Working Paper Series 777
(ISSN 2788-0443)

Why do some nudges work and others not?

Matej Lorko
Tomáš Miklánek
Maroš Servátka

CERGE-EI
Prague, March 2024

ISBN 978-80-7343-584-4 (Univerzita Karlova, Centrum pro ekonomický výzkum a doktorské studium)
ISBN 978-80-7344-712-0 (Národohospodářský ústav AV ČR, v. v. i.)

Why do some nudges work and others not?

Matej Lorko[a], Tomáš Miklánek[b], Maroš Servátka[c]

Abstract: While nudges have recently gained popularity, many nudging interventions fail, and the effects of successful ones are often short-lived. We conjecture that the success of a nudge depends on how it interacts with the underlying economic incentives that determine the payoff-maximizing behavior of the decision-maker. For example, in the domain of tax compliance, a nudge is likely to be effective only if it is financially optimal for the taxpayer to pay the tax. To test our conjecture, we run a multi-period experiment in which we manipulate tax audit probability, and nudge participants to report their income. In addition, we vary how often the nudge appears, to test whether more frequent nudging increases long-run compliance. We observe that the first application of a nudge has a positive immediate effect on income reporting irrespective of whether it is optimal to comply or not. However, subsequent nudges increase income reporting only if the nudge is aligned with the taxpayer's incentives. More frequent nudging in the direction opposite to incentives yields no effects on long-run compliance. Policy implications are discussed.

Keywords: nudge, incentives, tax compliance, experiment

This project was funded by the Czech Science Foundation (GAČR, project no. GA23-07136S) and by the Scientific Grant Agency of the Ministry of Education of Slovak Republic and the Slovak Academy of Sciences (VEGA Project No. 1/0167/21). The project has been reviewed and approved by the Ethical Committee of the University of Economics in Bratislava (no. EKEUBA-VEGA1/0167/21-2/2022)

[a] Corresponding author, Matej Lorko, University of Economics in Bratislava, Slovakia, and Faculty of Business Administration, Prague University of Economics and Business, Czech Republic. matej.lorko@gmail.com

[b] Tomáš Miklánek, Faculty of Business Administration, Prague University of Economics and Business, Czech Republic. tomas.miklanek@vse.cz

[c] Maroš Servátka, MQBS Experimental Economics Laboratory, Macquarie Business School, Sydney, Australia and the University of Economics in Bratislava, Slovakia. maros.servatka@mgsm.edu.au

1. Introduction

Thaler & Sunstein, (2008, p.6) define nudges as "aspects of the choices architecture that alter people's behavior in a predictable way without forbidding any options or significantly changing their economic incentives". These inconspicuous and cost-efficient policy interventions have recently gained prominence, with governments around the world establishing over 200 behavioral insights teams (also known as "nudge units") to suggest, test and apply behavioral nudges (DellaVigna & Linos, 2022; Hubble & Varazzani, 2023).

Behavioral research has shown that for nudges to work they must be easy, attractive, social, and timely (BIT Handbook, 2014). Yet, many nudges are found to be ineffective despite satisfying these criteria. The question of why some nudges work, and others do not, remains therefore unanswered, at least not in a satisfactory way. In this study we propose a conjecture that a nudge will work only if it is aligned with the underlying economic incentives. This simple yet fundamental idea seems to have been ignored by policy-makers, and to the best of our knowledge has not yet received attention in empirical research.[1]

We conjecture that if a nudge goes against economic incentives, it might temporarily perturb behavior but it will be ineffective if repeated. Its initial effect will dwindle over time as the economic forces correct behavior that is not in the best interest of the decision-maker. On the other hand, the effect of a nudge aligned with economic incentives will persist if it is reapplied, as it reinforces the optimal behavior. A subsequent question then emerges: Does more frequent nudging foster the desirable behavior?

We explore our general research question about the interaction of nudges and economic incentives in the context of tax compliance, where nudges have been extensively applied and studied, and where we can easily manipulate whether it is in the best (expected) monetary interest for a decision-maker to report income and pay tax owed or not.

[1] One can argue that careful wording of Thaler & Sunstein (2008) that a nudge does not change economic incentives implicitly encompasses our conjecture. Sunstein (2017) posits that the presence of strong opposing preferences is one potential reason for the inefficiency of nudges. However, this aspect has not been investigated causally.

2. Nudging and tax compliance

Tax compliance nudges are eminently attractive to governments, complementing systematic policy instruments such as enforcement (Slemrod, 2019; Slemrod & Yitzhaki, 2002). Given the dramatic magnitudes of tax evasion (see e.g., Internal Revenue Service, 2022; Kukk et al., 2020), even if the nudge increases the compliance rate only slightly, it can add immense revenues to government budgets. A rich variety of nudges have been proposed and applied, ranging from moral appeals, appeals to social norms or peer examples (Hallsworth et al., 2017), highlighting public goods provision (Bott et al., 2020), to correcting procrastination and limited attention (Hernandez et al., 2017; Loewenstein & Wojtowicz, 2023; Mascagni & Nell, 2022).

However, the general literature reviewing nudging studies (Benartzi et al., 2017; DellaVigna & Linos, 2022; Szaszi et al., 2018) reports mixed results regarding nudge efficacy, with some nudges having no effects (e.g., Gravert & Collentine, 2021) and some even backfiring (e.g., Holzmeister et al., 2022; Liu et al., 2016). Mixed effects of nudges are also reported in the domain of income reporting and taxation. While in some cases tax nudges increase compliance, both statistically and economically (e.g., Hallsworth et al., 2017; Holz et al., 2020; Vainre et al., 2020), they can backfire as well (e.g., De Neve et al., 2021). A recent meta-analysis of 45 field experiments (Antinyan & Asatryan, 2019) concludes that while deterrence-style nudges can positively affect tax compliance, non-deterrence nudges are usually less effective. However, deterrence-style nudges are not a panacea as their effect decreases over a longer horizon.

In this paper, we explore a from-an-economist-point-of-view fundamental reason why nudges yield mixed results by examining the effects of (mis)alignment between the underlying economic incentives that determine optimal behavior and the implemented nudge. Applying this general idea to the tax compliance context, we conjecture that even though nudges, by definition, bypass economic incentives, the interaction between the nudge and taxpayer incentives to comply (or not) determines whether the nudge will succeed or fail (see also Alm, 2019). If the nudge is misaligned with the incentives and pushes the decision-maker in the opposite direction relative to what is optimal, it is deemed to fail. For example, if the controlling

mechanisms and enforcing institutions are sufficiently weak, so there is a low probability of tax audit and/or low penalties for non-compliance, the decision to pay the tax is suboptimal for a risk-neutral (and in our setup, even for a realistically risk-averse) taxpayer. While a nudge may work the first time it is applied, due to the element of surprise and saliency, nudging is unlikely to increase tax compliance in the long run as both surprise and saliency dissipate over time while the economic incentives persist.[2]

Even though from a policy-making perspective and for welfare evaluation it is crucial to consider the effects of nudging over a longer horizon, a review by Szaszi et al. (2018) concludes that over 90 percent of nudging interventions implement only a one-off nudge, and the related studies usually evaluate only immediate effects. Notable exceptions include Manoli & Turner (2014), who provide evidence that the effect of a one-time nudge sharply decreases over time and disappears after three years, which suggests that nudging may need to be applied repeatedly. However, Antinyan et al. (2021) claim that after crossing some optimal threshold, more frequent nudging does not lead to further significant gains and decreases the overall efficiency, because of its repetitive costs. Even more strikingly, Dunning et al. (2017) show that a one-time nudging intervention may have unintended adverse effects in the long run by disrupting existing habits of desired behavior.

Our study contributes to the understanding of the effects of nudges on tax compliance in two ways. First, we make a distinction between optimal and non-optimal compliance. While the existing tax compliance literature varies whether it is optimal to comply by changing crucial parameters (e.g., Alm & Malézieux, 2021; Malézieux, 2018), our study conceptually relates the underlying incentives to nudge effectiveness. Our second contribution is particularly important for policy implications. Rather than focusing on the immediate effect of a one-off nudge, lessons from which may be limited for environments featuring a periodically repeated activity like reporting one's income and paying taxes, we explore whether the conjectured effectiveness of nudges aligned with economic incentives persists every time the nudge is applied and whether it increases with a higher frequency of nudging.

[2] Relatedly, we speculate that the main reason why deterrence nudges are more successful than non-deterrence nudges is the possibility that deterrence nudges are more likely to change the "perceived" incentives. For example, if nudge significantly increases taxpayer's subjective probability of being audited (in the spirit of Slemrod et al., 2001), the decision to pay taxes becomes optimal from the perspective of her perceived economic incentives. We note that our paper is not intended to offer insights into changes in the perceived audit rate since in our experiment the likelihood of audit is always stated explicitly.

