Are income and consumption taxes ever really equivalent? Evidence from a real-effort experiment with real goods

Tomer Blumkin¹, Bradley J. Ruffle*, Yosef Ganun²

Department of Economics, Ben-Gurion University, Beer Sheva 84105, Israel

Abstract

The public finance literature demonstrates the equivalence between consumption and labor-income (wage) taxes. We introduce an experimental paradigm in which individuals make real labor-leisure choices and spend their earned income on real goods. We use this paradigm to test whether a labor-income tax and an equivalent consumption tax lead to identical labor-leisure allocations. Despite controlling for subjects’ work ability and inherent labor-leisure preferences and disallowing saving, subjects reduce their labor supply significantly more in response to an income tax than to an equivalent consumption tax. We discuss the economic implications of a policy shift to a consumption tax.

© 2012 Elsevier B.V. All rights reserved.

1. Introduction

1.1. Background

In this paper, we offer an experimental test of the equivalence between consumption and wage taxes. We design an experimental environment in which, confronted with a real-effort task, subjects decide how to allocate their time between labor and leisure. We then compare labor-leisure allocations between two tax treatments, one with a consumption tax and another with an equivalent wage tax. Our experimental design allows us to control for possible sources of variation across subjects, notably, their ability at the real-effort task and their pre-tax labor-leisure preferences.

Since Hobbes (1651) and continuing with Mill (1871), the question of whether to tax consumption or income arises repeatedly in tax policy debate. Over the past three decades, most OECD countries – the US being a notable exception – have shifted their tax mix from income to consumption taxation, mostly in the form of VAT (see OECD, 2008).³

* Corresponding author. Tel.: +972 8 6472308.
E-mail addresses: tomerblu@bgu.ac.il (T. Blumkin), bradley@bgu.ac.il (B.J. Ruffle), yosig@bll.co.il (Y. Ganun).

¹ Tel.: +972 8 6472298.
² Tel.: +972 52 8619100.

³ The US also explored such a shift. In January 2005, the US President George W. Bush commissioned a panel to propose a comprehensive reform to the tax code (President’s Advisory Panel on Federal Tax Reform, 2005). The panel considered replacing the entire income-tax system with a national sales tax. However, the complexity of replacing the current tax regime with a broad-based consumption tax inhibited its adoption.

http://dx.doi.org/10.1016/j.euroecorev.2012.06.001
The classic Haig–Simons definition (according to which income equals consumption plus changes in wealth (savings))
highlights that the major distinction between an income tax (on all income sources) and a consumption tax is the former’s
taxation of savings.\footnote{See Haig (1921) and Simons (1938).} The taxation of savings distorts the taxpayer’s inter-temporal consumption allocation, an often-invoked argument in favor of shifting to a consumption tax. In a second-best setting, redistributive goals may justify the
taxation of savings despite its distortive effects.

The public finance literature has long been occupied with the question of the optimal mixture between income and
consumption taxes for redistributive purposes. In a seminal study, Atkinson and Stiglitz (1976) (AS) extend Mirrlees’
(1971) pioneering study on optimal income taxation by allowing for several consumption goods. AS demonstrate that,
with certain restrictions on the underlying preferences, taxing consumption becomes redundant and the social optimum is
attainable by levying a labor-income (wage) tax only. An implication of the AS result is the undesirability of taxing savings
(interest income). AS and a large body of the subsequent literature\footnote{See, for example, Deaton (1979), Saez (2002) and Kaplow (2006).} are premised on the equivalence between a linear
wage tax and a comprehensive uniform consumption tax (levied on the set of all consumption goods).\footnote{See Haig (1921) and Simons (1938).} In fact, wage taxes can be thought as pre-paid consumption taxes, while consumption taxes like the VAT and retail sales taxes are viewed as
post-paid consumption taxes. With a dearth of behavioral evidence on the topic, the literature regards these two taxes as
equivalent.

1.2. Our contribution

To test the equivalence of consumption and wage taxes, we introduce a new experimental paradigm in which both
work and leisure choices earn subjects payments in real goods. This paradigm consists of an incentivized, two-stage
individual choice problem that requires subjects to allocate their time between leisure (rewarded with a payment for each
unit consumed) and a real-effort task. Their performance at the real-effort task in the first stage earns them income, which
they allocate in the second stage between two consumption goods. After gaining experience in this environment, we
introduce a tax. In the labor-income tax (IT) treatment, a 50% flat wage tax is imposed on earned income. In the equivalent
consumption tax (CT) treatment, a 100% ad-valorem tax is levied on both consumption goods. No savings are permitted in
either treatment. Notice that both tax regimes entail a 50% erosion in the individual’s purchasing power without changing
the relative prices of the consumption goods. That is, in both tax treatments subjects face the same budget constraint.
Controlling for differences in labor-market productivity (i.e., differences in performance of the real-effort task) and
inherent labor-leisure preferences using pre-tax treatments, we test whether these two equivalent tax regimes in fact lead
to identical labor-leisure choices.

Although we have designed these two tax treatments to be theoretically equivalent, we propose a behavioral
hypothesis that suggests individuals will work more and consume less leisure in the CT treatment than in the equivalent
IT treatment. Our hypothesis is based on money illusion, that is, individuals’ observed tendency to think in nominal rather
than in real terms. An individual suffering from money illusion will typically display a reluctance to accept a nominal wage
increase. As a result, the impact of an equivalent (measured in real purchasing power terms) wage tax is apparent at the time individuals make
their labor-supply decisions. In our experimental framework, we expect that this differential perception of the taxes will
translate into individuals choosing to work less under a wage tax than under an equivalent consumption tax. Put

\footnote{In a neo-classical framework, any two tax schedules that yield the same choice set and budget constraint should have no impact on a rational
individual's choice (nor on government fiscal considerations) and hence should be equivalent for tax design purposes. The public finance literature
demonstrates the equivalency of several other pairs of tax instruments that are \textit{prima facie} different. Notable examples include social security taxes levied
on employees and employers and commodity taxes imposed on producers and consumers.}

\footnote{See Deaton (1979), Saez (2002) and Kaplow (2006).}

\footnote{See Shafir et al. (1997) for early survey evidence on money illusion as well as Fehr and Tyran (2001, 2007) for experimental evidence and
Brunnermeier and Julliard (2008), Cohen et al. (2005) and Koorenman et al. (2004) for evidence based on naturally occurring data.}
equivalence between the economic and statutory incidence of a unit commodity tax (also known as liability side tax equivalence). These papers show that for sufficiently competitive markets the economic incidence of a unit tax is independent of the side of the market that bears the statutory incidence.

**Riedl and van Winden (2011)** compare experimentally the economic performance of a small open economy subject to a wage tax (WT) with one subject to a sales-tax-cum-labor-subsidy (STLS). In the WT treatment, producers pay a wage tax on each unit of labor hired. In the STLS treatment, producers incur a sales tax on products sold, but also receive a subsidy for each unit of labor hired. According to most economic indicators, the STLS outperforms the WT. To explain their results, the authors propose that the upfront burden of a WT and uncertainty about product prices render producers reluctant to hire labor.

