A Work Project, presented as part of the requirements for the Award of a Master's degree in Economics from the Nova School of Business and Economics.

THE EQUIVALENCE OF INCOME AND CONSUMPTION TAXES: ARE MONEY ILLUSION, LOSS AVERSION AND SALIENCE THE SPOILSPORTS?

Tobias Tesing

Work project carried out under the supervision of:

Alexander Coutts

04-01-2021

The equivalence of income and consumption taxes: Are Money Illusion, Loss Aversion and Salience the Spoilsports?

Tobias Tesing

Abstract

Classic economic theory suggests the equivalence of sales and income taxation. However, Blumkin, Ruffle, and Ganun (2012) find in a real-effort experiment that subjects reduce their labor supply more in response to an income tax than to an equivalent sales tax. Using an adjusted experiment design, this paper aims to (i) replicate this result, (ii) test whether money illusion and/or loss aversion are the underlying behavioral biases, and (iii) examine the impact of the non-salience of sales taxes on labor supply. My difference-in-difference estimates do not confirm the results in Blumkin, Ruffle, and Ganun (2012). Moreover, money illusion, loss aversion and the non-salience of sales taxes do not appear to influence labor-leisure decisions. I argue that my adjusted experimental design may represent a more realistic approach and that the effect observed in Blumkin, Ruffle, and Ganun (2012) may be too small to have real-world implications.

Keywords: Taxation, Money Illusion, Loss Aversion, Behavioural Economics, Tax Salience

This work used infrastructure and resources funded by Fundação para a Ciência e a Tecnologia (UID/ECO/00124/2013, UID/ECO/00124/2019 and Social Sciences DataLab, Project 22209), POR Lisboa (LISBOA-01-0145-FEDER-007722 and Social Sciences DataLab, Project 22209) and POR Norte (Social Sciences DataLab, Project 22209).

1 Introduction

The tax system is a cornerstone of every economy. Taxes serve as a means for redistribution of income and wealth, not only from individual to individual but also across time. They allow governments to provide fiscal stimuli during bad times to smooth consumption and maintain existing (infra-)structures throughout recessions. During the current Coronavirus Crisis, governments worldwide are breaking records in their deficits, which are to be financed by future taxation[1]. Two major sources of state revenue are income and consumption taxes, which make up about 24% and 32%, respectively, of tax revenue in OECD countries (see OECD 2020a).

Given that a government wants to make those potentially welfare-enhancing expenditures, it is keen to find an efficient and equitable way to tax individuals. In this light, Blumkin, Ruffle, and Ganun (2012) (hereafter BRG) conduct a real-effort experiment which reveals that, contrary to standard economic theory, individuals' labor supply is more sensitive to the introduction of an income tax than to an equivalent sales tax, in a setting where both taxing schemes create the same budget line and, thus, equal incentives. While this finding suggests efficiency gains from shifting taxation from income to consumption, the authors do not test for the underlying reason for the observed behavior.

The present paper first aims to replicate the results in BRG and explore possible explanations for the diverging perceptions of the two taxation types. A more detailed insight into the underlying decision-making processes may give rise to better exploitation of the present biases in real life. First, I consider BRG's hypothesis of money illusion, the tendency to think in nominal rather than real terms. Following an income tax change, a money-illusioned individual would correctly adjust his labor supply because nominal and real wage change equally. In contrast, he would under-adjust after a sales tax change since he does not fully perceive his real wage change. Second, I suspect that individuals disproportionally dislike the explicit loss from gross to net income compared to an equivalent burden from gross to net prices. This hypothesis refers to the well-established concept of loss aversion outlined in Kahneman and Tversky (1979). Third, I connect the result of BRG with another key finding that consumers under-react to sales taxes

1. See projected household deficits for a double-hit scenario in OECD (2020b). For the first time in its history, the EU issues joint debt and distributes unprecedented transfers between member states (see European Commission 2020)

when they are not displayed in prices (Chetty, Looney, and Kroft 2009). Specifically, if workers underreact to sales taxes, it seems plausible that the non-salience of those taxes intensifies the underreaction. I test my hypotheses using a real-effort experiment in a set-up similar to the one of BRG. Participants face voluntary tasks, where every fulfilled task indirectly increases the probability of winning a prize. The experiment consists of two rounds of completing tasks and making consumption decisions in a virtual supermarket, where a tax increase between rounds is implemented.

The debate about whether and how to tax income and consumption has a long history in economic theory. Rules of thumb for efficient consumption taxation schemes derived by economists over the past century earmark low tax rates on wide tax bases, preferably on goods with inelastic supply and demand. On the other side, using income as a proxy for ability, income tax schemes may provide an efficient and equitable means to distribute the total tax burden across individuals in the economy. However, income taxation does not only indirectly tax consumption but also savings, which distorts intertemporal consumption decisions[2]. Facing this trade-off, most countries in the world tax both income and consumption. The present study adds to the exploration of this trade-off as it challenges BRG's argument in favor of consumption taxation in an adjusted setting and examines whether the framing of consumption taxation influences its distortionary impact in labor markets. Furthermore, the existence of salience effects in this context would suggest efficiency gains from making sales taxes non-salient because it would less distort individual labor-leisure decisions.

Importantly for taxation, a standard assumption of economic theory is Money Neutrality, that individuals base their decisions solely on real values and not on nominal values. Accordingly, equivalent reductions in real income should always yield equal reactions by individuals, regardless of the source of real income reduction: either a nominal price increase (sales tax increase) or a nominal income decrease (income tax increase). However, there is substantial evidence that this assumption may not always hold. Money illusion, first described in Fisher (1928) as the "tendency of people to think of currency in nominal, rather than real, terms", has been found in a variety of contexts, including stock markets (Cohen, Polk, and Vuolteenaho 2005; Schmeling

2. There is an immense literature on optimal taxation which is not the focus of this paper, see Sandmo (1976) or Mankiw, Weinzierl, and Yagan (2009) for overviews of this literature

and Schrimpf 2011), housing prices (Brunnermeier and Julliard 2008) and securities markets (Basak and Yan 2010). In an experimental study, Fehr and Tyran (2001) find money illusion to be the underlying reason for sluggish price adjustments to negative monetary shocks, though Petersen and Winn (2014) largely rebut those findings in an adjusted experimental setting. Similarly, Fehr and Tyran (2007) show in another experiment that money illusion prevents the market from reaching a Pareto-efficient equilibrium. For a broader, psychological account of money illusion see Shafir, Diamond, and Tversky (1997).

To my knowledge, this is the first paper to isolate the possible effect of money illusion on labor supply and establish a causal link between them[3]. The existence of money illusion in labor markets would have implications for both governments and central banks. First, governmental interventions that influence prices may have different effects than classical theory expects. Beyond the field of taxation discussed in this paper, for example, inflation adjustments of minimum wages may have a stronger positive impact on labor supply than typically assumed because workers only perceive the wage increase but not the price increase. Second, central banks would have to take money illusion into account considering the impact of monetary policy on labor markets and, consequently, on the economy as a whole.