We conduct a controlled multi-period laboratory experiment in which subjects execute an individual real-effort task to earn a taxable income.[3] In each period, a subject can report the earned income or not. By varying the audit probability we manipulate whether it is optimal to report income and thus pay tax (i.e., the audit rate is high enough that a risk-neutral taxpayer maximizes his expected payoff by complying) or not (i.e., the audit rate is low enough that unless the taxpayer is extremely risk averse, she should not report income). Depending on the experimental condition, subjects may be exposed to a nudge in the form of a reminder to report their income in some periods. A crucial feature of our experimental environment is that income reporting decisions are repeated. We investigate the interaction of audit rate (determining the economic incentives) and the presence and frequency of a nudge, leading to the following hypotheses:

- ***Hypothesis 1:*** *The first nudge increases immediate tax compliance regardless of whether it is optimal to pay tax or not.*
- ***Hypothesis 2:*** *Repeated nudges increase immediate tax compliance only if it is optimal to pay tax.*
- ***Hypothesis 3:*** *If it is optimal to pay tax, more frequent nudges increase long-run tax compliance. If it is not optimal to pay tax, more frequent nudges do not affect long-run compliance.*

In line with our hypotheses, we find that when an individual is nudged for the first time, she is significantly more likely to report her income, independently of whether the audit probability is low or high. However, repeated nudges increase immediate tax compliance (in the period when the nudge is applied) only under a high probability of audit, i.e., if the nudge is aligned with the underlying economic incentives of a taxpayer. Finally, we find that more

[3] While there is a growing number of field studies on tax nudges we conduct our experiment in a laboratory environment because it allows us to (1) control underlying economic incentives, (2) reduce the complexity of the decision-making environment, (3) control who is nudged and how often, (4) ensure that the nudge is acknowledged, (5) eliminate potentially confounding influences such as spillover effects from different nudges and/or nudges to which an individual was exposed indirectly, e.g., via communication with a different individual who was originally targeted (Mascagni, 2018), and (6) conduct a clean test of the long-term effect of repeated nudges without attrition of subjects. Many of these important necessary considerations for answering our research question are unattainable in the field. For a discussion about the external validity of laboratory tax experiments, see Alm et al. (2015).

frequent nudges have a positive effect on long-run tax compliance under the high probability of audit, while more frequent nudges under low probability yield no effect.

3. Experimental design

Our experiment employs the canonical tax game introduced by Allingham & Sandmo (1972), and consists of 16 periods in which subjects execute an individual real-effort task originally used in Lorko et al. (2019). In this task, the computer screen displays an inequality between a pair of two-digit numbers and the subject indicates whether the inequality is true or false. Immediately after the answer is submitted, a new, randomly generated inequality appears. Each period lasts 120 seconds.

To successfully complete the task and earn income for the period, a subject must provide 50 correct answers while providing fewer than 10 incorrect answers. The task is easy to understand and requires no prior training. However, it is also cognitively demanding and captivating, which opens space for inattention towards income reporting. We set the threshold for correct answers to 50 so that most of the subjects would take between 60 and 90 seconds to finish the task in each period. We implement the limit of 10 incorrect answers to discourage subjects from random clicking.

Income reporting and the reminder nudge

In each period the screen shows information about the current number of correct and incorrect answers, a counter displaying elapsed time, and a "Report your income" button (see a screenshot in the Appendix). Clicking the button reports the income for the current period, while not clicking the button results in the income being unreported. We chose to report the current period income instead of past period income because reporting the current period income makes all periods identical, meaning that there is no first period without reporting and no last period without the effort task. We reduce income reporting into a binary decision (report all income or none) to make the incentive structure easy to understand and to limit additional cognitive load (e.g., Deck et al., 2021) from thinking about the optimal level of underreporting.[4]

[4] A meta-study by Alm & Malézieux (2021) shows that even when subjects have the option to report any share of their income, approximately 64% of them report either full or zero income. The third most frequently used share - half of the income - occurs in fewer than 5% of the cases.

Importantly, while the button is displayed on the screen for the entire duration of each period, it is only active between the 30th and 60th seconds from the start of the period (see Table 1 for a period overview). We limit the time when income reporting is possible to parallel the outside-of-the-lab environment where the tax return window usually lasts for approximately one-third to one-quarter of a year, depending on the country and the tax system. The setup also allows for not reporting the income due to being inattentive while completing the task. In such a scenario, the nudge can serve as a helpful reminder.

The implemented nudge takes the form of a pop-up window that covers the entire task at the moment when the income reporting button becomes active. The message informs the subject that she can report her income in the next 30 seconds. To acknowledge the message and remove it from the screen, the subject must click the OK button. A screenshot of the nudge is provided in the Appendix.

Table 1: Overview of a period with a nudge reminder

Time from the start of the task	Event
Second 1	Task starts
Second 30	A nudge message appears
Second 30	The income-reporting button becomes active
Second 60	The income-reporting button becomes inactive
Second 120	Task ends

Earnings

In each period, a subject can earn 100 tokens if she provides 50 correct answers and fewer than 10 incorrect answers.[5] If the subject does not complete the task successfully, she earns zero income for the period. Reporting income leads to a fixed tax of 30 tokens being automatically subtracted from the 100 tokens earned in the given period.

[5] In the experiment, one experimental token was worth one Eurocent. The exchange rate was announced in the subject instructions.

In each period a subject can be audited with the audit probability of either 10% or 60%, depending on the experimental condition. The audit probability is kept constant for each subject across all periods. Failure to report income if audited results in having to pay the tax owed of 30 tokens and a fine of additional 30 tokens. No tax or fine is imposed if income for the period is zero, regardless of whether it was reported or not. The tax and fine remain constant for all subjects across all periods.[6] We postpone feedback about whether the audit occurred until all 16 periods have been completed to eliminate the potential confounding effects of being audited on the ensuing reporting decisions (e.g., the "bomb-crater" effect observed by Mittone et al., 2017). Table 2 shows per-period earnings for all possible combinations of income reporting and audit, provided that the task is completed successfully.

Table 2: Period earnings by income reporting and audit

	Audit	No audit
Income reported	Income – Tax = 70 tokens	Income – Tax = 70 tokens
Income not reported	Income – Tax – Fine = 40 tokens	Income = 100 tokens

Each subject is paid the cumulative sum of his/her payoffs from all 16 periods. We chose to pay for all periods instead of paying for one or more randomly chosen periods to establish a dominant strategy for almost all empirically relevant risk attitudes (Dave et al., 2010; Eckel & Grossman, 2008; Holt & Laury, 2002), in both the low-audit-probability condition and the high-audit-probability condition. Assuming Constant Relative Risk Aversion (CRRA), indifference between reporting and not reporting under 10% and 60% probability of audit occurs at the r coefficient values of 4.6 and -0.93, respectively.[7] In the risk attitude calibration by Holt & Laury (2002), 97% of subjects fall between these two values. Our payoff protocol, in which earnings accumulate across all 16 periods, accentuates the differences across audit probabilities even more. In case of the 10% audit probability, never reporting the income (across all 16 periods) results in higher earnings compared to always reporting 99.99% of the time. Thus, a payoff-maximizing agent with any risk aversion from a typically observed range will never report her income. If the probability of audit is 60%, never reporting leads to higher earnings compared to always reporting only 14.2% of the time and thus, the audit rate is high enough that it is favorable for all risk-averse, all risk-neutral and most risk-loving agents to

[6] For an overview of the effects of different parameter variations on tax compliance, see Blackwell et al. (2007), Malezieux (2018), or Alm & Malézieux (2021).

[7] The CRRA utility function is defined as $U(x) = \frac{x^{1-r}}{1-r}$ where x denotes a monetary payoff.

always report. In addition, the implemented protocol of paying for all periods mimics the actual tax-paying practice.

Experimental conditions

We employ a 2 x 4 factorial across-subject design (see Table 3 for an overview), in which we vary (1) the audit probability (10% or 60%) to identify how a nudge interacts with the underlying economic incentives and (2) how often a nudge appears on the screen to identify the effect of nudging frequency on reporting behavior. In the Baseline condition the nudge is never displayed. In the other three conditions a nudge appears for the first time in Period 5, as we allow subjects to become familiar with the real-effort task and income reporting in the first four periods. In the Single condition the nudge is never displayed again after Period 5. In the Periodic condition the nudge is displayed again in Periods 9 and 13, i.e., four periods apart. In the Regular condition the nudge is displayed in every following period after Period 5 until and including Period 13. The last three periods (i.e., Periods 14 to 16) in the Regular condition do not feature the nudge so that we can investigate whether the nudge effect persists after the intervention is removed.

Table 3: Experimental conditions

Nudge frequency	Low audit probability (10%)	High audit probability (60%)
Baseline (never)	Baseline-Low	Baseline-High
Single (Period 5)	Single-Low	Single-High
Periodic (Periods 5, 9, 13)	Periodic-Low	Periodic-High
Regular (Periods 5-13)	Regular-Low	Regular-High

Procedures

In each experimental session the same experimenter handed out printed instructions (all experiment-related materials are included in the Appendix) to subjects and read them aloud. Any subject questions were answered in private. After reading the instructions each subject was required to complete a set of 10 non-incentivized comprehension questions. Subjects were allowed to use the printed instructions and ask for help from the experimenter. If a comprehension question was answered incorrectly, a message window identifying the question with an incorrect answer appeared. The first period began only after each subject in the session correctly answered all 10 questions.