Our paper contributes to a recent fast-growing strand in the public finance literature on the misperception of taxes. **Sausgruber and Tyran (2004)** demonstrate that buyers systematically underestimate the tax burden of a tax levied on sellers and the consequences of this misperception for preferences redistribution. **Liebman and Zeckhauser (unpublished manuscript)** show that the optimal income-tax formulae change markedly when individuals mistake their average for their marginal tax rate. **Feldman and Katuscak (unpublished manuscript)** find that households respond to a predictable, lump-sum reduction in their tax schedules by reducing their labor supply, thereby suggesting a change in the average tax rate is mistaken for a change in the marginal tax rate. In a series of tax and redistribution decision-making scenarios, **McCaffery and Baron (2006)** elicit attitudes toward various fiscal policies and find that subjects prefer hidden to transparent taxes and ignore the longer-term effects of tax policies. Based on a grocery store field experiment and empirical state-level data on alcohol sales, **Chetty et al. (2009)** show that posting sales-tax-inclusive prices renders the tax more salient and thus reduces consumer demand relative to adding the sales tax at the cash register. **Feldman and Ruffle (unpublished manuscript)** confirm this finding in a controlled laboratory setting. But an additional tax-deduction treatment and purchases at the product-price level cast doubt on salience as the explanation. **Finkelson (2009)** demonstrates that toll rates increase in response to a switch from manual collection (where drivers pay in cash) to a less salient electronic collection system (where drivers are automatically debited).

The organization of the paper is as follows. In the next section, we present a simple theoretical framework to illustrate the mechanism underlying our main experimental finding and the potential welfare gain from a shift to a consumption tax. In section 3, we detail the experimental design and procedures. Section 4 presents the results. Section 5 concludes.

### 2. An illustrative model of agents’ misperception

In this section, we present a simple behavioral model to highlight how tax misperception affects individuals’ labor-leisure decisions. In particular, this model illustrates our main behavioral hypothesis that, due to tax misperception, individuals choose a higher labor supply under a consumption tax than under an equivalent labor-income tax. The model also demonstrates the potential welfare gain associated with a shift from a labor-income to a consumption-tax regime. The next section presents our experiment designed to test our behavioral hypothesis. The experiment does not directly test the welfare implications predicted by our model.

Consider a standard labor supply model with a representative individual whose utility is given by $U(c, l) = c - h(l)$, where $c$ denotes consumption, $l$ denotes labor and $h$ is assumed to be strictly increasing and strictly convex.$^9$

The production function employs labor only and exhibits constant returns to scale. We denote by $w > 0$ the individual’s hourly productivity (hence the competitive wage rate). We normalize the price of the consumption good to unity, without loss of generality. The individual is faced with the following (perceived) budget constraint:

$$
\ell \cdot w \cdot (1 - t) + \tau = c \cdot (1 + \alpha \cdot s),
$$

(1)

where $t < 1$ is the (flat) tax rate on labor income (wages), $\tau$ denotes a lump-sum transfer (a tax if negative) and $s$ is the tax rate on consumption. The parameter $0 \leq \alpha \leq 1$ measures the individual’s degree of misperception, possibly a result of money illusion (see our discussion in the Introduction). When $\alpha = 1$ the individual is fully rational and perceives the consumption tax and the corresponding budget constraint correctly, in accordance with neo-classical consumer theory. When $\alpha < 1$ the individual underestimates the burden associated with a consumption tax.$^9$ Based on his perceived budget constraint in (1), the individual determines his labor supply and earns the commensurate income. Finally, he spends this income on the consumption good, subject (regrettably) to the true budget constraint (as if $\alpha = 1$ in (1)).

---

$^8$ Salience may be the source of the money illusion we hypothesize in our setup. Wage taxes are tantamount to a direct wage cut; whereas price increases due to an equivalent consumption tax are more subtle and may go undetected. In the next section, we invoke the more general notion of misperception to model individuals’ under-reaction to the consumption tax.

$^9$ Our choice of a quasi-linear preference specification in the theoretical model (see Diamond (1998) and Salanie (2003) for applications in the optimal income tax literature), which rules out income effects, is made for tractability only to illustrate in sharpest relief our key argument on the role of tax misperception and the resulting potential welfare gain associated with a shift from an income to a consumption tax. In Appendix B, we examine a more general functional form of the utility function, which allows for an income effect by assuming diminishing marginal utility from income. We show that the welfare dominance of a consumption tax remains robust to the admittance of an income effect.

$^{10}$ Chetty et al. (2009) employ a similar formulation to study the role of tax salience in consumer purchasing decisions. In our model, tax misperception is exogenously given and is measured along a continuum, whereas Chetty et al. treat individuals’ misperception as a costly binary choice between full attention and complete inattention to the sales tax added at the cash register to the posted price.
We now demonstrate that when the individual underestimates the consumption tax burden, levying a consumption tax entails a smaller excess burden than that associated with an equivalent wage tax (while the individual is ex-post indifferent between the two tax regimes). Formally, we prove the following:

**Proposition 1.** When \( \alpha < 1 \), for any wage tax there exists a consumption tax that generates strictly higher tax revenues while leaving the individual's utility unchanged.

**Proof.** See Appendix A.

The misperception of the tax burden (reflected by a mitigated substitution effect between leisure and consumption) induces the agent to work more. In this respect, the government can raise more revenue through a consumption tax than a theoretically equivalent income tax. However, because the individual does not correctly perceive the true budget constraint (he overestimates his purchasing power), the utility derived under the tax regimes differs. In particular, the consumer does not end up at his most preferred consumption bundle given the true consumption-tax rate. In order to calibrate the two regimes to ensure that both yield the same level of utility, we reduce the consumption-tax rate below the level of equivalence. We then prove that due to the reduced substitution effect and corresponding higher labor supply, total tax revenues are higher under a consumption tax.

3. **Experimental design and procedures**

To test the equivalency between labor-income and consumption taxes, we design two tax systems that yield identical after-tax budget lines, thereby creating identical labor-leisure incentives if the taxes are correctly perceived. In a between-subject design, we determine whether subjects indeed make the same labor-leisure allocation as predicted by the theory or whether, according to our alternative hypothesis, subjects choose to work more in the consumption-tax condition. This comparison of labor-leisure choices in equivalent labor-income-tax and consumption-tax treatments is the third of our three-part experiment. The first two parts of the experiment are designed to measure and control for subjects' work ability and inherent labor-leisure preferences, respectively. We detail in turn each of these three parts below. The experimental instructions for all three parts appear in Appendix C.

3.1. **Three parts of the experiment**

3.1.1. **Part One (work ability)**

The common element to all three parts of the experiment is the real-effort work task: each subject solves by hand two-digit by two-digit multiplication questions. Part One serves to measure each subject's innate ability or productivity at this task. Our objective is to create a labor-income tax and a theoretically equivalent consumption-tax treatment balanced in terms of subjects' work abilities.

To measure each subject's work ability, subjects are asked to solve as many multiplication problems as they can in 3 min. Although individuals are generally internally motivated to prove to themselves (or to the experimenter) their ability to perform well in this task, we provided an additional monetary incentive of 0.5 shekel for each correctly answered question.\(^{11}\) Throughout this and the other two parts, the subject may observe both his numbers of correctly and incorrectly answered multiplication questions and his cumulative earnings. (See the screenshot for Part One in Appendix C.)