An additional treatment group, subject to a non-salient sales tax, in the present experiment intends to give new insights into the possible indirect effects of non-salient sales taxes on labor supply. Similar to money illusion, several studies show that individuals do not fully perceive the price level when taxes or fees are not entirely transparent, despite their awareness of the tax or fee. Comparing salient and non-salient taxation outcomes, these studies show how individuals under non-salient taxation either consume more or tolerate higher taxes. Xia and Monroe (2004) explore how the partitioning of prices in online shops, for example, into base price, shipping fee and handling fee, can enhance consumers' purchase intentions, perceived value and price satisfaction. In a field experiment in supermarkets in the US, where sales taxes are typically not included in displayed prices, Chetty, Looney, and Kroft (2009) find that sales drop by 8 percent when shelf prices include the sales tax. Finkelstein (2009) finds a 20-40 percent

3. There are regression studies on the influence of money illusion on the labor market (for example Fair (1971), Wachter (1972), Hadley (1982), Gustafson and Hadley (1989)). However, they focus on estimating correlations and do not provide clear causal evidence.

increase through the automation of toll collection compared to a more salient manual collection. Exploring the US property market, Cabral and Hoxby (2012) show that higher tax escrow rates – and, thus, lower tax salience for sellers - come with higher regional property tax rates. The extent of non-salience effects in the labor market, however, is somewhat controversial. While Abeler and Jäger (2015) find underreactions to tax changes in complex rather than in simple tax systems, participants in the lab experiment of Fochmann and Weimann (2013) underreact most to low and simple labor-income taxes. Here, complexity and size tend to increase the salience of the tax. Similarly, in a lab experiment of Hayashi, Nakamura, and Gamage (2013), participants under a wage tax work even less compared to a group under no tax but with equal net wage. However, when all-inclusive wages are made transparent, the differences disappear, implying a dampening effect of non-salience on labor supply. The present paper investigates a new channel of salience effects on labor supply, testing whether the salience of sales taxes influences labor decisions.

The experiment consists of four treatment groups that go through two rounds of labor and consumption decision. Subjects complete tasks to earn a budget for a virtual supermarket. The product choices made in this supermarket determine participants' payments. Each subject is randomly assigned to one of the treatment groups: salient sales tax (SST) group, non-salient sales tax (NST) group, money illusion (MI) group and income tax (IT) group. The former three are subject to a sales tax increase and, while the latter experiences an income tax increase between rounds. In contrast to the SST group, taxes are not included in the displayed prices for the NST group. In addition to the SST treatment, participants in the MI group see the increased prices of all goods in round two during work. To examine the influence of loss aversion on labor supply outcomes, participants answer six questions that measure their individual degree of loss aversion at the end of the experiment.

The results from a difference-in-difference estimation reveal no significant differences between all treatment groups. This holds for the whole sample on average and in the individual inspection of participants who either increase or decrease labor supply in response to the tax increase. Thus, my results do not confirm the finding of BRG that individuals decrease their labor supply more in response to an income tax than to a sales tax. Moreover, money illusion and the non-

salience of sales taxes do not appear to play a role in the labor-leisure trade-off participants face, although there are signs of underreaction to the non-salient sales tax at the moment of tax collection. Equivalently, the degree of loss aversion of a participant in the income tax group does not relate to her labor supply. The results are robust to the exclusion of outliers and logarithmic regression modeling.

The remainder of this paper is structured as follows: Section 2 describes the experiment design and its underlying logic, section 3 presents summary statistics and difference-in-difference estimation results, section 4 discusses differences to the experiment of BRG and possible implications of my results, and section 5 concludes.

2 Experiment Design

This real-effort experiment is inspired by and extends the experimental design in BRG, though resource limitations led to some adjustments of the original setup. My experiment consists of two rounds, where each round is split into two parts: a labor and a consumption decision. In the labor part, participants complete tasks to earn tokens, which serve as a budget for the consumption decision in the second part. The outcome of interest is the change in labor supply following a tax increase from round 1 to round 2, which is economically equivalent for all treatment groups.

The composition of consumption goods from a randomly chosen round determines participant payments in lottery tickets for a lottery to win amazon vouchers. The random choice avoids satiation effects and participants exerting all effort in one round to skip work in the other round. The number of tickets achieved in each round is only presented to the participant at the end of the experiment to avoid any learning effects concerning the supermarket.

In the first part of both rounds, the task for each participant is to count zeros in a table of 100 randomly placed zeros and ones, after the example of Abeler et al. (2011) and displayed in Figure 1. For each correctly indicated number of zeros, participants receive a wage of 9 tokens. This task does not require any prior knowledge and is relatively simple. Nevertheless, it is cognitively demanding, pointless and likely unpleasant for my subjects, thus creating an incentive to skip

working. Participants are free to stop working at any point, with a maximum of 50 tables per round. All tables are equal across treatment groups but not across rounds.

Figure 1: Example 0-1-Table

1	1	0	0	1	1	0	1	1	1
1	0	1	1	1	1	1	1	0	1
1	1	1	1	1	1	0	0	1	0
1	1	1	0	0	1	0	0	1	0
1	1	1	1	1	0	1	1	1	1
1	1	1	1	1	1	1	1	0	1
1	1	1	1	1	1	1	1	1	1
1	1	1	0	1	1	1	1	0	1
1	1	1	1	1	1	1	1	1	1
0	1	1	1	0	1	1	1	1	0

To familiarize all participants with the task, every subject must go through an introductory with three tables to specify, before round 1 and measurement of relevant metrics starts. Moreover, all participants must correctly answer two questions on their wages and tax levels before proceeding to the labor part in each round. These comprehension questions ensure that subjects correctly perceive all necessary factors to make an economically rational decision about the amount of labor they want to exert.

The virtual supermarket offers twelve everyday products, ranging from toothpaste to bananas. Prices range from 3 to 14 tokens. The purchase of each product yields a certain amount of lottery tickets in return, where every purchased unit of one product returns one ticket less than the previous unit of the same product. Moreover, combining products that are typically consumed together gives extra points, again with one combination returning one ticket less than the previous combination. A penalty of 5 tickets per budget exceeding token is to prevent participants from overshooting their budget.

A total of 217 Master and Bachelor students in Economics, Management and Finance participated in the experiment, with 212 enrolled at the Nova School of Business and Economics, Lisbon, and the remaining five at Copenhagen Business School, TU Berlin and HEC Montréal. The experiment took place from November 2nd to 8th, 2020, via an online Qualtrics survey. All participants were instructed not to use any tools (for example, calculators) during the experiment.

Screenshots of the virtual supermarket, explicit instructions and rules can be found in Appendix A.

2.1 Treatments and underlying logic

Three of the four treatment groups, namely SST, MI, and NST groups, are subject to a 20% sales tax in the first round and an 80% sales tax in the second round. The IT group does not face sales taxes but an income tax of 16.67% (1.5 tokens per table) in round 1 and 44.44% (4 tokens per table) in round 2. Both taxing schemes create equal post-tax budget lines and, thus, equal incentives. Consequently, all groups should show equal labor supply responses to the tax increase, on average, if participants are fully rational. Comparing sales taxation, the tax change in my experiment is 40 percentage points smaller than the one in BRG, yet it implements a far greater tax change than any real-world policy. If this change is not big enough to reveal differences in reactions across treatment groups, the differences may be too small to be relevant for real-world implications, if there are any.

