseline). Column (2) compares laundered lottery condition to unlaundered lottery condition (baseline); data from the no lottery condition are not included in this regression. Columns (3) and (4) are the same as (1) and (2) for truth incentives. Column (5) uses all six conditions. Standard errors in parentheses.

$^{*}p < 0.05$, $^{**}p < 0.01$, $^{***}p < 0.001$.

In contrast, performing the same analysis on donations made by participants in the Truth Incentive conditions, whose earnings were gained without deception, revealed no significant differences between the No Lottery, Unlaundered Lottery, and Laundered Lottery conditions (Columns 3 and 4). Table A3 in the Online Appendix, which presents summary statistics, shows that these results are consistent across the three waves.

Moreover, column 5 directly compares the relative effects of mental money laundering in the lie and truth incentive conditions in an OLS regression of the amount donated to charity on dummy variables corresponding to whether earnings were laundered or not (1 if laundered lottery, 0 otherwise), the initial earnings method (1 if lie incentives, 0 otherwise), and their interaction to capture the difference in difference. Dummy variables for the study waves were also included. As predicted, a significant interaction implies that laundering only affected the amount of money donated if it was unethically earned.

Since the populations of participants who chose to earn $20 differed between the two incentive schemes—the vast majority told the truth under truth Incentives while only a subset chose to lie under lie incentives—selection issues preclude direct

comparisons of donation *levels* between the two sets of conditions. Acknowledging this limitation, we find that for participants whose money was not laundered, those who lied donated significantly more than did those who told the truth ($\beta = 1.85$; $p < 0.001$). By contrast, for participants whose money was laundered, those who lied did not donate significantly more than participants who told the truth ($\beta = -0.20$; $p = 0.57$). Together, these comparisons between lie and truth incentive conditions provide support for Prediction 1.

3.2. Study 2: Mental Money Laundering with Multiple Sources

Our second experiment sought to demonstrate the generalizability of the effects documented in the first study, showing that people appear to engage in mental money laundering in contexts common to many compensation schemes—where earnings are obtained through multiple sources. The study also sought to test the prediction of motivated mental accounting: Given several possibilities for associating wealth with a particular source, an individual will choose the one that maximizes her utility. Specifically, we tested whether pooling money obtained through an unethical source together with money earned ethically would affect spending decisions relative to the case in which earnings from the two sources were not pooled.

Pooling the money could work in two ways. Due to the ambiguity of the whether any particular bill in pooled money was earned through clean or dirty sources, the ethically earned money could "clean" the unethical gains, or the unethical gains could "taint" the ethical gains. Motivated mental accounting predicts that the source ambiguity resulting from pooling would allow people greater flexibility in performing mental accounting operations, associating the entire sum to the account that enabled them to keep more money for themselves. By contrast, negativity bias has been documented in a wide variety of domains by psychologists (Baumeister et al. 2001; Rozin and Royzman 2001), and negative stimuli often induce a contagion effect whereby they change the valence of stimuli with which they become associated. As Rozin and Fallon observe in the original paper (Rozin and Fallon 1987), "a teaspoon of sewage will spoil a barrel of wine, but a teaspoon of wine will do nothing for a barrel of sewage". Negativity bias and contagion effects would predict that mixing clean money with even a small amount of tainted money should lead people to treat the entire sum as if it was tainted.

In Study 2, participants earned money from two tasks. In three of the four conditions, the first involved exerting real effort. The second involved a choice of whether to lie to another participant. In the no pool condition, participants received their earnings separately for each task, so that gains were segregated, and the source of dirty and clean money was easily identified. In the pool condition, money from both sources was *pooled* together; earnings from lying and exerting effort were mixed and given to the participant in the same envelope to obfuscate the source of the money. After receiving their earnings, participants chose how much to spend on a charitable donation.

In the pool lottery condition, we tested whether an additional exchange of the pooled money for clean money through a laundered lottery would lead to even lower charity donations than pooling alone. Money from the two sources was first routed

through a laundered lottery in which it was exchanged for physically different money and then pooled together into a single envelope.

In addition to the three core experimental conditions, we ran a fourth only effort condition in which participants earned clean money from two effort tasks. Their earnings were segregated, as in the no pool condition. Since all earnings in the only effort condition were obtained ethically, we predicted that donations will be lower than in the no pool condition—closer to those in the pool condition—despite being segregated by source. This allowed us to rule out the possibility that the increase in charitable spending in the No pool condition compared to the pool condition was due to the segregation of money rather than the mixing of earnings from different sources.

Method. Participants ($N = 360$) were recruited in even-numbered groups (minimum of 4, maximum of 8) from a student participant pool. In three of the four conditions, each completed two tasks for payment. The first was a real-effort task that asked participants to squeeze a hand dynamometer that recorded their force output in Newtons for 60 seconds. This task has been shown to capture changes in effort in response to differences in incentives (Imas 2014) and non-pecuniary treatments such as emotional manipulations (Gneezy and Imas 2014). Each participant was told that they would earn \$5 if their total force output exceeded 25,000 Newtons. This threshold was purposefully set so low that every participant met or exceeded it.