At the completion of Part One, while subjects proceeded to Part Two, the software ranked subjects according to the number of correctly solved multiplication questions. We applied the rank-sorting algorithm displayed in Table 1 to assign each subject to either the labor-income tax or consumption-tax treatment in Part Three. At the beginning of Part Three, each subject receives the instructions only for the treatment to which he has been assigned. Subjects are not made aware of the ranking algorithm, their overall ranking or the existence of the other tax treatment in which they do not participate.

The algorithm balances the two tax treatments in terms of the rankings of subjects' work abilities. Table 1 indicates how this balance is achieved: the subjects with the highest and fourth highest abilities are assigned to the consumption-tax treatment (CT), while the second and third highest ranking subjects are assigned to the labor-income tax treatment (IT). This "snake" pattern of subject assignment continues until all subjects are exhausted. The result is that if the number of subjects in a session is a multiple of four, the average ability ranking of the two treatments is identical; otherwise, the average ranking differs by a mere fraction for sessions with at least 17 subjects (applicable to all four of our sessions).

3.1.2. **Part Two (labor-leisure preferences)**

The second part of the experiment records subjects' (pre-tax) labor-leisure choices. This second part consists of a two-stage, full-information, individual-choice problem. In the first labor-leisure-allocation stage, the individual decides how much of the available 3 min to devote to work in the form of solving multiplication problems. For each correctly answered question, the subject earns two points that may be exchanged for either of the two consumption goods in the second stage.

\(^{11}\) One new Israeli shekel (NIS) equals about $0.3 USD or €0.2. To control for question difficulty across subjects, all subjects saw the same series of randomly chosen multiplication questions in the same order. To reduce the variance in question difficulty across questions, we excluded integers ending in "0" or "1".
The subject may stop working at any time during the 3-min round by pressing the “Stop” button. For each 15 s that the individual chooses not to work (leisure), he earns one unit of the leisure good (a voucher for a bottled soft drink). In the second, consumption stage of the round, the individual decides how to allocate the points earned from the labor task between the two consumption goods (vouchers redeemable for falafel sandwiches or for pizza slices). In this pre-tax treatment, each point earned can be exchanged for a half falafel sandwich or one pizza slice. Compared to Part One, this second part complicates the subject’s decision task in two respects: the subject must first decide how to allocate his 3-min endowment between labor and leisure, and he must subsequently decide how to allocate his earned income between the two consumption goods. Due to these additional complexities, we want to allow the subject to repeat Part Two. At the same time, we recognize that solving multiplication questions is mentally fatiguing and that this part’s main purpose is really to provide a control for subjects’ preferences. We resolved these tradeoffs by having each subject play two rounds of Part Two.

### Part Three (Income tax or Consumption tax)

Part Three of the experiment is identical to Part Two, except that the subject now faces either a labor-income tax or a consumption tax, depending on the treatment to which he is assigned (according to the sorting algorithm described in Part One). In the labor-income tax treatment (IT), a 50% wage tax is imposed on subjects’ earnings, meaning that for each correctly answered multiplication question the subject earns one point (instead of two). The prices of half falafel sandwiches and pizza slices remain the same as in the pre-tax treatment, namely one point each.

Alternatively, in the consumption-tax treatment (CT), a 100% consumption tax is imposed on each of the two consumption goods, meaning that a half falafel sandwich and a slice of pizza now each cost two points (instead of one point each). As in the pre-tax treatment, the value of a correctly answered question is two points.

The key feature of these two tax treatments is that the subjects in each treatment face the identical post-tax budget line: a half falafel sandwich and one pizza slice each cost one correctly answered multiplication question, while each soft drink costs 15 s not spent on solving multiplication questions. For consistency between treatments and to eliminate the tax-framing effect observed in Chetty et al. (2009), we presented subjects with tax-inclusive prices and income in each of the respective tax treatments. See the screenshots in Part three of Appendix C.

The decision complexity of Part Three suggests that allowing subjects to repeat it will provide us with more informed estimates of their preferences. Moreover, to the extent that subjects misperceive the consumption tax, feedback and repetition will allow them to correct their mistakes in subsequent rounds. For example, subjects might not think ahead to the consumption stage when making their labor-leisure choices in the first stage of round 1. However, after confronting the tax-inclusive prices in stage two of round 1, subjects can adjust their labor-leisure choices accordingly in round 2. At the same time, mental fatigue and the intensity associated with solving multiplication problems under time pressure severely limits the number of feasible repetitions. Given the centrality of this part of the experiment, we chose to have each subject play three rounds.

---

**Table 1**

<table>
<thead>
<tr>
<th>Treatment</th>
</tr>
</thead>
<tbody>
<tr>
<td>CT</td>
</tr>
<tr>
<td>1</td>
</tr>
<tr>
<td>4</td>
</tr>
<tr>
<td>5</td>
</tr>
<tr>
<td>8</td>
</tr>
<tr>
<td>9</td>
</tr>
<tr>
<td>12</td>
</tr>
<tr>
<td>13</td>
</tr>
<tr>
<td>16</td>
</tr>
</tbody>
</table>

To balance the two tax treatments in terms of labor-market productivity, subjects in a session are assigned to either the consumption-tax treatment (CT) or income-tax treatment (IT) according to the displayed ability-ranking algorithm.

---

12 Because subjects are confined to the lab for the duration of the experiment, the payment of leisure good ensures that they indeed derive utility from time not spent working.

13 Pizza and falafel are the two most popular fast foods in Israel. Pizza is sold whole and by the slice; falafel sandwiches are served in a half or full pita bread.
3.2. Subjects and payments

To ensure that our results are not attributable to subjects’ misunderstanding of the rules of the experiment or the implications of the tax, we administered a short quiz to subjects prior to beginning Parts Two and Three to verify their knowledge of the prices of all goods. More important for understanding the experiment and meaning of the tax, we limited participation to economics students. Eighty undergraduate majors in economics participated in one of the four sessions at Ben-Gurion University. The entire experiment took about 1 h and 45 min. The average cash payment from Part One was 3.0 NIS, while the average payments in vouchers based on one randomly chosen round from both Parts Two and Three were 7.7 bottled drinks, 6.3 half falafel sandwiches and 7.0 pizza slices. The vouchers were valid for redemption for up to 1 year. At the time the experiments were conducted, a bottled drink cost 6 NIS, while a half falafel sandwich and a pizza slice each cost 5 NIS. Thus, the average subject earned 115.7 NIS or about $34 USD or €24, considerably more than the student wage of 19 NIS an hour at the time the experiments were conducted.

3.3. Logic underlying design choices

Two distinctive features of our setup are the existence of a payment for not performing the work task and the form of payment for work and leisure—in kind rather than in cash. Both choices follow from our concern that individuals will reach a corner solution: all leisure or, more likely, all work. A prevalence of corner solutions inhibits our ability to measure individuals’ responsiveness to the addition of a tax.

The consumption of leisure is the benefit an individual receives from not working. Without an explicit payment for not working, individuals confined to the lab with no real opportunity cost of time may choose to work the entire time (rather than do ‘nothing’, which yields a plausibly limited benefit). To avoid this undesirable corner solution, we augment the benefit to not working (leisure) by paying for each unit of leisure consumed (one bottle of soft drink for each 15 s). In our context, this payment can be thought as a subsidy to leisure consumption. The payment for leisure is not aimed to substitute literally for real-world leisure, but rather serves to enhance its otherwise limited value in the laboratory.

In choosing the form of payment, we steered away from cash or cash equivalents such as cell-phone credit for fear that such payments may entail no satiation. This may encourage individuals to maximize their total cash payment, thereby leading to a corner solution.