Table 1: Treatment Summary

Group	Treatment		Expected Labor Supply Change compared to SST
	Round 1	Round 2	
SST	20% Sales Tax, included in prices	80% Sales Tax, included in prices	-
IT	16.67% Income Tax	44.44% Income Tax	↓ decrease more, ↔ equal when controlling for loss aversion
MI	20% Sales Tax, included in prices	80% Sales Tax, included in prices, new prices printed to screen during work	↓ decrease more, equal to IT
NST	20% Sales Tax, not included in prices	80% Sales Tax, not included in prices	↑ decrease less

Table 1 provides a quick overview of all treatments and expected results. The SST group and the IT group are conceptually equal to the two treatment groups in BRG; only the tax levels are

different. The main extension of their experiment consists of the following elements that aim to provide evidence on the underlying reasons for their finding.

2.1.1 Money Illusion

BRG explain their result through the presence of money illusion. If this is indeed the case, participants in the SST group do not fully perceive the real wage decrease following a tax increase. The money illusion treatment intends to tighten the gap in the perception of real wage changes between the SST and IT groups and hint at participants' real wages under a sales tax.

A money-illusioned participant considers his nominal wage rather than his real wage in his labor decision. Following a tax increase, a money-illusioned participant in the IT group adjusts his labor supply because of his nominal wage decrease. In his case, the nominal wage equals the real wage since the price level is constant. Thus, he bases his decision on a real wage change and takes the economically rational decision, although he suffers from money illusion. In contrast, a money-illusioned participant in the SST group does not fully adjust his labor supply. While he also bases his behavior on his nominal wage, his nominal wage is equal before and after the tax increase. He only adopts a behavioral change to the degree he perceives the price increase and the resulting real wage decrease.

Therefore, the money illusion treatment aims to strengthen the perception of the price change by printing the increased prices of all goods above all tables that participants work on in the second round. The direct hint at the tax increase makes it as visible as possible while subjects form decisions on how much to work. Participants in this group should then rather consider their real wage than their nominal wage and adjust their labor supply, on average, just as much as the participants in the IT group do. There should be no difference between the MI and the IT group observable in a difference-in-differences estimation.

2.1.2 Non-salient Sales Tax

Equal to the SST treatment, the NST group is subject to a 20% sales tax in round 1 and 80% in round 2. The only difference between groups is that under the NST, the tax is not included in the price printed below each product in the virtual supermarket. The comprehension questions before each working round and an additional hint in the supermarket make sure that participants

indeed know about the tax. As noted, previous studies (Chetty, Looney, and Kroft 2009; Finkelstein 2009; Cabral and Hoxby 2012) suggest that participants will underreact to the sales tax in the supermarket because they are not fully aware of it. If they are less aware of the tax change, they may also reduce their labor supply less in the second round compared to the SST group. In BRG, participants reduce labor supply in response to the salient sales tax, on average. At least to some degree, they are aware of the tax and the resulting real wage change and take it into account when deciding on how much to work. The documented underreactions to non-salient sales taxes are a sign of not being fully aware of the tax. Hence, participants under a non-salient sales tax may be less aware of the real wage decrease when forming their labor decision in the second round. If the differing labor supply responses in sales tax and income tax schemes are due to an underreaction to sales taxes (money illusion), rather than an overreaction to income taxes (loss aversion), the non-salience of sales taxes may translate into the labor market and intensify the underreaction to the tax.

2.1.3 Loss Aversion

To study the relevance of loss aversion in my experiment, I do not introduce another treatment group because of the limited number of expected participants. Instead, every subject must answer six questions at the end of the experiment to measure her loss aversion. In each question, the participant must decide if she accepts or rejects a lottery, with an equal probability of 50% of winning 6€ or losing x€, with x ranging from 2 to 7. The number of rejected lotteries serves as a measure of loss aversion.

Based on prospect theory of Kahneman and Tversky (1979), I hypothesize that the labor-income tax group in BRG reduces its labor supply more because it disproportionally dislikes the loss to the gross wage associated with an additional unit of labor supply. In the context of income taxation[4], one may argue that the gross wage serves as a reference point and that, consequently, individuals perceive their net wage as a loss relative to the gross wage. In contrast, the reference point for consumption with salient sales taxes is the net price consumers see on the shelves (outside North America). Hence, there would be no space for loss aversion to having an effect

4. Although loss aversion is widely studied, examples in the context of taxation are rare. Engström et al. (2015) provide evidence on the influence of loss aversion on tax compliance in Sweden. Rees-Jones (2014) finds similar behavior in the US.

with consumption taxes. Moreover, the payment of consumption taxes always comes with the reception of a good or service. Assuming that individuals value the purchased good or service higher than its net price, consumers receive a gain every time they pay consumption taxes. The tax reduces this gain, but individuals do not perceive this reduction as harsh as an equivalent reduction of their wage. When paying income tax, individuals usually also receive public goods in return. However, individuals may not value the public goods higher as their taxes paid, and the gain is not as visible as it is with consumption. Most strikingly, though, individuals also receive public goods in return for paying consumption taxes. Hence, as income taxes may be perceived as a loss while consumption taxes may not, loss aversion could drive the different results for two otherwise equivalent tax schemes. Controlling for loss aversion in the IT group may explain differences in labor supply between the IT and SST groups.

In contrast to BRG, all treatment groups are subject to taxation in both rounds. This ensures that participants in the NST group experience the non-salience of taxation before the tax increase between rounds 1 and 2 is implemented. Otherwise, if there was no taxation in round 1, the NST group would only be exposed to their treatment at the end of round 2, after all labor decisions are made. This inevitably leads to differing treatments already in round 1, especially between the IT group and the other groups that receive a sales tax treatment. However, because the taxation type is constant across rounds, possible differences in the first round cancel out in a difference-in-difference estimation.

The experimental supermarket fulfills several purposes, although supermarket outcomes are only of secondary interest in this study. Its main function is to serve as a place where sales taxes can be applied, enabling the comparison of income and sales tax treatments. Moreover, in combination with the fairly complex rules on the final lottery ticket return, it makes the whole experiment more exciting for participants. It distracts the focus of subjects from the working task. On the one hand, this aims to make participants take their labor decision more intuitively and naturally. On the other hand, it obscures the experiment's actual goal, which prevents participants from trying to please the experimenter with a behavior he expects, that is experimenter

demand effects[5]. Prices and lottery ticket returns are necessarily arbitrary, though the combination of rules, prices and returns is intended to complicate the maximization of lottery tickets. This obstruction reflects a real supermarket situation, where decisions are often difficult and arbitrary due to information asymmetries, lack of mental focus or the current emotional state[6].