After completing all 16 periods subjects participated in an incentivized risk-elicitation task where they had to choose between 70 tokens with certainty and a lottery which, to a certain extent, resembled the payoffs and probabilities used in the income-reporting task. Subjects

were informed that one of their choices would be randomly chosen and the resulting payoff would be added to their experimental earnings. The risk-elicitation task included two salience manipulations of contrast and prominence, to generate a proxy for being susceptible to nudging (Bordalo et al., 2022, see the Appendix for details).

Subjects then completed a three-item modified Cognitive Reflection Test (CRT, Primi et al., 2016) in which they could earn 50 tokens for each correctly answered question. The CRT was followed by unincentivized elicitation of dishonesty attitudes based on the World Values Survey (Inglehart et al., 2014), a set of basic demographics, and two open-ended questions. The first question asked subjects about how they made their reporting decisions and the second allowed them to leave any other comments about the experiment.

The final screen displayed per period information about the subject's performance, her reporting decision, whether an audit occurred, and what her resulting task earnings were. The screen also showed earnings from the randomly selected risk elicitation choice and the CRT. At the end of the experiment the experimental earnings were converted from tokens into cash and paid privately to the subjects.

Subjects

A total of 393 subjects, all students at the University of Economics in Bratislava, participated in the experiment. The number of subjects in each condition ranged from 45 to 53 due to variance in show-up rates. One of the subjects left the experiment after the third period and is therefore excluded from the analysis. The sample analyzed consists of 392 subjects (210 females) with a mean age of 21.5 (SD 2.0).

All experimental sessions were conducted in the Bratislava Behavioral and Experimental Economics Research Lab at the University of Economics in Bratislava between November 2022 and February 2023. The experiment used a computerized interface programmed in zTree (Fischbacher, 2007). Subjects were recruited using the Online Recruitment System for Economic Experiments (ORSEE; Greiner, 2015) from a subject database maintained by the lab. On average, an experimental session lasted approximately 90 minutes including the initial instructional period and payment of subjects. The subjects earned EUR 13 on average.[8]

[8] For comparison, the hourly wage that students could earn at the time of the experiment as a research assistant or in manual jobs typically ranged from EUR 5 to EUR 7.

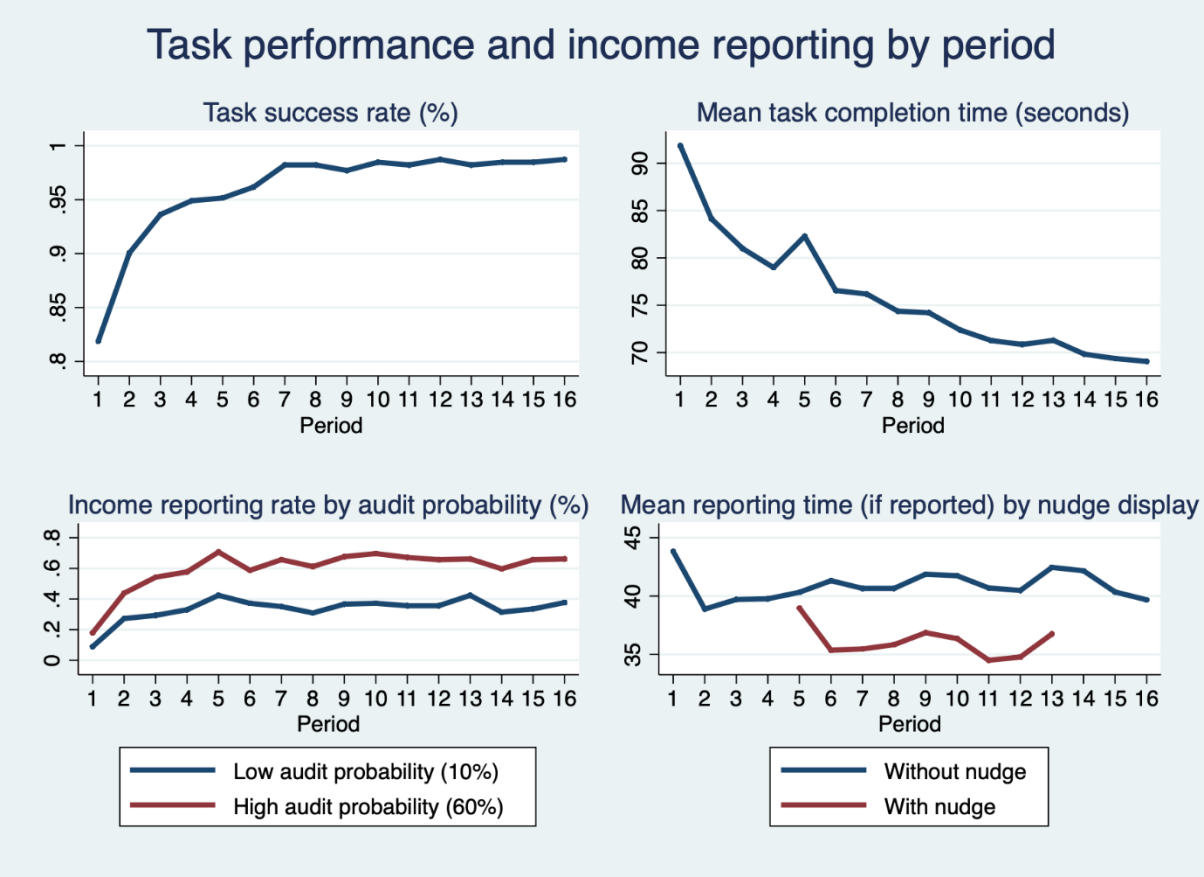
4. Results

Figure 1 displays summary statistics of task performance and income reporting. The upper-left panel shows that the task was completed successfully 96% of the time. The success rate increased to 98% after the initial four periods. The upper-right panel displays the average duration of task completion, which was 92 seconds in the first period and then gradually declined to 69 seconds in the last period. The sudden increase in task duration in Period 5 was caused by the nudge, which appeared in that period for the first time in all conditions except for the two Baseline conditions that did not feature a nudge in any period. Similar but less pronounced increases are detectable in Periods 9 and 13, in which the nudge reappeared in the two Periodic conditions.

The lower-left panel shows that the income reporting rate (which is our measure of subject tax compliance as reporting one's income automatically results in paying tax owed) was consistently higher under the high audit probability (59.9% across all 16 periods and all conditions) than under the low audit probability (33.3%). The income reporting rates increased and became relatively steady after the initial four periods (65.3% and 36.3%, for the high and low audit probability, respectively). Finally, the lower-right panel displays the average time, counted from the start of the period, at which income reporting took place and shows that whenever a nudge appeared, subjects who decided to report their income did so on average four seconds sooner compared to periods without the nudge (36.7 and 40.7 seconds, respectively).

Due to apparent learning effects occurring in the first four periods when subjects were familiarizing themselves with the experimental environment and reporting decisions, in what follows, we only analyze data from the fifth period onwards. Recall that this is when nudging started. We note that there were no statistical differences in income reporting rates across conditions given the audit probability in the initial four periods (Pearson's chi-square test yields p-values of 0.14 and 0.79 for the 10% and 60% audit probability, respectively).