The second task involved either playing a deception game or squeezing the handgrip again. Participants assigned to the deception game were anonymously paired and then assigned to either the role of sender or receiver. This task was labeled as the Communication Game in the instructions. All senders played the game in the lie incentives condition, which was the same as in the first experiment, except that option Y paid the sender \$5 and the receiver \$10, while option X paid the sender \$10 and the receiver \$5. After making their decision to lie or tell the truth, participants were paid their earnings.

There were four conditions in the experiment. In the no pool condition, earnings from the effort task were placed in one envelope labeled "Handgrip Payment" and earnings from the deception game were placed in a separate envelope and labeled "Communication Payment".[6] In the pool condition, earnings from the first and second tasks were mixed together and given to the participants in the *same* envelope labeled "Total Payment". In the pool lottery condition, participants first had their earnings from both tasks exchanged through the laundered lottery from the previous study. The earnings were then given to participants in one envelope as in the pool condition. A comparison between the pool and pool lottery conditions allowed us to examine the extent to which flexibility in mental accounting could be augmented further by adding a second stage of laundering.

For their second task, participants in the only effort condition squeezed the hand dynamometer again. The payment was raised to \$10 and the threshold was raised slightly to 30,000 Newtons, still low enough to ensure that all participants either met

6. See Online Appendix for pictures of envelopes used.

TABLE 3. Experimental design by treatments (in rows).

Stage 1	Stage 2	Earnings	Stage 3
Effort task	Deception game	Not pooled	Charitable donation
Effort task	Deception game	Pooled	Charitable donation
Effort task	Deception game	Pooled and laundered lottery	Charitable donation
Effort task	Effort task	Not pooled	Charitable donation

or exceeded it. Participants who lied in the deception game thus stood to earn the same amount as did participants in the only effort condition. There was a 5-minute break between the two tasks during which participants could solve anagrams in order to equate time across all four conditions and allow participants to recover from the first effort task.

As in the first experiment, in the third phase, all participants privately read a passage asking how much of their payment they would like to donate to a charity (Save the Children). Participants then placed the amount they would like to donate into a separate envelope and sealed it. Table 3 depicts all conditions and stages in the experiment.

Predictions. In the case of pooling earnings, the DM's wealth W_O^2 is a mix of earnings W_{UE}^1 and W_E^1 that would otherwise be linked to unethical and ethical accounts, respectively, where $W_{UE}^1 \gg W_E^1$. There are three possibilities for the guilt associated with this wealth. The most objective would simply correspond to the fraction of earnings associated with the unethical source, such that $g^3 = W_{UE}^1 /(W_{UE}^1 + W_E^1)$. The second possibility follows the negativity bias effect, such that the negative association with unethical earnings overwhelms feeling about the ethical earnings, and $g^3 > W_{UE}^1/(W_{UE}^1 + W_E^1)$. Specifically, source ambiguity prevents the individual from identifying which money is linked to the ethical versus unethical account, which leads her to treat the entire sum as if it were "dirty" money.

The third possibility of motivated mental accounting, by contrast, predicts that by creating ambiguity in its source, pooling the money should enable people to mentally dissociate the dirty money from its unethical source. Pooling generates ambiguity regarding the source of any particular amount of wealth, which leaves open several possibilities of linking wealth to particular accounts. Motivated mental accounting predicts that the DM will choose the option which maximizes utility, such that $g^3 < W_{UE}^1/(W_{UE}^1 + W_E^1)$ and is much closer, if not equal to, the case in which the entire sum is associated with the ethical account. Relative to the no pool condition, which segregated earnings by source, motivated mental accounting predicts that people in the pool condition will spend the sum as if it was all ethically earned, which will reduce charitable spending. If pooling is sufficient for dissociation, then the inclusion of the lottery should not further reduce donations: donations in the pool and pool lottery condition should be similar. Finally, because the only effort condition only involved ethically earned cash, donation rates in that condition are predicted to be lower than in the no pool condition and similar to those in the pool condition.