3 Results

For the analysis of my experiment results, I use a difference-in-differences estimation approach to control for heterogenous time-constant factors. The estimated differences between groups and rounds must be due to factors that change in the course of the experiment, that is an increase in a salient sales tax (SST group), an increase in the income tax (IT group), an increase in a salient sales tax and its emphasis during the work task (MI group) or an increase in a non-salient sales tax (NST group). All groups receive economically equivalent tax increases but under different framings. All possible differences in labor supply between taxation types in round 1 are accounted for because the taxation type is constant across rounds and initial differences cancel out.

I remove some observations in my analysis because participants show signs of misunderstanding or inconsistency. I drop ten participants who exceed their budget in the supermarket by more than 100% in one of the two rounds and appear not to understand the whole experiment. Moreover, 15 participants indicated the number of tokens they want to spend on a supermarket product rather than the number of products they want to buy. I do not consider their choices in the supermarket but keep their labor decisions because they seem to understand the experiment in general. Moreover, I define an ability variable that contains the number of correctly solved tables per second in the introductory round. In case ability levels differ, this measure allows to control for possible interactions of ability and treatment type[7] different ability levels between

5. See Zizzo (2010) for a literature review on experimenter demand effects in economic experiments

6. For example, it is well documented how the amount of calories purchased depends on the degree of hunger. For a literature review on how emotions influence consumption decisions see Achar et al. (2016).

7. for example, it may be the case that participants with high ability behave more rational in general. Given that differences between SST and IT group arise from an underreaction to sales taxes, a high-ability participant in the SST group may reduce his labor supply more (the rational behavior) than a low-ability participants in the same group

treatment groups. Finally, I code the loss aversion measure of eleven participants as missing because they either reject the lottery at low stakes and accept at high stakes, or switch from accepting to rejecting and back to accepting.

3.1 Summary Statistics

Table 2 presents summary statistics of the experiment. Group sizes range from 49 to 57. Analyses of variance do not reject the assumption of equality of means in loss aversion and ability at p-values of 0.62 and 0.81, respectively, indicating that random group assignment led to similar groups. Although the difference-in-differences approach controls for initial differences, the similarity of groups is nonetheless reassuring. Moreover, controlling for interactions of ability and treatment appears unnecessary. The high p-value stated before reflects the similarity of ability levels across groups. Including a corresponding control in regressions would, therefore, rather increase estimation uncertainty than reducing any bias.

Table 2: Summary Statistics

Treatment Group		SST	IT	MI	NST
Labor Supply	Round 1	18.14	21.10	14.29	19.52
	Round 2	16.49	19.41	13.02	17.04
	Δ	-1.65	-1.69	-1.27	-2.48
Share of overdrafters	Round 1	22.8%	22.5%	26.5%	36.5%
	Round 2	17.5%	20.4%	18.4%	28.9%
	Δ	-5.3	-2.1	-8.1	-7.6
Relative Overdraft (of overdrafters)	Round 1	27.1%	12.5%	13.1%	20.5%
	Round 2	19.5%	9.0%	24.9%	40.3%
	Δ	-7.6	-3.5	+11.8	+19.8
Loss Aversion		3.42	3.46	3.24	3.52
Ability		.0461	.0463	.0437	.0452
Observations		57	49	49	52

Note: Relative Overdraft is defined as overdraft divided by budget. All displayeed values are averages of the respective groups (and rounds, if applicable).

A quick look at average labor supply shows that subjects in the MI group work the least in both

rounds, with a significant difference compared to all other groups at the 5% level in round 1. Because treatments are not yet very different in round 1, especially comparing MI, SST and NST groups, this outcome comes as a surprise and may serve as a reminder that chance false positives are possible.

Also, it turns out that the IT group worked the most in both rounds, on average. This is at odds with the results of BRG, which predict that participants should work less under an income tax than participants under a sales tax. Although the significantly higher labor supply compared to the MI group may just be a statistical artifact, this already casts some doubt on the robustness of BRG's results. Moreover, BRG would predict a significant difference in the change of labor supply between SST and IT groups. However, this difference is only 0.04 in my experiment, where average labor supplies range from 21.10 to 13.02.

Looking at decisions taken in the supermarket, it appears that participants in the NST group underreact to the sales tax compared to the other groups. Over the two rounds, both the share of subjects who exceed their budget and the relative overdraft are significantly higher in the NST group. This result opens the possibility for a spillover of this underreaction into the labor market. However, in contrast to my expectation, those participants show the strongest reaction to the tax change, on average, with a drop in average labor supply of 2.48.

Figure 2 gives a more detailed insight into the dynamics of labor supply changes and splits the sample into participants who decreased or increased their labor supply (participants who did not change are omitted). This distinction is important because there are two possible reactions to a tax increase, depending on the size of income and substitution effect of each participant: she may increase labor supply because she wants to maintain her previous consumption level (income effect); or she may decrease her labor supply because the relative price of leisure relative to consumption decreased (substitution effect). When reactions to both sides are possible, looking at the average may be misleading to find whether one group or the other is more sensitive to tax changes.

The split shows a slightly different picture than Table 2. Although none of the differences between groups are significant, the SST group shows the harshest reaction in both directions, a finding that disappears in the average. Again, this does not confirm expectations that arise from

BRG, who suggest bigger downward reactions in the IT group. Moreover, the relatively large drop in average labor supply in the NST group appears to stem from comparably low increases of those participants who increased their labor supply in this group.

Figure 2: Labor Supply Change Split

Note: Increasers (Decreasers) are participants who increased their labor supply from round 1 to round 2. Participants who did not change labor supply are not included in this figure.

Taken together and in line with BRG, Figure 2 and Table 2 imply that the substitution effect dominates the income effect, on average. The relative price decrease of leisure influences labor decisions more than the willingness to keep the previous consumption level. Participants believe that working is not worth their time as much after the tax increase and do not want to compensate for the lower (real) wage by working more. However, the labor supply decrease is not significant in my experiment, neither for the individual groups nor for the whole sample.

3.2 Regression Results

Table 3 presents difference-in-differences regression results with the SST group as the baseline for the whole sample, only increasers and only decreasers. Regressions (1.1), (2.1) and (3.1) display simple differences in labor supply changes between groups. In addition, regressions (1.2), (2.2) and (3.2) include second-round-controls for loss aversion and an interaction term of loss aversion and the IT group dummy to explain differences in the IT group caused by loss

aversion.

Regarding the main predictions, I fail to replicate BRG's result. Column (1.1) shows no significant difference in the change of labor supply in response to a sales tax increase or an income tax increase. On the contrary, the small coefficient estimate and the large standard error suggest that reactions in both groups are basically equal, on average. Even though coefficients remain insignificant when splitting the sample into increasers and decreasers, (2.1) and (3.1) show tendencies contrary to my expectation. Increasers in the IT group increase their labor supply less, and decreasers decrease their supply less than in the SST group. Overall, the sales tax increase tends to provoke harsher responses than the income tax increase.