Figure 1: Task performance and income reporting by period

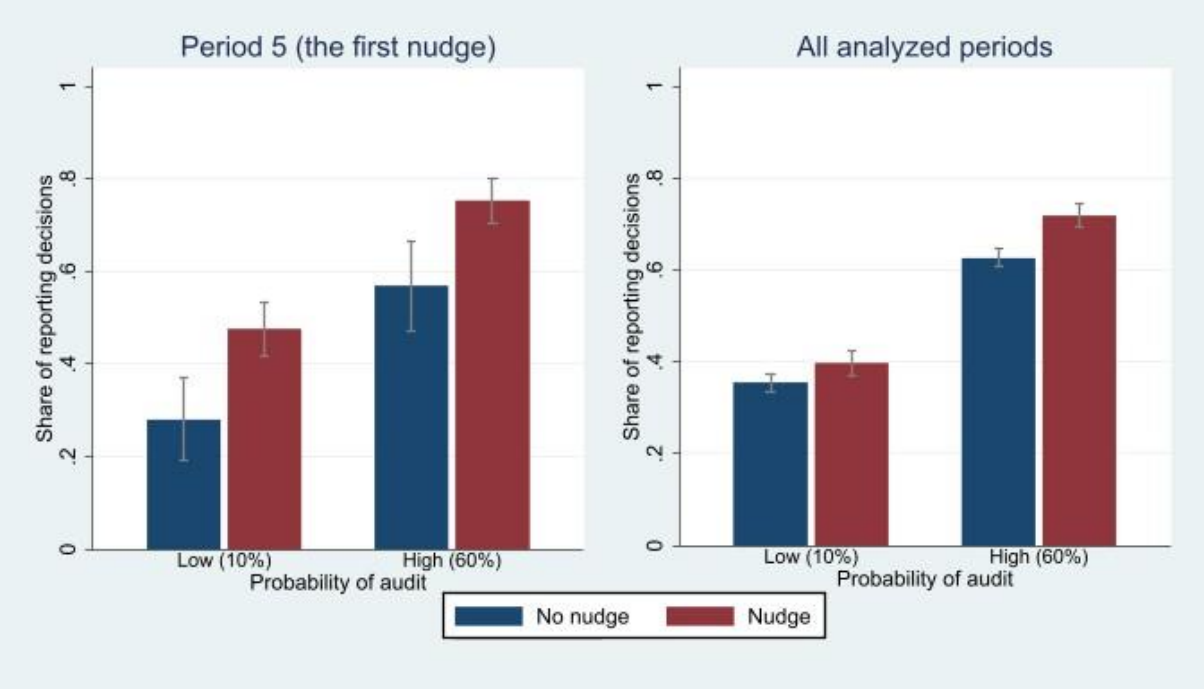


The first nudge

We begin by testing Hypothesis 1 that the first exposure to a nudge increases immediate tax compliance regardless of the audit probability. We find that in Period 5, subjects report their income significantly more often in the three nudge conditions (Single, Periodic, Regular) compared to the Baseline condition under both low and high audit probability. The differences (see the left panel of Figure 2) are 20 percentage points under the low audit probability and 18 percentage points under the high audit probability, with Pearson's chi-square tests yielding p-values of 0.02 and 0.01, respectively, providing support for Hypothesis 1.

Result 1: *The first nudge increases immediate tax compliance regardless of whether it is optimal to pay tax or not.*

Figure 2: Average income reporting rates with and without a nudge



Note: The figure displays 83.4% confidence intervals to illustrate two-sample means test results with α=0.05.

Repeated nudges

Next, we test Hypothesis 2 that repeated nudges increase immediate tax compliance only if the nudge is aligned with the taxpayer's economic incentives, i.e., only if it is optimal to pay tax. Table 4 presents OLS regressions with clustered standard errors at the subject level. For clarity of interpretation we run separate models for each probability of audit. In Models 1 and 2, we include Periods 5-16 and find that nudges indeed significantly increase the immediate income reporting rate only when the audit probability is high (see also the right panel of Figure 2). A higher cognitive reflection score is associated with a significantly higher probability of making optimal decisions of not reporting the income if the audit probability is low and reporting if the audit probability is high. Higher risk aversion is associated with a higher income-reporting rate under both audit probabilities while a higher dishonesty score is associated with a lower income-reporting rate.[9] We also find that females are less responsive to the audit probability than males. Inconsistency in the saliency manipulation choices is not significantly correlated with reporting decisions in any specification and its exclusion from the set of control variables does not qualitatively change any results. For brevity we do not include

[9] We calculate the dishonesty score from the set of five questions eliciting acceptance of different forms of cheating using the principal component analysis (Jackson, 2005).

it in the subsequent results. As a robustness check, in Models 3 and 4 we exclude Period 5 (i.e., the first nudge) and find qualitatively similar results.

Result 2: *Repeated nudges increase immediate tax compliance only if the nudge is aligned with incentives, i.e., only if it is optimal to pay tax.*

Table 4: Effect of nudge on income reporting

	Periods 5-16 (all analyzed periods)		Periods 6-16 (without the first nudge)	
	(1)	(2)	(3)	(4)
	Report rate	Report rate	Report rate	Report rate
	Low audit P	High audit P	Low audit P	High audit P
Nudged	0.05	0.10***	0.02	0.08**
	(0.04)	(0.03)	(0.04)	(0.04)
Cognitive ab.	-0.09***	0.05**	-0.09***	0.05**
	(0.02)	(0.02)	(0.02)	(0.02)
Female	0.17***	-0.09**	0.18***	-0.09**
	(0.04)	(0.04)	(0.04)	(0.04)
Risk aversion	0.07***	0.05***	0.07***	0.05***
	(0.01)	(0.01)	(0.01)	(0.01)
Dishonest	-0.02*	-0.03*	-0.02*	-0.03*
	(0.01)	(0.01)	(0.01)	(0.01)
Period	0.00	0.00	0.00	0.00
	(0.00)	(0.00)	(0.00)	(0.00)
Constant	0.12*	0.35***	0.10	0.33***
	(0.07)	(0.08)	(0.07)	(0.08)
N	2292	2412	2101	2211
R^2	0.16	0.07	0.16	0.07

Notes: Standard errors are reported in parentheses. All regressions include clustering of standard errors at the subject level. *, **, and *** indicate significance at the 10%, 5%, and 1%-level, respectively. Nudged is a dummy variable that equals 1 if the subject was nudged in the relevant period and 0 otherwise.

Nudging frequency

Our Hypothesis 3 states that more frequent nudges increase long-run tax compliance, but only if it is optimal to pay tax. In the OLS regressions presented in Table 5, we therefore estimate the effect of nudging once (Single condition), three times (Periodic condition), and nine times (Regular condition). We use the income reporting rate of individual subjects as the unit of observation and estimate the effects of nudging frequency across Periods 5-13 (i.e., periods in which the nudge appeared in at least one condition, Models 1 and 2), and across Periods 5-16 (Models 3 and 4). In line with Hypothesis 3, we find that under the high audit probability, the Periodic nudge condition and the Regular nudge condition weakly significantly increase the likelihood that a subject will report her income compared to the Baseline condition.

On the other hand, nudging more frequently under the low audit probability yields no significant effects on the income reporting rate.

Result 3: *If it is optimal to pay tax, nudging more frequently leads to (weakly significantly) higher long-run tax compliance. If it is not optimal to pay tax, more frequent nudging does not affect long-run compliance.*

Table 5: Effect of nudging frequency on long-run tax compliance

	Periods 5-13		Periods 5-16	
	(1)	(2)	(3)	(4)
	Report rate	Report rate	Report rate	Report rate
	Low audit P	High audit P	Low audit P	High audit P
Single nudge condition	-0.02	0.07	-0.02	0.08
	(0.06)	(0.06)	(0.06)	(0.06)
Periodic nudge condition	0.01	0.09	0.01	0.10*
	(0.06)	(0.06)	(0.06)	(0.06)
Regular nudge condition	-0.00	0.12**	-0.03	0.10*
	(0.06)	(0.06)	(0.06)	(0.06)
Control variables	YES	YES	YES	YES
Constant	0.15**	0.36***	0.15**	0.34***
	(0.07)	(0.09)	(0.07)	(0.09)
N	191	201	191	201
R^2	0.31	0.16	0.32	0.17

Notes: Standard errors are reported in parentheses. *, **, and *** indicate significance at the 10%, 5%, and 1%-level, respectively. Control variables include cognitive reflection, gender, risk aversion, and dishonesty score.

Nudge removal

Our design also allows us to investigate what happens after the nudge is removed. To that effect we analyze the likelihood that a subject reports income in three periods after the nudge appears for the first time (Periods 6-8) and three periods after the Regular nudge is removed (Periods 14-16). Table 6, Models 1 and 2 show that under high audit probability, subjects who were nudged in Period 5 continued to report income insignificantly more frequently in Periods 6-8 without a nudge. On the other hand, under low audit probability, we observe that nudge removal results in an insignificant decrease in the likelihood of reporting income in Periods 6-8. We find similar directional effects in Models 3 and 4, comparing Periods 14-16 in the Regular condition where the nudge was always displayed in the nine preceding periods to the Baseline condition. In summary, we find directional but statistically insignificant support for the intuition that once a nudge aligned with economic incentives is removed, its positive effects on compliance may persist, while the removal of a misaligned nudge may backfire into lower compliance.

Table 6: Effects of nudge removal

	Periods 6-8, all conditions		Periods 14-16, Baseline and Regular condition	
	(1)	(2)	(3)	(4)
	Report rate	Report rate	Report rate	Report rate
	Low audit P	High audit P	Low audit P	High audit P
Nudge removed	-0.03	0.03	-0.11	0.04
	(0.06)	(0.06)	(0.07)	(0.07)
Nudge displayed	0.03	0.13**		
	(0.06)	(0.07)		
Control variables	YES	YES	YES	YES
Constant	0.07	0.30***	0.08	0.20*
	(0.08)	(0.10)	(0.12)	(0.12)
N	191	201	99	100
R^2	0.27	0.16	0.29	0.17

Notes: Standard errors are reported in parentheses. *, **, and *** indicate significance at the 10%, 5%, and 1%-level, respectively. Nudge removed is a dummy variable that equals 1 if the subject was nudged in the previous period(s) and 0 otherwise. Nudge displayed is a dummy variable that equals 1 if the subject continued to be nudged in the relevant periods and 0 otherwise. Control variables include cognitive reflection, gender, risk aversion, and dishonesty score.