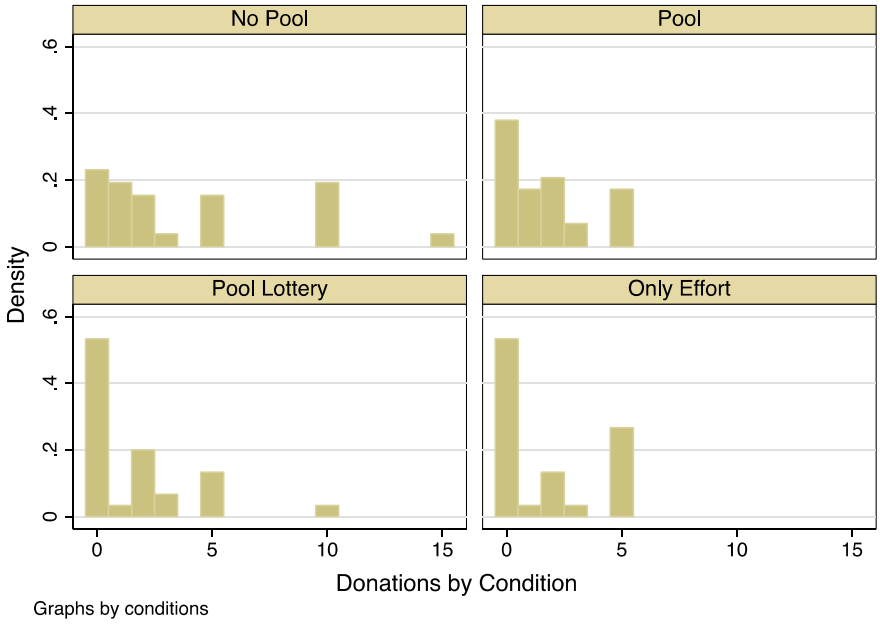

FIGURE 2. Donations by condition.

Results. Of the participants who participated in the deception game, 79% of receivers followed the sender's messages. This resulted in 260 participants for whom the receiver followed the sender's message. Of those senders, 65% sent a false message. As before, we examine the behavior of participants who earned the same amount, in this case \$15, across the four conditions ($N = 115$).

Figure 2 presents the distributions of charitable donations across the four conditions. Comparing donation levels across conditions, pooling appeared to effectively dissociate unethically earned money from its negative source. We regressed the amount donated on dummy variables corresponding to the pool and pool lottery conditions (Table 4).[7] Among participants whose earnings included unethically obtained cash, those in the pool and the pool lottery conditions donated significantly less than did those in the no pool condition (\$2.23, 14.9% effect relative to endowment; \$2.53, 15% effect relative to endowment; column 1), providing support for Prediction 2. Furthermore, the absence of a significant difference between participants in the pool and pool lottery condition ($p = 0.96$) suggests that the physical exchange of money that had already been pooled had no additional effect on donations, beyond pooling.

Finally, we compared pooling to the influence of simply segregating earnings by comparing the conditions in which participants played the deception game to the only

7. Results from Tobit regressions are presented in Table A4 of the Online Appendix; results are qualitatively unchanged compared to OLS. Summary statistics are presented in Table A5 in the Online Appendix.

TABLE 4. Multiple sources and donations: OLS regressions.

	Dependent variable:	
	Donation amount	
	(1)	(2)
Pool	−2.23**	−2.23**
	(0.80)	(0.75)
Pool lottery	−2.25**	−2.25**
	(0.79)	(0.74)
Only effort		−2.15**
		(0.74)
Constant	3.88***	3.89***
	(0.58)	(0.54)
N	85	115
R^2	0.11	0.10

Notes: Dependent variable for all models is donation amount. Column (1) compares pool and pool lottery conditions to the no pool condition (baseline). Column (2) compares the pool, pool lottery, and only effort conditions to the no pool condition (baseline). Standard errors in parentheses.
$^{*}p < 0.05$, $^{**}p < 0.01$, and $^{***}p < 0.001$.

effort condition. Note that comparisons to the only effort condition are subject to similar selection issues as in Study 1. Hence, comparisons of the only effort condition to the other three conditions should be taken as suggestive. As column 2 in Table 4 reveals, participants who lied in the no pool condition (whose earnings were also segregated in two envelopes) donated significantly more to charity than did participants in the only effort condition ($2.15, 14.3% effect relative to endowment). By contrast, participants who lied in the pool and pool lottery conditions donated as little to charity as did those in the only effort condition ($2.23, 14.9% effect relative to endowment; $2.25, 15% effect relative to endowment; $p > 0.85$ for all pairwise comparisons). Consistent with Prediction 1, the dissociation of dirty money from its source appears to underlie the differences in donations between the lie incentive conditions, not the mere segregation of earnings.

3.3. Sophistication

In our framework, associating money with an unethical source has a direct negative effect on utility. By de-coupling wealth from the unethical mental account, laundering should increase utility. The third and fourth experiments tested the degree to which people are sophisticated about the positive utility effects of laundering and proactively seek out opportunities to do it.