Table 3: Difference-in-Differences Regression Results

Sample	All		Increasers		Decreasers	
Labor Supply	(1.1)	(1.2)	(2.1)	(2.2)	(3.1)	(3.2)
Income	-0.04	8.20	-1.37	0.22	2.41	-0.31
	(4.20)	(7.25)	(5.46)	(8.94)	(5.12)	(7.50)
Money Illusion	0.38	0.96	-1.06	-1.10	3.42	3.52
	(3.89)	(3.99)	(5.80)	(6.04)	(4.59)	(4.83)
Non-salient	-0.83	-0.85	-3.69	-4.07	1.42	1.36
	(4.44)	(4.54)	(6.32)	(6.61)	(4.95)	(4.98)
*Income*Loss Aversion*		-2.20		-0.56		0.92
		(1.59)		(2.10)		(1.61)
Loss Aversion		1.09		0.27		0.42
		(0.80)		(1.41)		(0.79)
Observations	414	392	162	152	186	178
R^2	0.03	0.04	0.07	0.08	0.21	0.20

*** p<0.01, ** p<0.05, * p<0.1

Note: Robust standard errors in parentheses. Estimates reflect differences in differences across rounds between the displayed treatment groups and the baseline of the Salient Sales Tax group. Increasers (Decreasers) are participants who increased (decreased) their labor supply from round 1 to round 2.

Because there is no difference between the SST and IT groups, money illusion and loss aversion cannot explain any difference. This does not consequently mean that money illusion and loss aversion do not influence labor decisions. However, the estimation results do not support the hypothesis that these two biases affect labor supply responses. None of the estimates on the

difference in the change of labor supply between the MI group and the SST group is significant, and the subsample results suggest that the SST group rather responds with stronger reactions.

Controlling for loss aversion in the IT group in the whole sample reveals that participants in the IT group with high loss aversion decrease their labor supply more than those with low loss aversion, on average. While this estimate points in the expected direction (loss averse participants tend to avoid explicit losses from the income tax), it is insignificant and its effect becomes less visible when splitting the sample.

Finally, the underreaction to non-salient sales taxes observed in Table 2 appears not to translate into labor supply decisions. The NST group shows the most distinct reaction compared to the SST group in the whole sample, but again, the difference is insignificant and the reactions tend to be weaker when splitting the sample.

Robustness checks in log transformation and excluding participants who worked on 50 tables in both rounds or zero tables in the first round yield qualitatively equal results and can be found in Appendix B.

4 Discussion

The results of my experiment contradict those of BRG. In this section, I examine several reasons for the differing experiment outcomes and discuss the implications of my results.

The most compelling difference between the two experiments is the incentive structure. First, BRG incentivize leisure through the reception of one bottled soft drink per 15 seconds of leisure taken. Their approach of limiting a work round to three minutes forces them to pay leisure because the intrinsic value of a fraction of three minutes off work in a behavioral laboratory tends to zero. Instead, my experiment design lets participants work until they do not want to work anymore and should reflect a more natural trade-off between labor and leisure. Second, the bridge from labor over consumption to utility is more apparent in BRG. Every correctly solved task leads to immediate gratification in form of a voucher for a sandwich or a pizza slice in their experiment. In contrast, my participants face a complex maximization task in the virtual supermarket with the budget they earn and receive (only) a chance to win a reward. Again, my

approach appears to reflect real-life better because it includes the complexity of maximizing utility (lottery tickets) between labor and consumption.

Another apparent difference between experiments is the initial taxing status and the size of the tax increase. BRG start from zero taxation and then implement a 100% sales tax (50% income tax), whereas my participants are subject to a sales tax of 20% (income tax of 16.7%) in round 1 and a subsequent tax increase of 60 (27.7) percentage points. The tax increase in BRG is not only 40 percentage points larger but may also invoke a tax aversion effect[8] in the IT group. A corresponding effect would cancel out in my set up.

A recent study by Kurokawa, Mori, and Ohtake (2020) also adds some doubt to the external validity and relevance of BRG's results. In their experiment, participants can choose between an income or a sales taxing scheme. The authors test all pairs of 20%, 22% and 25% taxation of income and sales. It turns out that participants favor the lower nominal tax rate, although it may come with a higher tax burden as is the case with 20% income taxation vs. 22% sales taxation. Because sales taxes are in nominal terms always higher than economically equivalent income taxes, this finding suggests that sales taxes rather discourage economic activity than equivalent income taxes since individuals perceive them as a greater burden.

Furthermore, I do not only find no difference between the SST group and the IT group. It also turns out that participants do not suffer from money illusion in this context, which BRG speculate is the reason for their result. The lack of explanation for a wedge between labor supply levels in the different taxing schemes casts further doubt on the findings of BRG. In addition, my results add to the discussion on the role of money illusion in an economy and do not support the idea to account for money illusion effects in economic labor market modeling in central banks, governments or academics. In contrast to my experiment, previous studies (Fair 1971; Wachter 1972; Hadley 1982; Gustafson and Hadley 1989) only focused on estimating correlations and did not provide clear causal evidence.

The outlined observations hint at the possibility that the difference in reactions to income and sales taxes may, in fact, be negligible, and may be an artifact of the relatively strong direct in-

8. Kessler and Norton (2016) show in a lab experiment that participants decrease their labor supply more in response to a wage tax than to an equivalent wage drop. However, they do not test whether this effect is constant or increases with the tax.

centives and disproportionately large tax increases of BRG. Witness to those two factors may be the overall decrease in labor supply, which BRG finds significant, while it is not significant in my experiment. To the extent that my setting is more realistic, even with its large tax increases, it may suggest that a failure to find differences between treatment groups is an indication that the effects would be negligible in the real world. Moreover, my larger sample size, in total and per treatment group, would even allow to detect smaller differences at a given significance level. Overall, it seems that more studies are needed to validate or disprove BRG's and my results and their relevance, favorably in field or natural experiments.

I also do not find signs of overreactions to an income tax due to loss aversion. However, I hypothesize that loss aversion may still be at play in my experiment. On the one hand, participants may want to avoid a realization of a loss in their net wage relative to their gross wage and reduce labor supply. On the other hand, they may also be averse to a loss in total income from round 1 to round 2. Camerer et al. (1997) provide evidence for similar behavior at the example of New York cab drivers, who tend to set daily total income goals and work until they reached that goal, regardless of their hourly wage on that day. A lower wage rate would, therefore, lead to increased labor supply. Taken together, loss aversion may provoke both labor increasing and decreasing behaviors which, finally, outweigh each other.

5 Conclusion

I conduct a real-effort experiment to find the underlying reasons for differing reactions in labor supply to an introduction of income or sales taxation observed in Blumkin, Ruffle, and Ganun (2012). My initial hypothesis is that money illusion and/or loss aversion introduce a wedge between the two groups. It turns out that there is no difference in reaction in the sales tax group and the income tax group in my experiment. Moreover, I do not find evidence that money illusion or loss aversion influence labor supply decisions. This finding adds to the discussion on the relevance of money illusion in the economy and suggests that it is negligible in the labor market context.

I explore possible reasons for the diverging results of the present paper and BRG and find that the

difference in labor supply responses to income and sales taxation may be too small to be relevant for real-world implications. However, this hypothesis is to be confirmed by future studies in this context, preferably through field or natural experiments.