We include two consumption goods along with the leisure good to capture in the simplest way the two typical tradeoffs faced by individuals: time allocation between labor and leisure and earned-income allocation between commodities. We selected in-kind goods that are similarly attractive to one another for a wide range of students. In choosing three similarly well liked goods we aim, again, to minimize the number of subjects opting for a corner solution and to create sufficient variation across subjects’ labor-leisure allocation decisions to provide a persuasive test of the equivalence of the two tax regimes. Months prior to the experiment, we conducted a questionnaire to determine the set of goods to include in our design. The questionnaire elicits subjects’ preferences over different bundles of goods. The results revealed that falafel and pizza are equally well liked substitutes and neither is chosen to the exclusion of the leisure good, soft drinks, which serves as a complement.

Payment according to subjects’ cumulative earnings across all rounds would invite satiation, which would lead to different labor-leisure-consumption choices across rounds. This between-round variation is undesirable since it is an artifact of the payment calculation. Consequently, each subject was paid based on his results from one randomly chosen round from each of the second and third parts of the experiment. This payment method serves to avoid scenarios in which subjects concentrate their labor supply in one round and opt for leisure in the remaining round(s) of the treatment.

Our experiment consists of three parts with the first two parts serving as controls. The first part allows us to balance the work-ability composition of the IT and CT treatments and to explain a subject’s labor-leisure choice as a function of his observed labor productivity. The second part provides a within-subject control for inherent labor-leisure preferences. The time difference the subject devotes to the work task between the tax and the pre-tax treatments will serve as one of our dependent measures.

14 The subject answered the same set of questions in either IT or CT as in the previous no-tax treatment. The fact that the correct answers change from the no-tax to the tax treatment highlights for the subject the effect of the tax on prices.

15 The reader may wonder why participants do not simply choose all labor or all leisure according to the decision that maximizes the total monetary value of the vouchers for goods, subject to their labor market ability and the real wage rate. After the experiment, they could then sell or trade excess vouchers and purchase at market price the goods they did not consume during the experiment. The answer is that these vouchers (familiar only to the subjects in these experiments) are likely to trade at a substantial market discount. In addition, the number of bottled drinks, falafel sandwiches or pizza slices potentially earned from either complete labor-market or leisure specialization in this experiment would greatly exceed many subjects’ optimal quantity of this good in a year. Even students value a diversified diet. The paucity of observed corner solutions in our experiments validates this reasoning.

16 The questionnaire asked each subject to allocate a hypothetical income among three goods in each of the distinct bundles of goods. We conducted this questionnaire on 69 students at locations (Tel Aviv University and Sapir College) different from the site of our experiments (Ben-Gurion University) to avoid any subjects from participating in both the questionnaire and subsequent experiment. The detailed findings from the questionnaire are available upon request.
There is also an important between-subject aspect of the design that merits mention. Exposing all subjects to both tax treatments one after the other risks rendering transparent their equivalency, thereby unintentionally suggesting to subjects that they are expected to make identical labor-leisure choices. Moreover, order effects are non-existent in a between-subject design and session effects are rendered irrelevant since subjects in each session are equally divided between the two tax treatments.

A number of design choices were made to enhance the external validity of our results. The instructions to participants, for instance, explain the tax on goods and the tax on income precisely as such. In fact, the language of the instructions was purposely chosen to reflect the labor-leisure decisions and the taxation of goods and income outside of the laboratory. To frame our experiment in more neutral terms would subject our results to the critique that they may not be robust to a more realistic setting. Also, the subject repeated the same labor-leisure choice to allow for learning from the environment and past decisions. Finally, we recruited economics majors only. If economics students underestimate the impact of a reduction in real income through higher prices of consumption goods, as hypothesized, resulting in differential labor supply across the two tax treatments, then a less savvy subject pool may be even more susceptible to tax misperception.

4. Results

Descriptive statistics for the two tax treatments (IT and CT) appear in Table 2. The first row confirms the effectiveness of the ability-sorting algorithm in balancing the two tax treatments in terms of subjects’ abilities. The average ability in IT is 5.87 questions compared to 5.60 questions in CT. A t-test of means (p = 0.64) and the non-parametric Wilcoxon–Mann–Whitney test of distributions (p = 0.67) both indicate that abilities are similarly distributed in the two tax treatments.

The next three rows display the overall average time (in seconds) devoted to labor supply in the pre-tax treatment (before subjects knew of their assignment to, or even the existence of, the tax treatment), in the tax treatment, as well as the change between these two treatments, respectively. These numbers reveal that on average the substitution effect dominates any possible positive income effect: subjects respond to the imposition of a tax by significantly reducing their labor supply in both treatments (p < 0.01 for both IT and CT).

Of greater interest, subjects reduce their labor supply by 1/3 or 44.5 s on average in IT compared to the no-tax treatment, while in CT the decrease is only about 15% of the no-tax amount or a 20.3 s decline on average. The difference in these differences of each subject’s average response to the two taxes is 24.2 s (s.d. = 10.5) and is highly significant (p = 0.02). In words, subjects’ labor supply is significantly more responsive to the income tax than the consumption tax, as hypothesized. The implied arc elasticities of labor supply with respect to the real net wage are 0.56 in IT and 0.24 in CT.17

The differential reaction to the two taxes can also be seen in the distribution of subjects’ average labor supply responses displayed in Fig. 1. To begin, the figure reveals that the vast majority of subjects reduce their labor supply in response to

---

Table 2
Descriptive statistics by tax treatment.

<table>
<thead>
<tr>
<th>Treatment</th>
<th>IT</th>
<th>CT</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ability</td>
<td>5.87 (2.60)</td>
<td>5.60 (2.56)</td>
</tr>
<tr>
<td>Pre-tax labor supply</td>
<td>141.6 (27.6)</td>
<td>130.0 (36.2)</td>
</tr>
<tr>
<td>After-tax labor supply</td>
<td>98.7 (57.3)</td>
<td>109.1 (49.4)</td>
</tr>
<tr>
<td>Δ labor supply</td>
<td>-44.5 (49.7)</td>
<td>-20.3 (44.1)</td>
</tr>
<tr>
<td>Δ labor supply in IT - Δ labor supply in CT</td>
<td>-24.2 (10.5)</td>
<td></td>
</tr>
<tr>
<td>Pre-tax units of leisure</td>
<td>2.24 (1.79)</td>
<td>2.95 (2.32)</td>
</tr>
<tr>
<td>After-tax units of leisure</td>
<td>4.99 (3.67)</td>
<td>4.36 (3.26)</td>
</tr>
<tr>
<td>Δ units of leisure</td>
<td>2.75 (40)</td>
<td>1.41 (40)</td>
</tr>
</tbody>
</table>

Averages by tax and pre-tax treatment (standard deviations below in parentheses). Labor supply is measured in seconds, units of leisure in numbers of bottled drinks.

---

17 Since the functional form of the labor supply function is unknown and the percentage change in the real net wage is so large (50%), we compute arc elasticities. The arc elasticity calculates the percentage change relative to the midpoints of the change in labor supply and change in real net wage.
imposition of a tax, suggesting the dominance of the substitution effect. In line with our conjecture, the magnitude of the reduction of labor supply is larger in response to IT than to CT.