Predictions. The experiments that follow aim to test both parts of the following prediction:

PREDICTION 3. *The DM should strictly prefer laundering unethically earned funds. Moreover, she should be willing to pay some premium r to launder.*

Proof. Consider the case where $W_O^2 = W_O^1$, where W_O^2 corresponds to laundered wealth W_O^1 that is no longer associated with the unethical account. The first part of the prediction requires that the DM derives greater utility when more of her wealth is associated with the ethical account than the unethical one; specifically, that $u(c^*, d^*|W_O^2) > u(c^*, d^*|W_O^1)$. Guilt has a direct effect on utility, $u_g < 0$, and an indirect effect on utility as the DM reoptimizes her bundle. Define $\{c^*(g^x), d^*(g^x)\}$ as the optimal bundle when $g = g^x$. □

By the direct effect, we have

$$u(c^*(g^1), d^*(g^1), g^1) < u(c^*(g^1), d^*(g^1), g^2)$$

because $g^2 < g^1$. By the indirect effect we have

$$u(c^*(g^1), d^*(g^1), g^2) \leq u(c^*(g^2), d^*(g^2), g^2).$$

It follows that

$$u(c^*(g^2), d^*(g^2)|W_O^2) > u(c^*(g^1), d^*(g^1)|W_O^1).$$

Given the assumptions on u, the second part of Prediction 3 follows.

Study 3: Gambling to Launder. The third study had two stages. In the first stage, participants were incentivized to lie. In the second, they chose how much of their earnings to stake in a lottery that would, or would not, exchange their money for cash from a different source. If participants are sophisticated about the dissociative effects of laundering, then those who earned money by lying would be more likely to enter a lottery if it replaced unethical gains with cash from a different source (laundered lottery) than if it left them with the same cash as before (unlaundered lottery).

Method. Participants ($N = 220$) were recruited in even-numbered groups (minimum of 4, maximum of 8) through a student participant pool and assigned to either the role of sender or receiver in the deception game, which was conducted using paper and pencil instructions. All senders played the game with lie incentives, which were the same as in Study 1. After receiving their earnings, senders were randomly assigned to one of two lottery conditions: laundered lottery or unlaundered lottery. In this experiment, unlike in the first two studies, senders chose how much to stake in the lottery. In both conditions, each sender wrote down a success and a failure number between 1 and 12 for a lottery whose outcome would be determined by a digital die roll. Each sender then decided how much of her earnings from the first stage to stake in a lottery that paid double the amount staked if the outcome of the die roll matched her success number ($p = 0.083$), lost the amount staked if it matched her failure number ($p = 0.083$), and returned the same amount that was staked if the outcome was any other number ($p = 0.83$).

Prior to deciding how much of their earnings to stake, each participant read the complete instructions on how the lotteries would be carried out (i.e. as in the

TABLE 5. Experimental design by treatments (in rows).

Incentive	Stage 1	Stage 2
Lie	Sender–receiver game	Laundered lottery
Lie	Sender–receiver game	Unlaundered lottery

COLOUR ONLINE, B&W IN PRINT

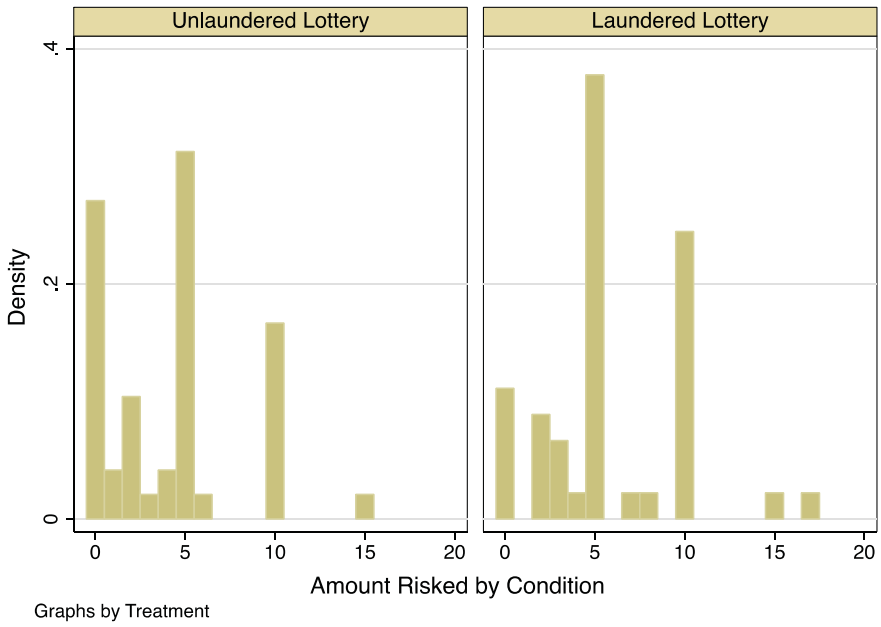

FIGURE 3. Amount risked by condition.

corresponding conditions in Study 1). These instructions made clear whether money entered in the lottery would be exchanged for "clean" bills from a different source (laundered lottery) or not (unlaundered lottery). Since any effect on giving behavior would have been confounded with individual differences based on the amounts participants were willing to stake in the lottery, we did not include a subsequent charitable donation stage. A depiction of the experimental design appears in Table 5.

In this study, the opportunity to launder required the participant to expose her earnings to greater risk. Assuming that participants are generally risk-averse, we interpret greater risk-taking in the laundered lottery condition as a willingness to pay premium to engage in mental money laundering (second part of Prediction 3).