Besides, I test whether non-salience effects translate from consumption decisions into the labor market. Although participants underreact to non-salient sales taxes at the point of sale, I do not observe any influence on labor supply. This strikes out the possibility of labor supply increases and subsequent output increases through the non-salience of sales taxation. However, I suspect a second channel of salience effects beyond the scope of my experiment and may be examined in future research. While income taxes are typically paid in twelve equally sized tranches and are subject to screening in a tax declaration, sales taxes are paid in many more and much smaller tranches and are (usually) not subject to a review at the end of a year. The resulting increased visibility of and emphasis on the income tax may further develop differences in the reaction to both taxation types[9].

9. However, as outlined in section 2, the literature produces ambiguous results on salience effects in income taxation, see Abeler and Jäger (2015), Fochmann and Weimann (2013) and Hayashi, Nakamura, and Gamage (2013).

References

Abeler, Johannes, Armin Falk, Lorenz Goette, and David Huffman. 2011. "Reference points and effort provision." *American Economic Review* 101 (2): 470–92.

Abeler, Johannes, and Simon Jäger. 2015. "Complex tax incentives." *American Economic Journal: Economic Policy* 7 (3): 1–28.

Achar, Chethana, Jane So, Nidhi Agrawal, and Adam Duhachek. 2016. "What we feel and why we buy: the influence of emotions on consumer decision-making." *Current Opinion in Psychology* 10:166–170.

Basak, Suleyman, and Hongjun Yan. 2010. "Equilibrium asset prices and investor behaviour in the presence of money illusion." *The Review of Economic Studies* 77 (3): 914–936.

Blumkin, Tomer, Bradley J Ruffle, and Yosef Ganun. 2012. "Are income and consumption taxes ever really equivalent? Evidence from a real-effort experiment with real goods." *European Economic Review* 56 (6): 1200–1219.

Brunnermeier, Markus K, and Christian Julliard. 2008. "Money illusion and housing frenzies." *The Review of Financial Studies* 21 (1): 135–180.

Cabral, Marika, and Caroline Hoxby. 2012. *The hated property tax: salience, tax rates, and tax revolts.* Working Paper w18514. National Bureau of Economic Research.

Camerer, Colin, Linda Babcock, George Loewenstein, and Richard Thaler. 1997. "Labor supply of New York City cabdrivers: One day at a time." *The Quarterly Journal of Economics* 112 (2): 407–441.

Chetty, Raj, Adam Looney, and Kory Kroft. 2009. "Salience and taxation: Theory and evidence." *American economic review* 99 (4): 1145–77.

Cohen, Randolph B, Christopher Polk, and Tuomo Vuolteenaho. 2005. "Money illusion in the stock market: The Modigliani-Cohn hypothesis." *The Quarterly journal of economics* 120 (2): 639–668.

Engström, Per, Katarina Nordblom, Henry Ohlsson, and Annika Persson. 2015. "Tax compliance and loss aversion." *American Economic Journal: Economic Policy* 7 (4): 132–64.

European Commission. 2020. *Recovery plan for Europe.* https://ec.europa.eu/info/live-work-travel-eu/health/coronavirus-response/recovery-plan-europe_en. Accessed: 2020-11-12.

Fair, Ray C. 1971. "Labor force participation, wage rates, and money illusion." *The Review of Economics and Statistics,* 164–168.

Fehr, Ernst, and Jean-Robert Tyran. 2001. "Does money illusion matter?" *American Economic Review* 91 (5): 1239–1262.

———. 2007. "Money illusion and coordination failure." *Games and Economic Behavior* 58 (2): 246–268.

Finkelstein, Amy. 2009. "E-ztax: Tax salience and tax rates." *The Quarterly Journal of Economics* 124 (3): 969–1010.

Fisher, Irving. 1928. *The Money Illusion.* New York: Adelphi.

Fochmann, Martin, and Joachim Weimann. 2013. "The effects of tax salience and tax experience on individual work efforts in a framed field experiment." *FinanzArchiv/Public Finance Analysis,* 511–542.

Gustafson, Elizabeth, and Lawrence Hadley. 1989. "Labor supply and money illusion: a dynamic simultaneous equations model." *Quarterly Review of Economics and Business* 29 (4): 63–76.

Hadley, Lawrence H. 1982. "Some evidence on the role of money illusion in labor supply behavior." *Journal of Economics and Business* 34 (2): 153–157.

Hayashi, Andrew T, Brent K Nakamura, and David Gamage. 2013. "Experimental Evidence of Tax Salience and the Labor–Leisure Decision: Anchoring, Tax Aversion, or Complexity?" *Public finance review* 41 (2): 203–226.

Kahneman, Daniel, and Amos Tversky. 1979. "Prospect theory: an analysis of decision under risk." *Econometrica* 47 (2): 263–291.

Kessler, Judd B, and Michael I Norton. 2016. "Tax aversion in labor supply." *Journal of Economic Behavior & Organization* 124:15–28.

Kurokawa, Hirofumi, Tomoharu Mori, and Fumio Ohtake. 2020. "A Choice Experiment on Taxes: Are Income and Consumption Taxes Equivalent?" *Journal of Behavioral Economics and Finance* 13:53–70.

Mankiw, N Gregory, Matthew Weinzierl, and Danny Yagan. 2009. "Optimal taxation in theory and practice." *Journal of Economic Perspectives* 23 (4): 147–74.

OECD. 2020a. *Global Revenue Statistics Database.* https://stats.oecd.org/Index.aspx?datasetcode=RS_GBL. Accessed on: 2020-12-16.

———. 2020b. *OECD Economic Outlook No 107 - Double-hit scenario – (Edition 2020/1).* https://www.oecd-ilibrary.org/content/data/82024563-en. Accessed on: 2020-11-12.

Petersen, Luba, and Abel Winn. 2014. "Does money illusion matter? Comment." *American Economic Review* 104 (3): 1047–62.

Rees-Jones, Alex. 2014. "Loss aversion motivates tax sheltering: Evidence from US tax returns." *Available at SSRN 2330980.*

Sandmo, Agnar. 1976. "Optimal taxation: An introduction to the literature." *Journal of public economics* 6 (1-2): 37–54.

Schmeling, Maik, and Andreas Schrimpf. 2011. "Expected inflation, expected stock returns, and money illusion: What can we learn from survey expectations?" *European Economic Review* 55 (5): 702–719.

Shafir, Eldar, Peter Diamond, and Amos Tversky. 1997. “Money illusion.” *The Quarterly Journal of Economics* 112 (2): 341–374.

Wachter, Michael L. 1972. “A labor supply model for secondary workers.” *The Review of Economics and Statistics,* 141–151.

Xia, Lan, and Kent B Monroe. 2004. “Price partitioning on the internet.” *Journal of Interactive Marketing* 18 (4): 63–73.

Zizzo, Daniel John. 2010. “Experimenter demand effects in economic experiments.” *Experimental Economics* 13 (1): 75–98.