Auxiliary analyses

Finally, Table 7 presents our auxiliary regression analyses of behavior that our hypotheses are not explicit about. We find that when the nudge appears, income is reported more promptly and there are fewer late report attempts (i.e., trying to report the income after the reporting window is closed). Although the nudge also prolongs the task duration (because subjects must acknowledge it once it appears on the screen), it does not significantly affect whether the task is completed successfully, and thus does not reduce the period income, nor the period cash earnings after tax and audit.

While we find that the high audit probability (unsurprisingly) leads to lower period earnings, it also increases late reporting rates. A higher cognitive reflection is associated with more frequent successful task completion as well as with higher earnings due to a higher frequency of optimal reporting choices. Finally, we find that females take significantly longer to complete the task than males but report their income sooner.

Table 7: Auxiliary analysis

	(1) Late report	(2) Reporting time	(3) Task duration	(4) Period income	(5) Period earnings
Nudged	-0.03***	-4.34***	2.11**	-2.21	-0.02
	(0.01)	(0.66)	(1.03)	(1.39)	(0.01)
Period	-0.01***	-0.10*	-0.90***	0.12	0.00
	(0.00)	(0.06)	(0.07)	(0.08)	(0.00)
High audit prob.	0.03**	-0.47	1.70	-1.92	-0.18***
	(0.01)	(0.92)	(1.16)	(1.50)	(0.01)
Cognitive refl.	-0.03***	-0.76*	-0.49	1.71**	0.03***
	(0.01)	(0.46)	(0.67)	(0.73)	(0.01)
Female	0.02*	-1.66**	6.67***	-0.48	-0.03**
	(0.01)	(0.81)	(1.21)	(1.50)	(0.01)
Risk averse	0.00	-0.04	0.05	1.26*	-0.00
	(0.00)	(0.28)	(0.39)	(0.72)	(0.01)
Dishonest	0.00	0.20	0.30	-0.60	-0.00
	(0.00)	(0.25)	(0.35)	(0.44)	(0.00)
Constant	0.13***	44.63***	78.10***	90.27***	0.80***
	(0.03)	(1.62)	(2.23)	(3.45)	(0.03)
N	4704	2408	4605	4704	4704

Notes: Standard errors are reported in parentheses. *, **, and *** indicate significance at the 10%, 5%, and 1%-level, respectively.

5. Discussion

Due to their simplicity and inexpensive implementation nudging interventions have been studied extensively in many domains of economic behavior, including tax reporting, on-time utility payments, healthy lifestyle, retirement saving, pro-environmental behavior, and others. Governments find nudges attractive because of their capability to shift behavior in the desired direction without a need for mandates, choice restrictions, and changes in economic incentives. However, at least in the area of tax reporting, a large fraction of nudging experiments find no or very weak effects. Based on our experimental findings, we speculate that a potential explanation behind many failed tax nudges is the possibility that the underlying economic incentives were ignored. While nudges are capable of changing behavior, they are unable to overturn economic incentives.

Our findings resonate with recent research investigating factors that may interact with nudges. Löfgren & Nordblom (2020) conclude that the effect of nudging is stronger for relatively less important decisions. This claim is supported by de Ridder et al. (2022) who add that agents cannot be nudged into decisions that are not aligned with their initial preferences. On the other hand, Farhi & Gabaix (2020) allow for this possibility in their theoretical model of optimal taxation by assuming that susceptibility to nudges is an exogenous (and

heterogenous) parameter and that nudges can directly affect utility. Similarly, Bordalo et al. (2022) argue that while FAST agents (Forgetful and Salient Thinkers) could be nudged into decisions that are against their initial preferences, nudges can influence FBOR agents (Forgetful But Otherwise Rational) only if the nudging direction is aligned with their initial preferences. Finally, Goldin (2015) and Johnson & Goldstein (2003) claim that nudging could be particularly effective in an environment in which preferences are not clear or agents are inconsistent in their choices.

We also explore the effects of repeated nudges. Interestingly, extant research differs in theoretical predictions of behavior when nudging is repeated over several periods. On the one hand, Taubinsky (2013) argues that nudges in the form of reminders lead to habit formation and therefore repeated nudges increase the probability that an individual will act again in the future, even without a nudge. The same reasoning is suggested by Mullainathan (2002) and empirically observed in the works of Tobias (2009) and Henderson et al. (2011). On the other hand, Sellen et al. (1997) claim that the increased probability of acting on a reminder might not be permanent, and Ericson (2017) argues that anticipated and unanticipated reminders yield different theoretical predictions. While unanticipated reminders increase the probability of task completion, anticipated reminders can both increase and decrease this probability.

We find that if it is optimal for a taxpayer to report income, more frequent nudging leads to higher long-run tax compliance. While our analysis shows that the effect of more frequent nudging is relatively weak, we note that nudges generally do not overturn preferences but instead make it easier for an individual to act in accordance with her underlying preferences. From this perspective, a weakly significant increase in long-run compliance can be seen as the upper limit of what nudges can achieve. In line with the earlier findings that nudges can sometimes backfire we find that when it is not optimal for the taxpayer to comply, the nudge removal yields an insignificantly negative effect on compliance. In practice, this could mean that if the taxpayer perceives the nudge as just an instant call to action, she may not pay tax in the future unless she is called to action by a subsequent nudge again.

All in all our study shows that economic incentives interact with nudging. While nudges can be effective if they are aligned with underlying economic incentives, policy-makers should not expect nudges to reliably increase tax compliance if the tax audit rates (or at least subjectively perceived audit probability) and/or penalties are too low, parallel to the observed null effects of honesty nudges in an environment with zero risk of penalty for not reporting

truthfully (Dimant et al., 2020). This is particularly important if the cost of the nudging intervention is non-negligible. It is therefore crucial from the policy perspective to understand the perceived economic incentives of taxpayers.

Laboratory experiments offer many advantages over field experiments, especially when it comes to repetition and control. The fundamental value of laboratory experiments is in testing theoretical ideas in the most rudimentary conditions. If a laboratory test delivers proof of concept, it is not prudent to generalize laboratory results to the outside-the-lab world. However, the underlying theoretical idea that receives empirical support can be applied to the field setting and tested further for robustness. A laboratory test is thus a part of the process that in the end expands our understanding of everyday life phenomena. Having said that, it is crucial to consider the external validity of laboratory experiments and conduct trials testing the robustness of laboratory findings in a naturally occurring field setting. In the case of the current study, future research should attempt to replicate our findings, ideally over multiple reporting periods and controlling for the subjective probability of being tax audited.

References

Allingham, M. G., & Sandmo, A. (1972). Income tax evasion: A theoretical analysis. *Journal of Public Economics*, *1*(3–4), 323–338.

Alm, J. (2019). WHAT MOTIVATES TAX COMPLIANCE? *Journal of Economic Surveys*, *33*(2), 353–388. https://doi.org/10.1111/joes.12272

Alm, J., Bloomquist, K. M., & McKee, M. (2015). On the external validity of laboratory tax compliance experiments. *Economic Inquiry*, *53*(2), 1170–1186.

Alm, J., & Malézieux, A. (2021). 40 years of tax evasion games: a meta-analysis. *Experimental Economics*, *24*(3), 699–750.

Antinyan, A., & Asatryan, Z. (2019). Nudging for tax compliance: A meta-analysis. *ZEW-Centre for European Economic Research Discussion Paper*, *19–055*.

Antinyan, A., Asatryan, Z., Dai, Z., & Wang, K. (2021). Does the frequency of reminders matter for their effectiveness? A randomized controlled trial. *Journal of Economic Behavior & Organization*, *191*, 752–764.

Benartzi, S., Beshears, J., Milkman, K. L., Sunstein, C. R., Thaler, R. H., Shankar, M., Tucker-Ray, W., Congdon, W. J., & Galing, S. (2017). Should governments invest more in nudging? *Psychological Science*, *28*(8), 1041–1055.

Blackwell, C., & others. (2007). A meta-analysis of tax compliance experiments. *International Studies Program Working Paper*, *7*, 24.

Bordalo, P., Gennaioli, N., & Shleifer, A. (2022). Salience. *Annual Review of Economics*, *14*, 521–544.

Bott, K. M., Cappelen, A. W., Sørensen, E. Ø., & Tungodden, B. (2020). You've got mail: A randomized field experiment on tax evasion. *Management Science*, *66*(7), 2801–2819.

Dave, C., Eckel, C. C., Johnson, C. A., & Rojas, C. (2010). Eliciting risk preferences: When is simple better? *Journal of Risk and Uncertainty*, *41*, 219–243.

De Neve, J.-E., Imbert, C., Spinnewijn, J., Tsankova, T., & Luts, M. (2021). How to improve tax compliance? Evidence from population-wide experiments in Belgium. *Journal of Political Economy*, *129*(5), 1425–1463.

de Ridder, D., Kroese, F., & van Gestel, L. (2022). Nudgeability: Mapping Conditions of Susceptibility to Nudge Influence. *Perspectives on Psychological Science*, *17*(2), 346–359. https://doi.org/10.1177/1745691621995183

Deck, C., Jahedi, S., & Sheremeta, R. (2021). On the consistency of cognitive load. *European Economic Review*, *134*, 103695.