Results. In total, 85% of receivers followed the sender's messages, leaving 186 participants. Of the 93 senders in these pairs, 66% sent a false message. As in the first experiment, our analysis focuses on senders who lied and earned the same amount ($20) in both lottery conditions ($N = 61$).

Figure 3 presents the distributions of the amount staked in the lotteries by condition. Consistent with our prediction, participants in the laundered lottery

TABLE 6. Experimental design by treatments (in rows).

Incentive	Stage 1	Stage 2
Lie	Sender–receiver game	Exchange of bills offered
Truth	Sender–receiver game	Exchange of bills offered

condition staked significantly more in the lottery ($M = \$6.17$; $SD = 4.15$) than did participants in the unlaundered lottery condition ($M = \$3.78$; $SD = 3.92$; $p = 0.024$). Here, participants valued the opportunity to launder money earned unethically, choosing to expose more of it to risk when the lottery dissociated their earnings from the unethical mental account.

Study 4: Requesting to Launder. The fourth study also had two stages. In the first stage, participants were incentivized to either lie or tell the truth. In the second, they could request that the experimenter exchange their earnings for money from a different source. Sophistication about the effects of mental money laundering in this setting would predict that those who lied would be more likely to exchange their earnings than participant who told the truth.

Method. Participants ($N = 186$) were recruited in even-numbered groups (minimum of 4, maximum of 8) through a student participant pool and assigned to either the role of advisor or chooser in the deception game, which was conducted using paper and pencil instructions. All advisors played the game with either lie or truth incentives, similar to Study 1, except for the following changes. In the lie incentives conditions, option Y paid the advisor \$1 and the chooser \$5; option X paid the advisor \$5 and the chooser \$1. The advisor earned more money and the chooser earned less money if the former lied and the latter followed the message. In the truth incentive conditions, option Y earned both players \$5 while option X earned both \$1. After the chooser made a decision, both players were paid their earnings in cash (in \$1 bills).

After being paid, all participants were called one by one to the front of the room to sign for their earnings before leaving. At this stage, a research assistant blind to the hypothesis told those who had earned \$5 that there was an opportunity to exchange the \$1 bills for a single \$5 bill from a visibly separate supply of funds. The research assistant then discretely recorded whether the participant chose to exchange her earnings or not.[8] Given evidence from the first study on how the exchange of money affects prosocial spending, we interpret the request to exchange earnings as demand for the opportunity to engage in mental money laundering (first part of Prediction 3). A depiction of the experimental design appears in Table 6.

8. This study was run based on a pilot experiment conducted using a student sample from a different East Coast university. There, 94 participants were recruited and assigned to either the lie incentives or truth incentives condition. After removing pairs where the chooser did not follow the advisor's message, we found that 38% of Advisors in the lie incentives condition chose to exchange their funds, compared to 0% of advisors in the truth incentives condition.

Results. In total, 89% of receivers followed the senders' messages, leaving 166 participants. Of the 83 advisors in this sample, 65% sent a false message in the lie incentives condition and 7% sent a false message in the truth incentives condition. Our analysis focuses on advisors who earned the same amount ($5) in both the lie and truth incentive conditions ($N = 65$).

As noted in the analyses of previous studies, comparisons between those who lied in the lie incentives condition and those who did not in the truth incentives condition should be taken as suggestive, given potential selection issues. That being said, 31% of advisors who lied for $5 asked to have their earnings laundered, whereas 18% of those who told the truth made the same request ($p = 0.24$). While the 13% effect is trending in the predicted direction, it is not significant. Note that in some ways our fourth experiment is a more direct test of the third prediction, in that the dependent variable is a binary choice of whether to launder rather than a decision of how much money to risk. We interpret the findings from the last two studies as providing supportive but mixed evidence for sophistication about the dissociative effects of laundering.

Discussion. The mixed support for sophistication can be interpreted in several ways. One possibility is that ex ante beliefs about laundering effects are situation specific. While the instructions for Study 3 clearly delineate that money entered in the laundered lottery will be exchanged for cash from a different source, some participants in Study 4 may not have noticed that exchanging money would lead to clean funds. There was no novel activity with which the clean money would be associated, so the "source" of the exchanged funds may still have been the most salient activity in the experiment (i.e. the sender–receiver game). This may explain that while 13% more people chose to exchange their money when they lied for it than when they told the truth, this effect was weaker than in Study 3. Another possibility is selection: as discussed earlier in the paper, the group of people who lied in the lie incentives condition was different than the group who told the truth in the truth incentives condition. If, for some reason, those in the latter condition had a greater baseline propensity to exchange funds, this could explain the more muted effect. Notably, selection was not an issue in Study 3. Finally, experimenter demand may have played a larger role in Study 4. Unlike the other studies in which decisions were made in private cubicles, the choice to exchange money was an explicit public request to the experimenter. If participants who lied did not want to signal their deception by asking for an exchange of funds, this would moderate their propensity to request for it. Despite these potential explanations for the diminished effect in Study 4, we cannot rule out that our results simply suggest limited sophisticated about the effects of mental money laundering.