Appendix A: Experiment Screenshots

Figure 3: Virtual Supermarket Products, Prices and Ticket Returns

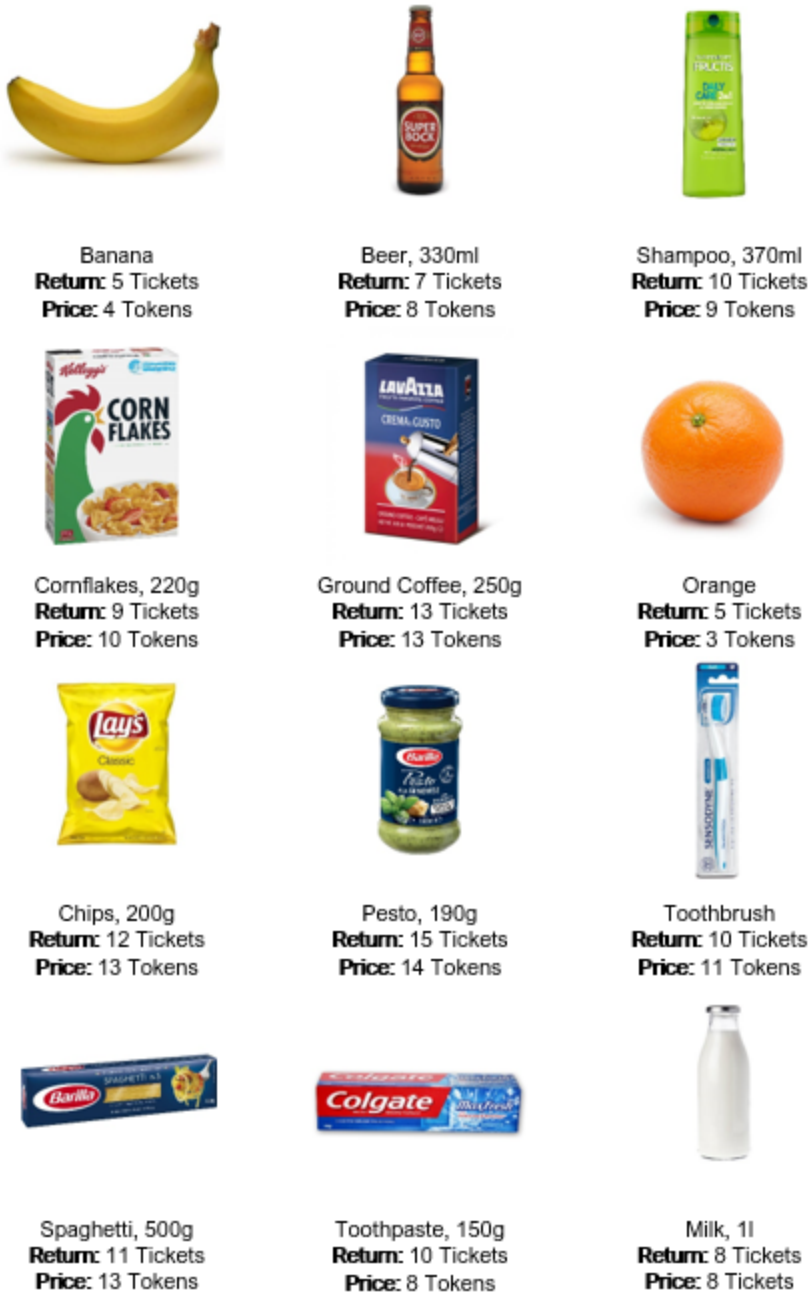


Figure 4: Rules for Lottery Ticket Payment

In addition to the tickets per product, there are four more rules that determine the number of tickets you receive in total:

- **Combining products** that are typically consumed together gives bonus tickets. More specifically, for your first combination you receive 10 extra tickets, for your second combination you receive 9 extra tickets, for your third combination you receive 8 extra tickets, and so on. From the 11th combination on, you receive 0 extra tickets for each combination.
  Example: If you have 4 combinations, you get 10 + 9 + 8 + 7 = 34 bonus tickets

- Buying the **same product again** reduces the tickets received from it by 1
  Example: If you buy 3 beers, you receive 7 + 6 + 5 = 18 tickets from those beers

- If you **exceed your budget** of tokens, tickets will be deducted from your account. **For each token that exceeds your budget, 5 tickets will be deducted from your account.**
  Example: If your budget is 100 tokens but you spend 103 tokens, your penalty is 3 * 5 = 15 tickets

- Tokens that are left over do not yield any tickets. **Only products generate lottery tickets** as return.

Figure 5: SST and MI Round 1 Description and Comprehension Questions

This is the first round out of two in this experiment.

In this round, you will receive a **wage of 9 tokens per correctly specified table**. You will have unlimited time and you can specify **as many tables as you want** up to a maximum of 50 tables. You will see one table per page.

When you entered the number of zeros into the field below the table and you want to go to the next table, click on “Next”. When you do not want to specify any more tables and want to proceed to the supermarket, please select the field “I want to end the task and proceed to the supermarket” before you click on “Next”.

Moreover, when you enter the supermarket, there will be a **sales tax of 20% on all products**. The tax will already be **included in the price** of each good.

Before you proceed to the task, please answer the following two questions:

How many tokens do you get per correctly specified table? [ ]

How much is the sales tax in %? [ ]