The figure also reveals that about a quarter of the subjects curiously increase their time devoted to the work task following the introduction of the tax. For these subjects, the income effect seems to dominate the substitution effect, assuming, plausibly, that leisure is a normal good. However, misperception of the consumption tax suggests that the increase in labor supply should be smaller under a consumption tax than under an income tax—precisely the opposite of what is observed in the figure.

Table 3 compares the group of subjects who lowered their post-tax labor supply with their counterparts who raised it. The table shows no significant differences between the two groups in terms of gender composition, labor-market ability or labor-leisure preferences as reflected in their pre-tax labor supply.

Returning to our main result, regression (1) of Table 4 reports a highly significant difference-in-difference estimate, with the subject's own labor-market ability (ability), measured as the number of multiplication questions subject i correctly answered in Part One, as a control. The coefficient of −23.9 on the treatment indicator IT reveals that subjects reduce their labor supply by an additional 24 s on average in response to the income tax compared to the equivalent consumption tax. The coefficient on ability is small and not significantly different from zero.

Let us now make use of the entire panel dataset for all 80 subjects, each of whom participated in two no-tax rounds and three tax rounds facing either an income tax (IT) or a consumption tax (CT). Regressions (2)–(5) in Table 4 report the regression results with subject i’s labor supply in round t in seconds as the dependent variable. The random-effects GLS regression results in (2) reveal that subjects reduce their labor supply by 39 s on average when exposed to the income tax, while those who face the consumption tax work 21 s less compared to their pre-tax labor supply. Both the IT and CT coefficients are significantly different from 0 (p < 0.01 in both cases), implying that subjects reduce their labor supply in both tax treatments relative to the no-tax regime. Furthermore, these coefficients are significantly different from each other (p < 0.01), namely labor supply falls more under the IT than the CT regime.

Learning is a common phenomenon in individual-choice experiments. With successive rounds a subject may become more adept at solving multiplication questions or more fatigued. To determine whether subjects’ labor supply decisions display a time trend, regression (2) also includes dummies for rounds 2, 4 and 5, none of which is significantly different from zero. The implication is that the only significant reduction in subjects’ labor supply occurs in round 3 when the income or consumption tax is first introduced. Prior to the tax, subjects spend as much time solving multiplication problems in round 2 as in round 1. Also, subsequent to the tax’s introduction (rounds 4 and 5), subjects’ labor supply stabilizes.

18 Another explanation for this increase in labor supply is that subjects aspire to a target income in order to purchase a specific number of falafel sandwiches or pizza slices. Camerer et al. (1997) present evidence that New York City cab drivers set a daily earnings target, leading them to work longer hours on slow days and to quit early on busy days. See Farber (2005) and Fehr and G¨otte (2007) for critiques of this finding.

19 The significance and non-significance of all the estimates in all of our regressions remain unchanged if as the dependent variable we instead use the coarser measure of subject i’s number of units of leisure consumed in round t (i.e., the number of 15 s blocks the subject chose not to work).
Table 3
Descriptive statistics by response to tax.

<table>
<thead>
<tr>
<th>Variable/grouping</th>
<th>Lower after-tax labor supply</th>
<th>Higher after-tax labor supply</th>
</tr>
</thead>
<tbody>
<tr>
<td>Male</td>
<td>46 (78.0%)</td>
<td>14 (66.7%)</td>
</tr>
<tr>
<td>Ability</td>
<td>5.69</td>
<td>5.86</td>
</tr>
<tr>
<td></td>
<td>(2.63)</td>
<td>(2.46)</td>
</tr>
<tr>
<td>Pre-tax LS</td>
<td>139.2</td>
<td>126.3</td>
</tr>
<tr>
<td></td>
<td>(33.6)</td>
<td>(27.8)</td>
</tr>
<tr>
<td>After-tax LS</td>
<td>90.0</td>
<td>143.1</td>
</tr>
<tr>
<td></td>
<td>(54.3)</td>
<td>(24.8)</td>
</tr>
<tr>
<td>Δ LS</td>
<td>−49.2</td>
<td>16.8</td>
</tr>
<tr>
<td>Pre-tax leisure</td>
<td>2.37</td>
<td>3.21</td>
</tr>
<tr>
<td></td>
<td>(2.14)</td>
<td>(1.88)</td>
</tr>
<tr>
<td>After-tax leisure</td>
<td>5.60</td>
<td>2.08</td>
</tr>
<tr>
<td></td>
<td>(3.50)</td>
<td>(1.61)</td>
</tr>
<tr>
<td>Δ Leisure</td>
<td>3.23</td>
<td>−1.13</td>
</tr>
<tr>
<td>Pre-tax consumption</td>
<td>9.92</td>
<td>8.90</td>
</tr>
<tr>
<td></td>
<td>(4.80)</td>
<td>(3.76)</td>
</tr>
<tr>
<td>After-tax consumption</td>
<td>3.56</td>
<td>5.68</td>
</tr>
<tr>
<td></td>
<td>(2.69)</td>
<td>(2.31)</td>
</tr>
<tr>
<td>Δ Consumption</td>
<td>−6.36</td>
<td>−3.22</td>
</tr>
<tr>
<td>Number of subjects</td>
<td>59</td>
<td>21</td>
</tr>
</tbody>
</table>

Means (std. dev.) by those who lowered and those who raised their labor supply in response to introduction of the tax (CT and IT pooled).

Table 4
Diff-in-diff, random-effects GLS and random-effects double-censored Tobit panel regressions.