4. Discussion and Conclusion

Our studies document that people exploit flexibility in mental accounting to psychologically dissociate earnings from undesirable sources and spend money in more selfish ways. In Study 1, participants who earned money by lying donated significantly less

of it to charity when their earnings were replaced with physically different bills, via a lottery, than when the lottery returned the original "dirty" money. Study 2 demonstrated the generalizability of mental money laundering to the more common situation in which people pool multiple sources of earnings. Finally, the results of Studies 3 and 4 provide mixed support for sophistication—that people are aware of opportunities to dissociate earnings from undesirable sources ex ante, and actively seek them out.

Our findings are the first to document the phenomenon of motivated mental accounting using incentive-compatible experiments. The results show that people exploit flexibility in mental accounting to do what they want to do rather than what they believe they should do. More generally, our findings contribute to a developing literature exploring how people rationalize their unethical behavior. People avoid obtaining information that might compel them, out of a sense of ethicality, to sacrifice self-interest (Dana, Weber and Kuang 2007). They hire others to take selfish actions on their behalf (Hamman, Loewenstein and Weber 2010) so that they feel less responsible for the resulting outcomes than they would if they had taken the actions themselves (Bartling and Fischbacher 2012). Moreover, third parties hold selfish agents less responsible if they employ intermediaries to act selfishly on agent's behalf as well (Coffman 2011; Fershtman and Gneezy 2001). In a similar vein, we find that people seek to dissociate the rewards of unethical past acts from their sources by exploiting flexibility in mental accounting operations.

Our evidence is confined to the laboratory, but Study 2 hints at the likely ubiquity of mental money laundering. Most people and organizations that are engaged in morally questionable activities are also engaged in legal, ethical ventures. Tax cheats with unreported income typically also earn legitimate income that they duly report. Firms that earn money through dubious means also earn money in legitimate ways. Professors who engage in questionable consulting activities, such as serving as an expert witness for corporations in lawsuits against the public, also earn salaries from their normal academic work. Prior research has shown that in most situations "bad is stronger than the good" (Baumeister et al. 2001). A small vial of fecal matter can ruin an entire vat of wine, but no amount of wine can make a vat of fecal matter appetizing (Rozin and Rozyman 2001). In contrast, Study 2 shows that a smaller amount of "clean" money (ethical earnings) purified a larger amount of "dirty" money (unethical earnings). That mental accounting seems to be an exception to the "bad is stronger than the good" rule is a testament to the scope of flexible accounting practices, and more generally to the power of human rationalization.

New technologies that blur the line between mental accounting and "real" accounting, such as payment rails and budgeting software (e.g. Venmo and Mint, respectively), offer novel opportunities for creatively pooling resources. Helping consumers create virtual accounts that mirror mental accounts can certainly have benefits, such as enhancing saving and making spending more efficient (e.g. Loewenstein et al. 2012). Our findings suggest that it may also provide a tool aiding mental money laundering and encouraging behaviors with social costs. Consumers may eliminate undesirable source effects simply by depositing unethical earnings into an account containing money that was ethically earned.

The malleability of mental accounting has significant implications for how empirical results from the literature should be interpreted and applied in practice. Our results suggest that the labeling of income, shown in a number of papers to affect spending patterns, will be less effective when the label encourages patterns of spending that conflict with the person's broader desires. When this is the case, for labeling to be maximally effective, policymakers should be careful to remove any source ambiguity that would allow individuals to exploit this malleability to license self-interested or short-sighted behavior. For example, the observed fungibility violations in Kooreman (2000) may have been due to the unambiguous labeling of funds as "child benefits". If the same funds were labeled as "family benefits", malleability in mental accounting may have led recipients to spend them in more self-interested ways that benefited the adults in the family (e.g. consumer electronics or vacations). Non-fungible payment mechanisms that prevent physical laundering and are associated with a specific expense category, like electronic benefit transfer (EBT) cards used to deliver Supplemental Nutrition Assistance Program benefits, may be particularly effective at inhibiting mentally money laundering of funds intended for family-oriented spending (Hastings and Shapiro 2018).

In situations in which labeling does align with a person's broader desires, our results suggest that sophisticated firms may use this to, for example, pay a smaller wage premium through the practice of "greenwashing". These implications connect to research in labor economics, which finds that workers demand a wage premium to work at firms where either the means of production or the good itself has negative externalities (Frank 1996), and are willing to take a pay cut when the virtuous signals provided by their work are presumed to supplement wages (e.g. "care work"; England, Budig, and Folbre 2002). Consumers similarly prefer to avoid goods with negative externalities, and firms respond to this behavior by lowering prices relative to goods without these externalities (Bartling, Weber, and Yao 2015). Given the ability of people to engage in mental money laundering, firms may respond by introducing elements of CSR into the means of production, or by coupling purchases with donations to charity. By pooling unethical elements with the (albeit smaller) ethical elements, such "greenwashing" may lead potential employees to be less averse to working for the firm and consumers to pay more for goods (Dubé, Luo, and Zhang 2017), enabling the company to reap greater rents.