Figure 6: Round 1 Supermarket Snippet for SST and MI Group

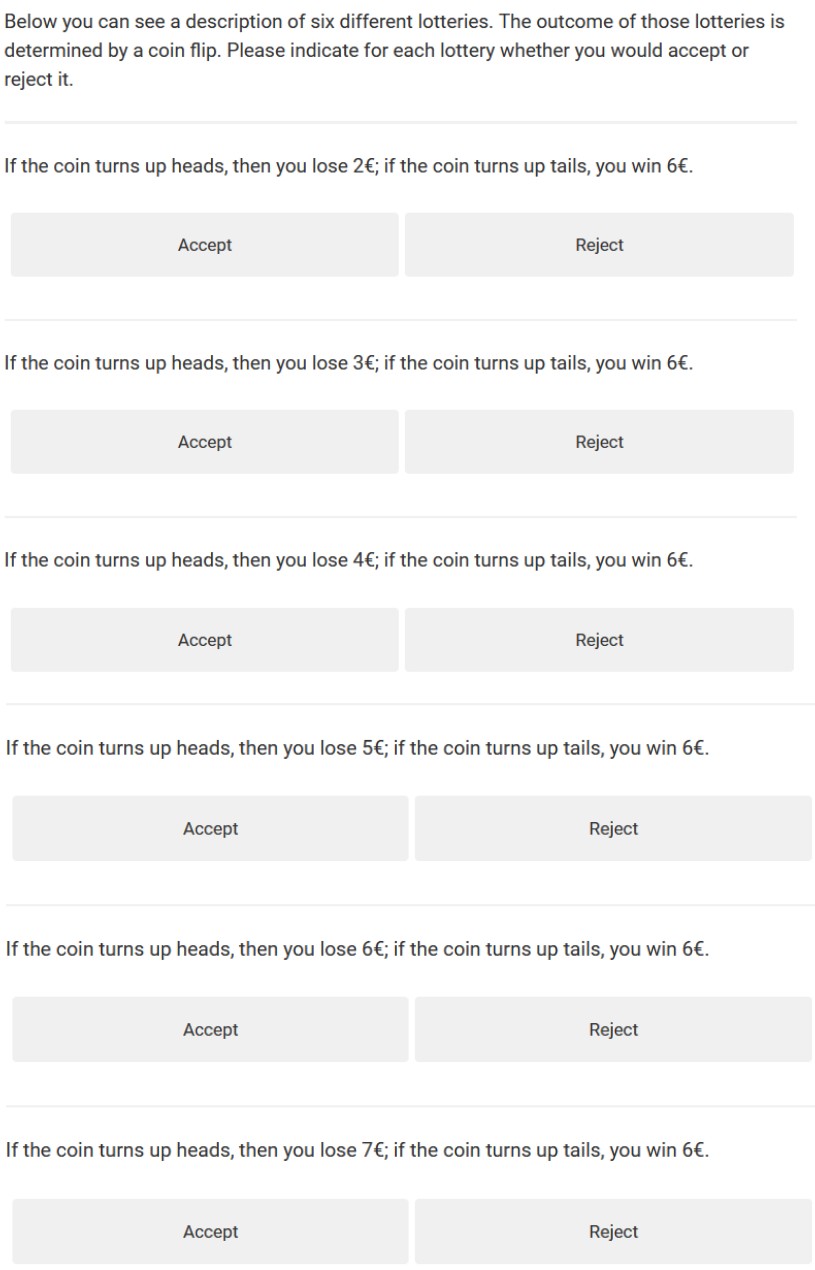
Figure 7: Money Illusion Table Round 2

Remember: New Prices after tax increase!

Banana:	7.2 tokens
Beer:	14.4 tokens
Shampoo:	16.2 tokens
Cornflakes:	18 tokens
Ground Coffee:	23.4 tokens
Orange:	5.4 tokens
Chips:	23.4 tokens
Pesto:	25.2 tokens
Toothbrush:	19.8 tokens
Spaghetti:	23.4 tokens
Toothpaste:	14.4 tokens
Milk:	14.4 tokens

1	1	1	1	1	1	0	1	0	1
0	0	0	1	0	0	1	0	0	1
1	1	1	0	1	1	1	1	1	1
1	1	1	1	0	1	1	0	1	0
1	1	1	1	1	1	1	1	0	1
1	1	1	0	1	1	1	1	0	1
1	1	1	1	1	1	1	1	1	1
0	1	0	1	1	0	1	1	1	0
0	1	1	1	1	1	1	1	1	1
0	1	1	1	0	1	1	1	1	1

Figure 8: Loss Aversion Questions



Appendix B: Robustness Checks

Table 4: Difference-in-Differences Estimation in log-Transformation

Sample	All		Increasers		Decreasers	
$\log(LaborSupply)$	(1.1)	(1.2)	(2.1)	(2.2)	(3.1)	(3.2)
Income	0.004	0.336	-0.116	0.068	0.195	-0.569
	(0.253)	(0.549)	(0.332)	(0.465)	(0.326)	(0.801)
Money Illusion	-0.090	-0.028	-0.053	-0.033	-0.020	-0.010
	(0.253)	(0.260)	(0.366)	(0.391)	(0.316)	(0.330)
Non-salient	-0.073	-0.075	-0.117	-0.130	-0.076	-0.087
	(0.267)	(0.270)	(0.361)	(0.378)	(0.342)	(0.342)
*Income*Loss Aversion*		-0.092		-0.074		0.220
		(0.128)		(0.097)		(0.193)
Loss Aversion		0.087		0.029		0.073
		(0.058)		(0.072)		(0.091)
Observations	414	392	162	152	186	178
R^2	0.034	0.044	0.081	0.095	0.242	0.260

*** p<0.01, ** p<0.05, * p<0.1

Note: Robust standard errors in parentheses. Estimates reflect logarithms of differences in differences across rounds between the displayed treatment groups and the baseline of the Salient Sales Tax group. Increasers (Decreasers) are participants who increased (decreased) their labor supply from round 1 to round 2.

Table 5: Difference-in-Differences Estimation excluding Outliers

Sample	All		Increasers		Decreasers	
Labor Supply	(1.1)	(1.2)	(2.1)	(2.2)	(3.1)	(3.2)
Income	0.03	-0.07	-0.07	-0.05	0.20	-0.57
	(0.24)	(0.54)	(0.32)	(0.44)	(0.33)	(0.80)
Money Illusion	-0.09	-0.04	-0.06	-0.06	-0.02	-0.01
	(0.24)	(0.25)	(0.33)	(0.35)	(0.32)	(0.33)
Non-salient	-0.06	-0.06	-0.07	-0.09	-0.08	-0.09
	(0.26)	(0.27)	(0.35)	(0.36)	(0.34)	(0.34)
*Income*Loss Aversion*		0.03		-0.03		0.22
		(0.12)		(0.10)		(0.19)
Loss Aversion		0.04		-0.01		0.07
		(0.06)		(0.07)		(0.09)
Observations	382	360	158	148	186	178
R^2	0.05	0.05	0.08	0.09	0.24	0.26

*** p<0.01, ** p<0.05, * p<0.1

Note: Robust standard errors in parentheses. Estimates reflect differences in differences across rounds between the displayed treatment groups and the baseline of the Salient Sales Tax group. Outliers are defined as participants who either worked on 50 tables in both rounds or on zero tables in the first round. Increasers (Decreasers) are participants who increased (decreased) their labor supply from round 1 to round 2.

Table 6: Difference-in-Differences Estimation in log-Transformation excluding Outliers

Sample	All		Increasers		Decreasers	
$\log(LaborSupply)$	(1.1)	(1.2)	(2.1)	(2.2)	(3.1)	(3.2)
Income	0.026	-0.068	-0.068	-0.045	0.195	-0.569
	(0.237)	(0.544)	(0.317)	(0.443)	(0.326)	(0.801)
Money Illusion	-0.091	-0.039	-0.057	-0.060	-0.020	-0.010
	(0.245)	(0.250)	(0.333)	(0.351)	(0.316)	(0.330)
Non-salient	-0.064	-0.057	-0.069	-0.092	-0.076	-0.087
	(0.260)	(0.265)	(0.348)	(0.362)	(0.342)	(0.342)
*Income*Loss Aversion*		0.032		-0.031		0.220
		(0.123)		(0.097)		(0.193)
Loss Aversion		0.041		-0.015		0.073
		(0.059)		(0.071)		(0.091)
Observations	382	360	158	148	186	178
R^2	0.048	0.053	0.077	0.088	0.242	0.260

*** p<0.01, ** p<0.05, * p<0.1

Note: Robust standard errors in parentheses. Estimates reflect logarithms of differences in differences across rounds between the displayed treatment groups and the baseline of the Salient Sales Tax group. Outliers are defined as participants who either worked on 50 tables in both rounds or on zero tables in the first round. Increasers (Decreasers) are participants who increased (decreased) their labor supply from round 1 to round 2.