<table>
<thead>
<tr>
<th>Method</th>
<th>Diff-in-diff</th>
<th>GLS Labor supply&lt;sub&gt;it&lt;/sub&gt;</th>
<th>GLS Labor supply&lt;sub&gt;it&lt;/sub&gt;</th>
<th>GLS Labor supply&lt;sub&gt;it&lt;/sub&gt;</th>
<th>GLS Labor supply&lt;sub&gt;it&lt;/sub&gt;</th>
<th>Tobit Labor supply&lt;sub&gt;it&lt;/sub&gt;</th>
<th>GLS Net earnings&lt;sub&gt;it&lt;/sub&gt;</th>
<th>GLS Net earnings&lt;sub&gt;it&lt;/sub&gt;</th>
</tr>
</thead>
<tbody>
<tr>
<td>Variable/equation</td>
<td>(1)</td>
<td>(2)</td>
<td>(3)</td>
<td>(4)</td>
<td>(5)</td>
<td>(6)</td>
<td>(7)</td>
<td>(8)</td>
</tr>
<tr>
<td>T</td>
<td>−23.9***</td>
<td>−39.9***</td>
<td>−36.5***</td>
<td>−38.2***</td>
<td>−37.6***</td>
<td>−16.4***</td>
<td>−15.0***</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(10.6)</td>
<td>(6.0)</td>
<td>(7.0)</td>
<td>(6.7)</td>
<td>(6.2)</td>
<td>(6.9)</td>
<td>(6.2)</td>
<td></td>
</tr>
<tr>
<td>CT</td>
<td>−</td>
<td>−20.7**</td>
<td>−18.3**</td>
<td>−21.4**</td>
<td>−19.4**</td>
<td>−14.8**</td>
<td>−13.4**</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>(6.0)</td>
<td>(6.9)</td>
<td>(6.7)</td>
<td>(6.2)</td>
<td>(6.9)</td>
<td>(6.2)</td>
<td></td>
</tr>
<tr>
<td>ability&lt;sub&gt;i&lt;/sub&gt;</td>
<td>−1.17</td>
<td>0.90</td>
<td>0.90</td>
<td>0.51</td>
<td>0.66**</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(1.68)</td>
<td>(1.59)</td>
<td>(1.59)</td>
<td>(1.60)</td>
<td>(0.17)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ability&lt;sub&gt;i&lt;/sub&gt; notax</td>
<td>−</td>
<td>−</td>
<td>1.22</td>
<td>−</td>
<td>−</td>
<td>−</td>
<td>0.85**</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>(1.67)</td>
<td>(1.67)</td>
<td></td>
<td></td>
<td></td>
<td>(0.18)</td>
<td></td>
</tr>
<tr>
<td>ability&lt;sub&gt;i&lt;/sub&gt; tax</td>
<td>−</td>
<td>−</td>
<td>0.57</td>
<td>−</td>
<td>−</td>
<td>−</td>
<td>0.46**</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>(1.67)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>(0.18)</td>
<td></td>
</tr>
<tr>
<td>round 2</td>
<td>−</td>
<td>1.2</td>
<td>1.2</td>
<td>1.2</td>
<td>2.3</td>
<td>−0.30</td>
<td>−0.30</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>(5.2)</td>
<td>(5.2)</td>
<td>(5.2)</td>
<td>(5.2)</td>
<td>(0.75)</td>
<td>(0.75)</td>
<td></td>
</tr>
<tr>
<td>round 4/round 4* IT</td>
<td>−</td>
<td>−5.4</td>
<td>−4.0</td>
<td>−10.3</td>
<td>−6.8</td>
<td>0.26</td>
<td>1.09</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>(5.2)</td>
<td>(5.5)</td>
<td>(7.3)</td>
<td>(5.2)</td>
<td>(0.75)</td>
<td>(0.80)</td>
<td></td>
</tr>
<tr>
<td>round 5/round 5* IT</td>
<td>−</td>
<td>1.0</td>
<td>2.3</td>
<td>3.5</td>
<td>0.0</td>
<td>0.59</td>
<td>1.41**</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>(5.2)</td>
<td>(5.5)</td>
<td>(7.3)</td>
<td>(5.2)</td>
<td>(0.75)</td>
<td>(0.80)</td>
<td></td>
</tr>
<tr>
<td>round 4* CT</td>
<td>−</td>
<td>−</td>
<td>−</td>
<td>−0.5</td>
<td>−</td>
<td>−</td>
<td>−</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>(7.3)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>round 5* CT</td>
<td>−</td>
<td>−</td>
<td>−</td>
<td>−</td>
<td>−</td>
<td>−</td>
<td>−</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>constant</td>
<td>−13.8</td>
<td>135.2</td>
<td>130.1</td>
<td>128.2</td>
<td>130.1</td>
<td>−</td>
<td>15.7</td>
<td>14.6</td>
</tr>
<tr>
<td></td>
<td>(11.2)</td>
<td>(5.2)</td>
<td>(10.5)</td>
<td>(10.9)</td>
<td>(10.5)</td>
<td></td>
<td>(1.1)</td>
<td>(1.2)</td>
</tr>
<tr>
<td>Number of obs.</td>
<td>80</td>
<td>400</td>
<td>400</td>
<td>400</td>
<td>400</td>
<td>400</td>
<td>400</td>
<td>400</td>
</tr>
<tr>
<td>ρ</td>
<td>−</td>
<td>.505</td>
<td>.508</td>
<td>.510</td>
<td>.508</td>
<td>.532</td>
<td>.300</td>
<td>.296</td>
</tr>
<tr>
<td>R²log L</td>
<td>.06</td>
<td>.11</td>
<td>.11</td>
<td>.11</td>
<td>−1817</td>
<td>.64</td>
<td>.65</td>
<td></td>
</tr>
</tbody>
</table>

*** p-value less than 0.01. **p-value less than 0.05. *p-value less than 0.10. Difference-in-difference, random-effects GLS estimates and marginal effects from random-effects, double-censored Tobit regressions with treatment indicators (T, CT), the subject’s own labor productivity (ability<sub>i</sub>) and round dummies as the regressors. Heteroskedasticity-robust standard errors in parentheses. The dependent variable is subject i’s average change in labor supply between the tax and no-tax treatment (1), subject i’s labor supply in round t ((2)–(5)), subject i’s labor supply in round t with values 0–6 seconds treated as 0 in (6) and subject i’s net earnings in round t ((7) and (8)).
The coefficients on IT, CT and the round dummies remain unchanged when the subject’s own ability at the labor task is included in regression (3). The ability variable is not significantly different from zero in this or any other regression we ran, nor is it or ability’s ever significant when the latter is included. In (4), we interact ability, with both the no-tax and the tax-inclusive rounds to allow for a subject’s labor productivity to affect differently his labor supply in the presence of the tax. A substitution effect would render the distortive effect of a tax more pronounced for high-ability individuals, with an income effect potentially counteracting this pattern to some extent. The insignificance of the ability*tax coefficient suggests these two forces offset one another. In fact, neither of the interaction terms (or their squared terms when included) is significant, while the gap in the labor supply between the two tax treatments remains highly significant (p < 0.01). We also tried interacting ability, separately with IT and CT. The coefficients on both interaction terms are close to, and not significantly different from, zero.

To explore the magnitude of the labor-supply gap between the tax treatments over time, we interact the round dummies separately with IT and CT in (5). None of the round-dummy interactions is significant, implying that there is no significant time trend in labor supply in either of the tax treatments. The absence of a time trend suggests that neither fatigue nor increased deftness at solving questions constitutes a prevailing force in either tax treatment. The coefficients illustrate a non-monotonic relationship in the difference in labor supply between tax treatments. The 17 s difference between the IT and CT coefficients (in round 3) is highly significant (p=0.05). In round 4, the tax-treatment difference grows to 27 s (p < 0.01); while in round 5, the difference shrinks to 12 s and is no longer significant (p=0.17).

To account for the fact that the decision space in our experiment is censored from below at 0 s and from above at 180 s, we report the marginal effects from a random-effects, double-censored Tobit regression in (6). Eleven percent of the total observations (45 of the total 400 rounds) are right-censored with subjects devoting all of their time to the labor market. The fact that a disproportionate number of these observations (26/45) appear in the no-tax rounds (despite having only 2/5 tax-free rounds) attests to subjects’ labor supply sensitivity. Ten additional right-censored observations appear in CT with the remaining nine observations in IT.

At the other extreme, because consuming all leisure involves pressing three buttons in sequence, each located on a different location of the computer screen, it is physically impossible to stop the clock with the entire 180 s intact.20 Still, the intention to consume all leisure is revealed in 21 instances in which the subject stopped the clock between 1 and 6 s after the round began and did not answer a single question. Indeed, with no instance in which the clock was stopped after 7–10 s had elapsed, six seconds elapsed stands out as the natural threshold for subjects desiring to consume exclusively leisure.

Revealingly, 17 of the 21 instances in which the clock was stopped before 7 s elapsed occurred in IT; the other four observations being in CT, with no single attempt to consume all leisure in the no-tax treatment. Put another way, all instances of voluntary unemployment take place after taxes are introduced, with over 80% of them in the income tax treatment. These differences are highly significant (χ²(2) = 102.8, p < 0.01).