DellaVigna, S., & Linos, E. (2022). RCTs to scale: Comprehensive evidence from two nudge units. *Econometrica*, *90*(1), 81–116.

Dimant, E., van Kleef, G. A., & Shalvi, S. (2020). Requiem for a Nudge: Framing effects in nudging honesty. *Journal of Economic Behavior & Organization*, *172*, 247–266. https://doi.org/10.1016/j.jebo.2020.02.015

Dunning, T., Monestier, F., Pineiro, R., Rosenblatt, F., & Tuñón, G. (2017). Is paying taxes habit forming? theory and evidence from Uruguay. *University of California-Berkeley Working Paper*.

Eckel, C. C., & Grossman, P. J. (2008). Forecasting risk attitudes: An experimental study using actual and forecast gamble choices. *Journal of Economic Behavior & Organization*, *68*(1), 1–17.

Ericson, K. M. (2017). On the interaction of memory and procrastination: Implications for reminders, deadlines, and empirical estimation. *Journal of the European Economic Association*, *15*(3), 692–719.

Farhi, E., & Gabaix, X. (2020). Optimal taxation with behavioral agents. *American Economic Review*, *110*(1), 298–336.

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

Goldin, J. (2015). Which way to nudge: Uncovering preferences in the behavioral age. *Yale LJ*, *125*, 226.

Gravert, C., & Collentine, L. O. (2021). When nudges aren't enough: Norms, incentives and habit formation in public transport usage. *Journal of Economic Behavior & Organization*, *190*, 1–14.

Greiner, B. (2015). Subject pool recruitment procedures: organizing experiments with ORSEE. *Journal of the Economic Science Association*, *1*(1), 114–125.

Hallsworth, M., List, J. A., Metcalfe, R. D., & Vlaev, I. (2017). The behavioralist as tax collector: Using natural field experiments to enhance tax compliance. *Journal of Public Economics*, *148*, 14–31.

Henderson, C. M., Beck, J. T., & Palmatier, R. W. (2011). Review of the theoretical underpinnings of loyalty programs. *Journal of Consumer Psychology*, *21*(3), 256–276. https://doi.org/10.1016/j.jcps.2011.02.007

Hernandez, M., Jamison, J., Korczyc, E., Mazar, N., & Sormani, R. (2017). *Applying Behavioral Insights to Improve Tax Collection*.

Holt, C. A., & Laury, S. K. (2002). Risk aversion and incentive effects. *American Economic Review*, *92*(5), 1644–1655.

Holz, J. E., List, J. A., Zentner, A., Cardoza, M., & Zentner, J. (2020). *The 100 million nudge: Increasing tax compliance of businesses and the self-employed using a natural field experiment*.

Holzmeister, F., Huber, J., Kirchler, M., & Schwaiger, R. (2022). Nudging debtors to pay their debt: Two randomized controlled trials. *Journal of Economic Behavior & Organization*, *198*, 535–551.

Hubble, C., & Varazzani, C. (2023). *Mapping the global behavioural insights community - Observatory of Public Sector Innovation*. https://oecd-opsi.org/blog/mapping-behavioural-insights/

Inglehart, R., Haerpfer, C., Moreno, A., Welzel, C., Kizilova, K., Diez-Medrano, J., Lagos, M., Norris, P., Ponarin, E., Puranen, B., & others. (2014). World values survey: Round six-country-pooled datafile version. *Madrid: JD Systems Institute*, *12*.

Internal Revenue Service. (2022). *Federal Tax Compliance Research: Tax Gap Estimates for Tax Years 2014-2016*. www.irs.gov

Jackson, J. E. (2005). *A user's guide to principal components*. John Wiley & Sons.

Johnson, E. J., & Goldstein, D. (2003). Do defaults save lives? In *Science* (Vol. 302, Issue 5649, pp. 1338–1339). American Association for the Advancement of Science.

Liu, C.-W., Agarwal, R., & Gao, G. (2016). *The dark side of positive social influence*.

Loewenstein, G., & Wojtowicz, Z. (2023). The Economics of Attention. *Available at SSRN 4368304*.

Löfgren, Å., & Nordblom, K. (2020). A theoretical framework of decision making explaining the mechanisms of nudging. *Journal of Economic Behavior & Organization*, *174*, 1–12. https://doi.org/10.1016/j.jebo.2020.03.021

Lorko, M., Servátka, M., & Zhang, L. (2019). Anchoring in project duration estimation. *Journal of Economic Behavior & Organization*, *162*, 49–65.

Malezieux, A. E. (2018). *A practical guide to setting up your tax evasion game*.

Manoli, D. S., & Turner, N. (2014). *Nudges and learning: evidence from informational interventions for low-income taxpayers*.

Mascagni, G. (2018). From the lab to the field: A review of tax experiments. *Journal of Economic Surveys*, *32*(2), 273–301.

Mascagni, G., & Nell, C. (2022). Tax compliance in Rwanda: Evidence from a message field experiment. *Economic Development and Cultural Change*, *70*(2), 587–623.

Mittone, L., Panebianco, F., & Santoro, A. (2017). The bomb-crater effect of tax audits: Beyond the misperception of chance. *Journal of Economic Psychology*, *61*, 225–243.

Mullainathan, S. (2002). A memory-based model of bounded rationality. *The Quarterly Journal of Economics*, *117*(3), 735–774.

Primi, C., Morsanyi, K., Chiesi, F., Donati, M. A., & Hamilton, J. (2016). The development and testing of a new version of the cognitive reflection test applying item response theory (IRT). *Journal of Behavioral Decision Making*, *29*(5), 453–469.

Sellen, A. J., Louie, G., Harris, J. E., & Wilkins, A. J. (1997). What Brings Intentions to Mind? An In Situ Study of Prospective Memory. *Memory*, *5*(4), 483–507. https://doi.org/10.1080/741941433

Slemrod, J. (2019). Tax compliance and enforcement. *Journal of Economic Literature*, *57*(4), 904–954.

Slemrod, J., Blumenthal, M., & Christian, C. (2001). Taxpayer response to an increased probability of audit: evidence from a controlled experiment in Minnesota. *Journal of Public Economics*, *79*(3), 455–483.

Slemrod, J., & Yitzhaki, S. (2002). Tax avoidance, evasion, and administration. In *Handbook of public economics* (Vol. 3, pp. 1423–1470). Elsevier.

Sunstein, C. R. (2017). Nudges that fail. *Behavioural Public Policy* , *1*(1), 4–25. https://doi.org/10.1017/BPP.2016.3

Szaszi, B., Palinkas, A., Palfi, B., Szollosi, A., & Aczel, B. (2018). A systematic scoping review of the choice architecture movement: Toward understanding when and why nudges work. *Journal of Behavioral Decision Making*, *31*(3), 355–366.

Taubinsky, D. (2013). From intentions to actions: A model and experimental evidence of inattentive choice. *Work Pap*.

Team, B. I. (2014). *Four Simple Ways to Apply EAST Framework to Behavioural Insights | The Behavioural Insights Team*. https://www.bi.team/publications/east-four-simple-ways-to-apply-behavioural-insights/

Thaler, R. H., & Sunstein, C. R. (2008). *Nudge: Improving decisions about health, wealth, and happiness*. Penguin.

Tobias, R. (2009). Changing behavior by memory aids: A social psychological model of prospective memory and habit development tested with dynamic field data. *Psychological Review*, *116*(2), 408–438. https://doi.org/10.1037/a0015512

Vainre, M., Aaben, L., Paulus, A., Koppel, H., Tammsaar, H., Telve, K., Koppel, K., Beilmann, K., Uusberg, A., & others. (2020). Nudging towards tax compliance: A fieldwork-informed randomised controlled trial. *Journal of Behavioral Public Administration*, *3*(1).

ONLINE APPENDIX

Contents

INSTRUCTIONS

Thank you for coming to the experiment. Please do not talk with anyone and do not look at other people screens. Also, please turn off all your phones and personal electronic devices, and place all your personal items under your desk for the duration of this experiment. If at any moment you have any questions, please raise your hand. The experimenter will approach you and answer your questions privately.

Please read these instructions carefully. The instructions explain how you can earn money in this experiment.

All your decisions in the experiment will be anonymous. No one else during or after the experiment will learn about your decisions in the experiment. Your name and your identity will not be linked to your decisions.

Your earnings will be paid to you in cash at the end of the experiment. If you leave before the experiment finishes or if you break the rules, we will not be able to pay you your earnings.

16 ROUNDS OF THE SAME TASK

The experiment consists of 16 rounds. In each round, you will perform the same comparison task. The screen will show an inequality between two numbers ranging from 10 to 99. You will evaluate whether the presented inequality is true or false. Immediately after you submit your answer, a new inequality will show up.

Examples:

To successfully complete the task and earn income for the round, you must provide 50 correct answers and make less than 10 mistakes. In each round, the time limit to provide 50 correct answers is exactly 120 seconds.