More generally, our results inform the literature on how mental accounting affects financial decision making. Thaler and Shefrin (1981) and Shefrin and Thaler (1988) posited that mental accounting allows people to rationalize spending decisions and to exert more self-control, and Loewenstein et al. (2012) discuss how mental accounting can be used to mitigate consumers' pain of paying, enhancing pleasure from consumption. Our findings suggest that these disparate benefits can be drawn together to form a more unified perspective on mental accounting. Mental accounting is more than a side effect of the categorization of sources and pools of resources. It is a strategy people flexibly apply to facilitate a variety of saving and spending goals, whether their end is self-constraint, maximizing experiential utility, or reducing the disutility of engaging in selfish and unethical behavior.

References

Abeler, Johannes and Felix Marklein (2017). "Fungibility, Labels, and Consumption." *Journal of the European Economic Association*, 15, 99–127.

Andreoni, James, Justin M. Rao, and Hannah Trachtman (2017). "Avoiding the Ask: A Field Experiment on Altruism, Empathy, and Charitable Giving." *Journal of Political Economy*, 125, 625–653.

Bartling, Björn and Urs Fischbacher (2012). "Shifting the Blame: On Delegation and Responsibility." *The Review of Economic Studies*, 79, 67–87.

Bartling, Björn, Roberto A. Weber, and Lan Yao (2015). "Do Markets Erode Social Responsibility?" *The Quarterly Journal of Economics*, 130, 219–266.

Baucells, Manel and Woonam Hwang (2017). "A Model of Mental Accounting and Reference Price Adaptation." *Management Science*, 63, 4201–4218.

Baumeister, Roy F., Ellen Bratslavsky, Catrin Finkenauer, and Kathleen D. Vohs (2001). "Bad Is Stronger than Good." *Review of General Psychology*, 5, 323–370.

Bénabou, Roland, Armin Falk, and Jean Tirole. (2018). "*Narratives, Imperatives, and Moral Reasoning*." NBER Working Paper No. w24798.

Bénabou, Roland and Jean Tirole (2011). "Identity, Morals, and Taboos: Beliefs as Assets." *The Quarterly Journal of Economics*, 126, 805–855.

Cheema, Amar and Dilip Soman (2006). "Malleable Mental Accounting: The Effect of Flexibility on the Justification of Attractive Spending and Consumption Decisions." *Journal of Consumer Psychology*, 16, 33–44.

Coffman, Lucas C (2011). "Intermediation Reduces Punishment (and Reward)." *American Economic Journal: Microeconomics*, 3, 77–106.

Crawford, Vincent P. and Joel Sobel (1982). "Strategic Information Transmission." *Econometrica: Journal of the Econometric Society*, 50, 1431–1451.

Cryder, Cynthia E., Stephen Springer, and Carey K. Morewedge (2012). "Guilty Feelings, Targeted Actions." *Personality and Social Psychology Bulletin*, 38, 607–618.

Dana, Jason, Daylian M. Cain, and Robyn M. Dawes (2006). "What You Don't Know Won't Hurt Me: Costly (but Quiet) Exit in Dictator Games." *Organizational Behavior and Human Decision Processes*, 100, 193–201.

Dana, Jason, Roberto A. Weber, and Jason Xi Kuang (2007). "Exploiting Moral Wiggle Room: Experiments Demonstrating an Illusory Preference for Fairness." *Economic Theory*, 33, 67–80.

Dubé, Jean-Pierre, Xueming Luo, and Zheng Fang (2017). "Self-Signaling and Prosocial Behavior: A Cause Marketing Experiment." *Marketing Science*, 36, 161–186.

England, Paula, Michelle Budig, and Nancy Folbre (2002). "Wages of Virtue: The Relative Pay of Care Work." *Social Problems*, 49, 455–473.

Epley, Nicholas, Dennis Mak, and Lorraine Chen Idson (2006). "Bonus of Rebate?: The Impact of Income Framing on Spending and Saving." *Journal of Behavioral Decision Making*, 19, 213–227.

Erat, Sanjiv and Uri Gneezy (2012). "White Lies." *Management Science*, 58, 723–733.

Exley, Christine L (2016). "Excusing Selfishness in Charitable Giving: The Role of Risk." *The Review of Economic Studies*, 83, 587–628.

Fershtman, Chaim and Uri Gneezy (2001). "Strategic Delegation: An Experiment." *RAND Journal of Economics*, 32, 352–368.