In Tobit regression (6), the dependent variable labor supply is adjusted to take on the value of 0 for all values 0–6 s and equals labor supply for all other values. As in the GLS regressions, the marginal effects on IT and CT reported in (6) continue to differ significantly from zero and from one another (p < 0.01 in both cases). Again, neither ability, nor any of the round dummies is significantly different from zero.21

The inability of subjects’ labor market productivity to explain the variation in their labor market allocation decision is surprising. Higher ability individuals are, by definition, more productive in a given amount of time. If, across the spectrum of abilities, subjects’ average time devoted to the labor market is similar, then we would expect labor market earnings to rise with ability. To examine this possibility, we regress ability, as well as the tax treatment and round indicators on subject i’s net real labor market earnings (henceforth abbreviated as “net earnings”) in round t.22

The highly significant coefficient of 0.66 on ability, in (7) indicates that each additional question a subject is capable of solving when he devotes the entire 3 min to the labor task earns the subject 0.66 more points in net labor-market income. In (8), we allow for ability to explain differentially a subject’s net earnings in the tax and no-tax treatments. Indeed, each additional unit of ability translates into 0.85 additional points of income per round in the no-tax treatment compared to 0.46 points per round in the tax treatments.

Since abilities are similarly distributed across tax treatments, our main finding that subjects reduce their labor supply by a larger amount in IT than in CT should translate into more sharply reduced labor-market earnings with the introduction of the tax in IT than in CT. The 16.4-point drop in net labor income in IT is in fact significantly larger than the 14.8-point decrease in net labor income in CT (p=0.038). The IT and CT estimates are robust to the alternative specification of ability, in (8) and continue to differ significantly from one another (p=.045).

20 The subject first needs to press the “Start” button which starts the clock, then presses “Cancel” on the message box containing the multiplication question and finally presses the “Stop” button (see the screenshot in Appendix C).
21 The Tobit regression is estimated using Gauss-Hermite quadrature with 12 points of evaluation (Stata’s default). As a robustness check, we re-estimate (6) based on both 8 and 16 quadrature points. None of the coefficients changes by more than 0.001%. We also estimated the GLS specifications in (3)–(5) using Tobit regressions. The qualitative results are identical.
22 In each of the two no-tax rounds, a subject’s net earnings simply equal his gross earnings (two points times the number of correctly answered multiplication questions), whereas net earnings are half of gross earnings for the three tax-inclusive rounds in both IT and CT.
5. Conclusions

In this paper, we test experimentally the equivalence between a consumption and a labor-income tax. Our main finding is that, across the spectrum of labor-market productivities, individuals work more under the consumption tax. We provide a simple theoretical model that invokes tax misperception to demonstrate our main finding and its robustness to the incorporation of an income effect. An important normative implication of the model is the potential welfare gain associated with a shift from income to consumption taxes. Such a shift can enhance individuals’ utility without changing government tax revenue.

Two often-raised arguments in favor of a shift to a consumption tax are its administrative advantages compared to an income tax (i.e., simplicity of measuring consumption versus labor income and ease of collection and enforcement) and the elimination of the inter-temporal distortion of consumption allocation caused by the taxation of capital income (see McCaffery, 2002). Our paper uncovers evidence for an additional, perceptual advantage. Our policy implication of a shift to consumption-based taxation is in line with documented trends in most OECD countries.

Acknowledgments

We thank Ofer Azar, Bram Cadsby, Danny Cohen-Zada, Naomi Feldman, Todd Kaplan, Esteban Klor, Yoram Margalioth, Klaus Schmidt, Jeremy Tobacman, Nat Wilcox, seminar participants at numerous conferences and departmental seminars and especially two anonymous referees for valuable comments. Inbal Fitusi, Gal Gabai, Yoram Gabay and Ayala Waichman provided excellent research assistance. We gratefully acknowledge the cooperation of Falafel El Hanegev and Mister Pizza and the financial support of the Maurice Falk Institute for Economic Research in Israel.

Appendix A. Proof of proposition

Before outlining the formal argument, it would be constructive to demonstrate the role of tax misperception in the decision-making process. The individual, unlike in the standard case, solves a two-stage (sequential) optimization process. In the first stage, based on the misperceived consumption tax burden and the associated budget constraint, the individual makes a labor-leisure choice. Then, in the second stage, given the time allocation already made, the individual is faced with the true budget constraint with its associated actual level of consumption (lower than the one anticipated in the first stage). Now we turn to the formal argument. We will distinguish between two cases. We first examine the extreme case of $\alpha = 0$, namely the individual is completely oblivious to the tax. Next we discuss the general case of $0 < \alpha < 1$.

Consider first the case $\alpha = 0$. Consider a wage-tax regime ($B^{WT}$ in Fig. 2; that is, we set $t > 0$ and $s = t = 0$. Denote the wage-tax rate by $t^{WT}$. Let $l^{WT}$ and $U^{WT}$ denote the individual’s choice of labor supply under the wage-tax regime and his corresponding utility level, respectively. Let $l^{NT}$ be the individual’s choice of labor supply under a no-tax regime (i.e., setting $s = t = \tau = 0$, given by $B^{NT}$ in Fig. 2). By virtue of our quasi-linear specification, $l^{NT}$ equals the choice of labor

![Fig. 2. Efficiency gain from a shift to a consumption tax.](image)
supplied under a lump-sum tax regime (i.e., setting $\tau < 0$ and $s = t = 0$), which entails no deadweight loss. Note further that $t_{NT} > t_{WT}$, since labor supply is strictly increasing with respect to the net-of-tax wage rate (as we assume that $h$ is convex). Denote by $c_{NT}$ the implicit solution to
\[
U(c_{NT}, t_{NT}) = U_{WT}.
\] (2)

In words, $c_{NT}$ is the level of consumption that yields the individual the same utility level as that obtained under a wage-tax regime when the individual works $t_{NT}$ hours.

Consider a consumption-tax regime (i.e., $s > 0$ and $t = \tau = 0$). Denote the consumption-tax rate by $s_{CT}$, where $s_{CT}$ is given by the implicit solution to
\[
t_{NT} \cdot w = c_{NT} \cdot (1 + s_{CT}).
\] (3)

In words, the consumption-tax rate is set such that the consumption-labor pair $(c_{NT}, t_{NT})$ lies on the after-tax budget line $B_{CT}$ shown in Fig. 2. Since $\alpha = 0$, an individual faced with the consumption-tax regime chooses to work $t_{NT}$ hours. His resultant utility equals $U_{WT}$ according to (2). Notice that the bundle $(c_{NT}, t_{NT})$ lies at the intersection of the $U_{WT}$ indifference curve and $B_{CT}$, rather than at the point of tangency. This follows from the individual's misperception that induces him to work more hours than the (ex-post) optimal amount given the actual (true) budget constraint. The tax revenues raised by the consumption tax thus equal the tax revenues that would be raised by a lump-sum tax (for the same utility level).\(^2\) The tax revenues raised by a lump-sum tax are, by definition, higher than the tax revenues from the wage-tax regime. To sum up, with complete misperception a consumption tax would yield the same allocation as that obtained with a lump-sum tax (due to the absence of income effects), which obviously entails no deadweight loss and is hence superior to a wage-tax regime. This completes the proof for the case $\alpha = 0$.