EARNINGS

In each round, you will earn 100 tokens if you successfully complete the task (that is, if you provide 50 correct answers and make less than 10 mistakes). If you do not complete the task successfully, your income from the round will be 0 tokens.

Your income for each round is subject to a tax. In each round, the screen will therefore also show a "Report your income" button. If you click the button, the income you earn in that round will be reported for tax purposes. The tax is fixed and always equal to 30 tokens. If you do not click the "Report your income" button, you will not report your income and therefore you will not pay the tax in that round.

In each round, your income will be audited with a [10% / 60%] probability. If you have not reported your income and you are audited, you will pay a tax of 30 tokens and a penalty of additional 30 tokens.

The following table shows earnings for all possible combinations of reporting and audit if you earn your income by successfully completing the task. Note that if your income for the round is 0 (you did not successfully complete the task), there is no tax and no penalty. Thus, your earnings are 0, no matter whether you reported your income or not.

	Were you audited ([10% / 60%] probability in each round)	
Income reported?	Audited	Not audited
Reported	Income (100) – Tax (30) = 70 tokens	Income (100) – Tax (30) = 70 tokens
Not reported	Income (100) – Tax (30) – Penalty (30) = 40 tokens	Income (100) = 100 tokens

STRUCTURE OF EACH ROUND

In each round the "Report your income" button will only be active for a 30 second window. This window opens after 30 seconds from the start of the round and closes at 60 seconds from the start. The upper-right corner of the screen will display a timer.

Second 0	Task starts
Second 30	"Report your income" button becomes active
Second 60	"Report your income" button becomes inactive
Second 120	Task ends

WHEN YOU FINISH

At the end of the experiment, the software will show a summary for all 16 rounds. For each round you will learn whether you completed the task successfully, whether you reported your income, whether there was an audit and what were your resulting earnings.

Your cash earnings in this experiment will be the sum of your earnings from all 16 rounds. The exchange rate is 100 tokens = 1 EUR.

When you finish the experiment, please remain quietly seated until the experimenter calls your cubicle number. You will then go to the room at the back of the laboratory to privately collect your experimental earnings in cash.

If you have any questions, please raise your hand now.

COMPREHENSION QUESTIONS

1. How many rounds are there in the experiment?
   Correct answer: 16

2. How many correct answers do you need to provide in a round to earn your income?
   Correct answer: 50

3. How many incorrect answers in a round result in not earning an income?
   Correct answer: 10

4. In each round, there is a possibility of audit. What is the exact probability (in %)?
   Correct answer: [10 / 60]

5. In a given round, how many tokens do you earn if you successfully complete the task, report your income, and there is no audit?
   Correct answer: 70

6. In a given round, how many tokens do you earn if you successfully complete the task, report your income, and there is an audit?
   Correct answer: 70

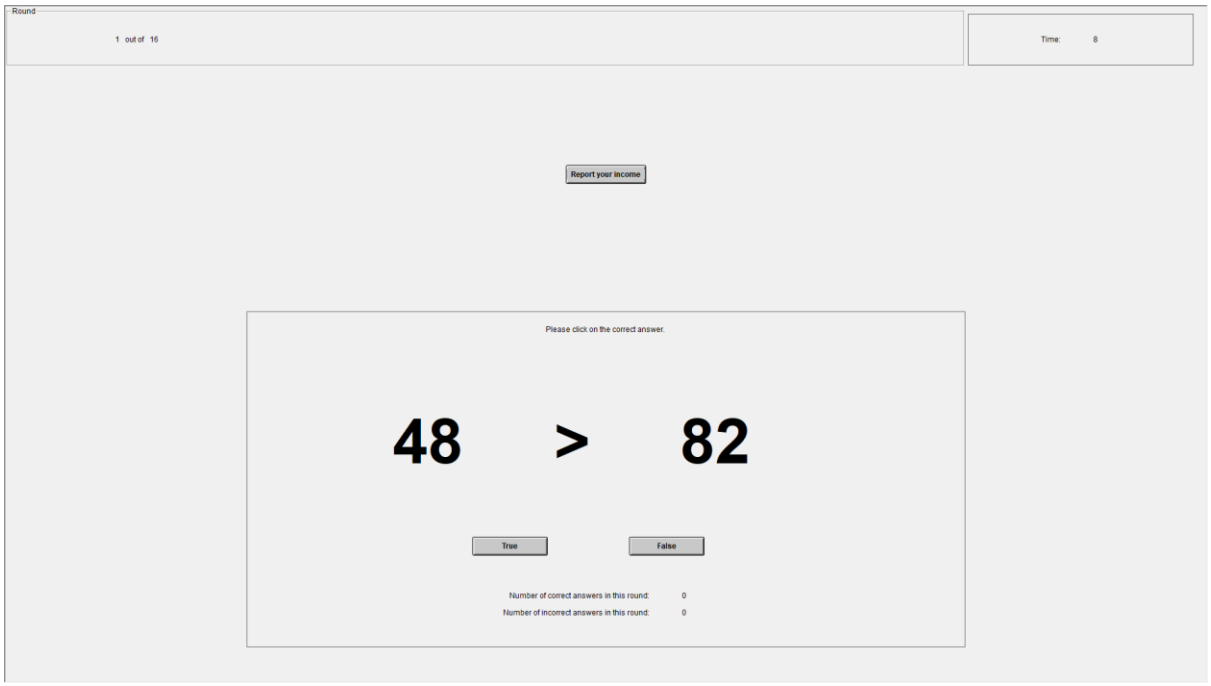
7. In a given, how many tokens do you earn if you successfully complete the task, do not report your income, and there is no audit?
   Correct answer: 100

8. In a given round, how many tokens do you earn if you successfully complete the task, do not report your income, and there is an audit?
   Correct answer: 40

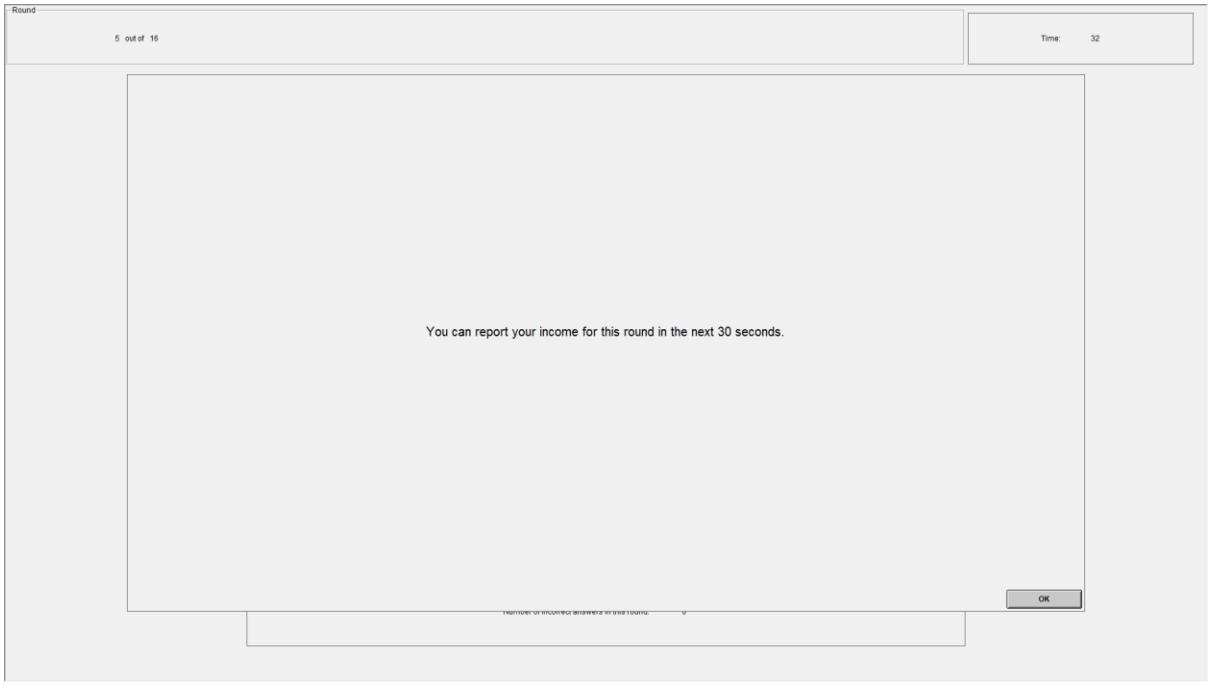
9. In a given round, how many tokens do you earn if do not successfully complete the task?
   Correct answer: 0

10. How much EUR is 100 tokens?
    Correct answer: 1

SCREENSHOT OF THE REAL EFFORT TASK



SCREENSHOT OF THE NUDGE



RISK ELICITATION PROCEDURE

Lotteries 1-2(3): Eliciting individual risk attitudes, risk-type classification

The first two (or three) lotteries estimate subject's risk aversion and categorize the subject by gamma coefficient to one of six risk aversion types (A-F).