Frank, Robert H (1996). "What Price the Moral High Ground?" *Southern Economic Journal*, 63, 1–17.

Gneezy, Uri (2005). "Deception: The Role of Consequences." *American Economic Review*, 95(1), 384–394.

Gneezy, Uri and Alex Imas (2014). "Materazzi Effect and the Strategic Use of Anger in Competitive Interactions." *Proceedings of the National Academy of Sciences of the USA*, 111, 1334–1337.

Gneezy, Uri, Alex Imas, and Kristóf Madarász (2014). "Conscience Accounting: Emotion Dynamics and Social Behavior." *Management Science*, 60, 2645–2658.

Gneezy, Uri, Silvia Saccardo, Marta Serra-Garcia, and Roel van Veldhuizen (2020). "Bribing the Self." *Games and Economic Behavior*, 120, 311–324.

Hamman, John R., George Loewenstein, and Roberto A. Weber (2010). "Self-Interest through Delegation: An Additional Rationale for the Principal–Agent Relationship." *American Economic Review*, 100(4), 1826–46.
Hastings, Justine and Jesse M. Shapiro (2018). "How Are SNAP Benefits Spent? Evidence from a Retail Panel." *American Economic Review*, 108(12), 3493–3540.
Heath, Chip and Jack B. Soll (1996). "Mental Budgeting and Consumer Decisions." *Journal of Consumer Research*, 23, 40–52.
Henderson, Pamela W. and Robert A. Peterson (1992). "Mental Accounting and Categorization." *Organizational Behavior and Human Decision Processes*, 51, 92–117.
Hines, James R. and Richard H. Thaler (1995). "The Flypaper Effect." *Journal of Economic Perspectives*, 9(4), 217–226.
Imas, Alex (2014). "Working for the "Warm Glow": On the Benefits and Limits of Prosocial Incentives." *Journal of Public Economics*, 114, 14–18.
Kooreman, Peter (2000). "The Labeling Effect of a Child Benefit System." *American Economic Review*, 90(3), 571–583.
Lazear, Edward P., Ulrike Malmendier, and Roberto A. Weber (2012). "Sorting in Experiments with Application to Social Preferences." *American Economic Journal: Applied Economics*, 4, 136–163.
Levav, Jonathan and A. Peter McGraw (2009). "Emotional Accounting: How Feelings about Money Influence Consumer Choice." *Journal of Marketing Research*, 46, 66–80.
Loewenstein, George, Cynthia E. Cryder, Shlomo Benartzi, and Alessandro Previtero (2012). "Addition by Division: Partitioning Real Accounts for Financial Well-Being." In *Transformative Consumer Research for Personal and Collective Well-Being*. Routledge, pp. 441–450.
Morewedge, Carey K., Leif Holtzman, and Nicholas Epley (2007). "Unfixed Resources: Perceived Costs, Consumption, and the Accessible Account Effect." *Journal of Consumer Research*, 34, 459–467.
Nyilasy, Gergely, Harsha Gangadharbatla, and Angela Paladino (2014). "Perceived Greenwashing: The Interactive Effects of Green Advertising and Corporate Environmental Performance on Consumer Reactions." *Journal of Business Ethics*, 125, 693–707.
O'Curry, S (1997). "Income Source Effects." Working paper, Department of Marketing, DePaul University Chicago.
Park, Hyun Young and Tom Meyvis (2011). "Feeling Moral about Money: How Moral Emotions Influence Consumer Spending Decisions." *ACR North American Advances*, 39, 233–234.
Prelec, Drazen and George Loewenstein (1998). "The Red and the Black: Mental Accounting of Savings and Debt." *Marketing Science*, 17, 4–28.
Rozin, Paul and April E. Fallon (1987). "A Perspective on Disgust." *Psychological Review*, 94, 23.
Rozin, Paul and Edward B. Royzman (2001). "Negativity Bias, Negativity Dominance, and Contagion." *Personality and Social Psychology Review*, 5, 296–320.
Shefrin, Hersh M. and Richard H. Thaler (1988). "The Behavioral Life-Cycle Hypothesis." *Economic Inquiry*, 26, 609–643.
Sussman, Abigail B. and Adam L. Alter (2012). "The Exception Is the Rule: Underestimating and Overspending on Exceptional Expenses." *Journal of Consumer Research*, 39, 800–814.
Tasimi, Arber and Susan A. Gelman (2017). "Dirty Money: The Role of Moral History in Economic Judgments." *Cognitive Science*, 41, 523–544.
Thaler, Richard (1985). "Mental Accounting and Consumer Choice." *Marketing Science*, 4, 199–214.
Thaler, Richard H (1999). "Mental Accounting Matters." *Journal of Behavioral Decision Making*, 12, 183–206.
Thaler, Richard H. and Hersh M. Shefrin (1981). "An Economic Theory of Self-Control." *Journal of Political Economy*, 89, 392–406.

Supplementary Data

Supplementary data are available at *JEEA* online.