<table>
<tr><td colspan="6">Lottery 1</td></tr>
<tr><td colspan="3">Lottery 1
Option A (γ<0.45)
40 tokens w/p 45% or 100 tokens w/p 55%</td><td colspan="3">Lottery 1
Option B (γ>0.45)
70 tokens w/p 100%</td></tr>
<tr><td colspan="3">Option A in Lottery 1
↓
Lottery 2</td><td colspan="3">Option B in Lottery 1
↓
Lottery 2</td></tr>
<tr><td colspan="2">Lottery 2
Option A (γ<0.21)
30 tokens w/p 40% or
100 tokens w/p 60%</td><td>Lottery 2
Option B (γ>0.21)
70 tokens w/p 100%</td><td>Lottery 2
Option A (γ<0.77)
40 tokens w/p 50% or
115 tokens w/p 50%</td><td colspan="2">Lottery 2
Option B (γ>0.77)
70 tokens w/p 100%</td></tr>
<tr><td colspan="2">Option A in Lottery 2
↓
Lottery 3</td><td>Option B in Lottery 2
↓</td><td>Option A in Lottery 2
↓</td><td colspan="2">Option B in Lottery 2
↓</td></tr>
<tr><td>Lottery 3
Option A
(γ<0)
40 tokens
w/p 50% or
100 tokens
w/p 50%</td><td>Lottery 3
Option B
(γ>0)
70 tokens
w/p 100%</td><td>↓</td><td>↓</td><td>Lottery 3
Option A
(γ<1.85)
40 tokens
w/p 30% or
100 tokens
w/p 70%</td><td>Lottery 3
Option B
(γ>1.85)
70 tokens
w/p 100%</td></tr>
<tr><td>↓</td><td>↓</td><td>↓</td><td>↓</td><td>↓</td><td>↓</td></tr>
<tr><td>γ<0</td><td>0<γ<0.21</td><td>0.21<γ<0.45</td><td>0.45<γ<0.77</td><td>0.77<γ<1.85</td><td>γ >1.85</td></tr>
<tr><td>Risk type A</td><td>Risk type B</td><td>Risk type C</td><td>Risk type D</td><td>Risk type E</td><td>Risk type F</td></tr>
</table>

Lotteries 4 (and possibly 5): Contrast manipulation

In Lotteries 4 and 5 we test for inconsistencies in risk preferences by adding either an extremely high or an extremely low payoff with probability 0.01. Such inconsistencies could be explained by the salience theory and should be observed for FAST subjects, but not for FBOR subjects.

	Type A ($\gamma<0$)	**Type B ($0<\gamma<0.21$)**	**Type C ($0.21<\gamma<0.45$)**	**Type D ($0.45<\gamma<0.77$)**	**Type E ($0.77<\gamma<1.85$)**	**Type F ($\gamma>1.85$)**
Lottery 4 (Challenging the upper bound of elicited gamma)						
Option A: Consistent choice	• 40t, 45% • 100t, 54% • 1t, 1%	• 35t, 40% • 100t, 59% • 1t, 1%	• 40t, 40% • 100t, 59% • 1t, 1%	• 40t, 30% • 100t, 69% • 1t, 1%	• 65t, 10% • 190t, 89% • 1t, 1%	**N/A**
Option B **FAST?**	• 70t, 100% Optimal if $\gamma>0.27$	• 70t, 100% Optimal if $\gamma>0.34$	• 70t, 100% Optimal if $\gamma>0.64$	• 70t, 100% Optimal if $\gamma>1.17$	• 70t, 100% Optimal if $\gamma>1.94$	**N/A**
Lottery 5 (Challenging the lower bound of elicited gamma)						
Option A: Consistent choice	**N/A**	• 70t, 100%	• 70t, 100%	• 70t, 100%	• 70t, 100%	• 70t, 100%
Option B **FAST?**	**N/A**	• 20t, 45% • 100t, 54% • 300t, 1% Optimal if $\gamma<-0.25$	• 30t, 45% • 100t, 54% • 300t, 1% Optimal if $\gamma<0.05$	• 30t, 40% • 100t, 59% • 300t, 1% Optimal if $\gamma<0.35$	• 40t, 45% • 100t, 54% • 300t, 1% Optimal if $\gamma<0.61$	• 40t, 35% • 100t, 64% • 300t, 1% Optimal if $\gamma<1.44$

Lotteries 6 (and possibly 7): Prominence manipulation

In Lotteries 6 and 7 we test for inconsistencies in risk preferences by using exactly the same options as in Lotteries 1-3, but attracting subject attention to one possible payoff of the lottery choice. Again, such inconsistencies could be explained by the salience theory and should be observed for FAST subjects, but not for FBOR subjects.

	Type A ($\gamma<0$)	**Type B ($0<\gamma<0.21$)**	**Type C ($0.21<\gamma<0.45$)**	**Type D ($0.45<\gamma<0.77$)**	**Type E ($0.77<\gamma<1.85$)**	**Type F ($\gamma>1.85$)**
Lottery 6 (Challenging the upper bound of elicited gamma)						
We highlight the worst potential outcome by explicitly stating that the worst possible outcome of the choice is X tokens which can occur only if subject chooses the Option A.						
Option A: Consistent choice	• 40 t, 50% • 100t, 50%	• 30t, 40% • 100t, 60%	• 40t, 45% • 100t, 55%	• 40t, 50% • 100t, 50%	• 40t, 30% • 100, 70%	**N/A**
Option B **FAST?**	• 70t, 100%	• 70t, 100%	• 70t, 100%	• 70t, 100%	• 70t, 100%	**N/A**
Lottery 7 (Challenging the lower bound of elicited gamma)						
We highlight the best potential outcome by explicitly stating that the best possible outcome of the choice is 100 tokens which can happen only if subject chooses the Option B.						
Option A: Consistent choice	**N/A**	• 70t, 100%	• 70t, 100%	• 70t, 100%	• 70t, 100%	• 70t, 100%
Option B **FAST?**	**N/A**	• 40 t, 50% • 100t, 50%	• 30t, 40% • 100t, 60%	• 40t, 45% • 100t, 55%	• 40t, 50% • 100t, 50%	• 40t, 30% • 100, 70%

COGNITIVE REFLECTION TEST QUESTIONS

1. If 3 kids can wrap 3 toys in 1 hour, how many kids are needed to wrap 6 toys in 2 hours?
   Correct answer: 3

2. Jerry received both the 15th highest and the 15th lowest mark in the class. How many students are there in the class?
   Correct answer: 29

3. In an athletics team, tall members are three times as likely to win a medal than short members. This year the team has won 60 medals so far. How many of these have been won by short athletes?
   Correct answer: 15

Abstrakt

Přestože nudging (postrčení) získal velkou popularitu, mnoho nudgingových intervencí selhává, a účinky těch úspěšných jsou často krátkodobé. Předpokládáme, že úspěch nudgingu závisí na tom, jak interaguje se základními ekonomickými motivacemi, které určují chování jednotlivce s cílem maximalizovat výnos. Například v oblasti dodržování daňových povinností je nudging pravděpodobně účinný pouze tehdy, pokud je pro daňového poplatníka finančně optimální zaplatit daně. Pro ověření našeho předpokladu provádíme multi-periodový experiment, ve kterém měníme pravděpodobnost daňové kontroly a nudgingem motivujeme účastníky k reportování jejich příjmů. Navíc měníme četnost, s jakou se nudging objevuje, abychom otestovali, zda častější náznakování zvyšuje reportování dlouhodobě. Pozorujeme, že první použití nudgingu má okamžitý pozitivní účinek na reportování příjmů bez ohledu na to, zda je to finančně optimální nebo ne. Avšak další nudging zvyšuje reportování příjmů pouze tehdy, pokud je nudging v souladu s motivacemi daňového poplatníka. Častější nudging opačným směrem než jsou pobídky nevede k žádným efektům na dlouhodobé chování. Diskutujeme o důsledcích pro implementaci veřejných politik.

Working Paper Series
ISSN 2788-0443

Individual researchers, as well as the on-line version of the CERGE-EI Working Papers (including their dissemination) were supported from institutional support RVO 67985998 from Economics Institute of the CAS, v. v. i.

Specific research support and/or other grants the researchers/publications benefited from are acknowledged at the beginning of the Paper.

Published by
Charles University, Center for Economic Research and Graduate Education (CERGE)
and
Economics Institute of the CAS, v. v. i. (EI)
CERGE-EI, Politických vězňů 7, 111 21 Prague 1, tel.: +420 224 005 153, Czech Republic.
Phone: + 420 224 005 153
Email: office@cerge-ei.cz
Web: https://www.cerge-ei.cz/

Editor: Byeongju Jeong

The paper is available online at https://www.cerge-ei.cz/working-papers/.

ISBN 978-80-7343-584-4 (Univerzita Karlova, Centrum pro ekonomický výzkum a doktorské studium)
ISBN 978-80-7344-712-0 (Národohospodářský ústav AV ČR, v. v. i.)