We turn next to the case $0 < \alpha < 1$. Denote by $t'$ the amount of labor chosen by an individual faced with a consumption-tax rate given by the implicit solution to (3). Due to the strict concavity of the utility function ($h'' > 0$), which implies the convex-shaped indifference curve, it follows that the marginal tax rate associated with the consumption-tax regime is lower than that associated with the wage-tax regime; formally
\[
\frac{1}{1 + s_{CT}} > 1 - t_{WT}
\]
(that is, $B_{CT}$ is steeper than $B_{WT}$). Notice that otherwise the bundle $(c_{NT}, t_{NT})$ would have lied below the indifference curve $U_{WT}$ which would have violated condition (2). Thus, $t_{WT} < t' < t_{NT}$. The first inequality follows from the fact that labor supply strictly increases with respect to the net-of-tax wage rate and the individual's misperception of the consumption tax ($\alpha < 1$), implying that the misperceived budget line is steeper than the true one; whereas the second inequality follows from both the positive slope of labor supply with respect to the net-of-tax wage rate and $\alpha > 0$ (that is, some positive fraction of tax is being internalized, inducing the individual to work less than under a no-tax regime). Let $c'$ represent the individual's consumption level from $t'$. It follows from the strict concavity of the utility function that $U(c', t') > U_{WT}$. In other words, the bundle $(c', t')$ lies above the indifference curve $U_{WT}$ (see Fig. 2).

Consider now an alternative consumption-tax regime. Denote the consumption tax by $s_{WT}$, where $s_{WT}$ is given by the implicit solution to
\[
t_{WT} \cdot w = c_{WT} \cdot (1 + s_{WT}).
\] (4)

That is, we set a consumption-tax rate that yields an after-tax budget line identical to that of the wage-tax regime ($B_{WT}$). Denote by $c'$ and $t'$, respectively, the consumption level and the amount of labor chosen by the individual faced with the consumption-tax regime in (4). Similar to the above reasoning, it follows that $t_{WT} < t' < t_{NT}$. Thus, $U(c', t') < U_{WT}$. In other words, the bundle $(c', t')$ lies below the indifference curve $U_{WT}$ (see Fig. 2). Notice again that due to misperception the individual chooses his labor supply presuming that the budget line is steeper than $B_{WT}$. This misperception implies an ex-post suboptimal choice (with respect to the actual budget line $B_{WT}$) yielding $U < U_{WT}$. Moreover, as $s_{WT} > s_{CT}$, by virtue of the convexity of $h$, it follows that $t' < t$. Since the utility is continuous in both arguments and the budget set is convex, the indirect utility function is continuous in $s$, then by the intermediate value theorem it follows that there exists some consumption-tax rate, $\hat{s}$, where $s_{CT} < \hat{s} < s_{WT}$, with the individual's corresponding consumption and labor choices given by $\hat{c}$ and $\hat{t}$, such that $U(\hat{c}, \hat{t}) = U_{WT}$ (given by point $A$ in the figure). Moreover, $t_{WT} < \hat{t} < t_{NT}$. A shift from a wage tax to a consumption tax at the rate $\hat{s}$ moves the individual along the indifference curve $U_{WT}$ toward the bundle $(c_{NT}, t_{NT})$ chosen under a lump-sum tax (set to maximize tax revenues, by construction). Consequently, the tax revenues from the consumption tax $\hat{s}$ are strictly higher than those from the wage tax $t_{WT}$. In sum, we have constructed a budget constraint imposing a consumption-tax rate lower than the theoretically equivalent wage-tax rate such that the utility level is the

\(^2\) Notice that a lump-sum tax would be given by a budget constraint tangent to the indifference curve $t_{WT}$ at the point $(c_{NT}, t_{NT})$ (parallel to $B_{NT}$).
same under both regimes and established that the tax revenues attained under the consumption tax would be higher than those attained under a wage tax.

### Appendix B. Robustness to income effect

The theoretical model employs a quasi-linear functional form for the utility function to illustrate how, in the presence of tax misperception, a shift from an income to a consumption tax yields a welfare gain. The exercise reveals that the government could raise higher tax revenues under a consumption tax, while providing the individual with the same level of utility as under an income tax.

A quasi-linear utility function implies that no income effect exists. In what follows, we examine the robustness of our qualitative results to the incorporation of an income effect. We maintain our separability assumption (between consumption and labor) but allow for diminishing marginal utility of consumption. Formally, we assume that utility takes the following iso-elastic functional form: 

\[ U(c,h) = c^k / k - \hat{h}^b / b, \]

where \( 0 < k < 1 \) and \( b > 1 \).

Thus, marginal utility from consumption is diminishing whereas marginal cost (disutility) of labor is increasing (as in the quasi-linear specification). Notice that for the case \( k = 1 \), we obtain the quasi-linear specification employed in the theoretical model (with no income effect). Because the model does not allow for a closed-form solution (we were unable to generalize the proposition proved for the quasi-linear specification), we turn to numerical analysis and make the following parametric assumptions: \( w = 1 \), \( t = 0.5 \) and \( b = 2 \). The simulation results are robust to a wide range of parameters, as long as the substitution effect dominates the income effect, implying an upward-sloping labor supply curve.

We conduct comparative-static analysis with respect to two parameters: (i) the parameter \( k \), capturing the relative magnitude of the income effect (as the parameter \( k \) decreases, the marginal utility from consumption diminishes more rapidly; hence, the income effect becomes larger; where \( k = 1 \) implies that there is no income effect), (ii) the parameter \( \epsilon \), capturing the degree of misperception (where \( \epsilon = 0 \) implies full misperception and \( \epsilon = 1 \) implies no misperception).

The model is solved repeatedly for different pairs of \( k \) and \( \epsilon \). We first solve for the optimal choice of the individual under an income tax (\( t = 0.5 \), namely a flat marginal tax rate of 50%). We then calibrate the consumption-tax rate so that the utility derived under an income tax is maintained. Finally, we compare the tax revenues obtained under the two regimes. The following table summarizes the results. As a guide to interpreting the table, consider the first row: with no income effect in place (\( k = 1 \)) and with a high degree of misperception (\( \epsilon = 0.1 \)), the tax revenues obtained under an income-tax regime are 0.25, whereas those obtained under the calibrated consumption-tax regime are 0.37.

Several observations follow from close inspection of the table. First and foremost, the welfare dominance of the consumption tax over the income tax holds for all degrees of misperception and over the entire range of income effects examined. This is reflected by the fact that tax revenues are consistently higher under the consumption tax. Second, and as expected, as the income effect becomes more manifest (i.e., the smaller the value of \( k \), tax revenues increase under both tax regimes (for any degree of misperception \( \epsilon \)) as individuals choose to work more; although the difference between the two tax regimes in terms of the tax revenues generated (hence the welfare dominance of the consumption tax) decreases. Thus, the presence of an income effect mitigates the perceptual advantage of the consumption tax, but does not eliminate it altogether. Moreover, as one would expect, the higher the degree of misperception (i.e., the lower the value of \( \epsilon \)), the larger is the difference between the tax revenues raised by the two tax regimes, for any given level of income effect (i.e., for any value of \( k \)). Put differently, the welfare gain from a shift to a consumption tax is higher as the extent of misperception becomes more pronounced. While not reported in the table, one can show that when \( \epsilon = 1 \) (no misperception), the two tax regimes are equivalent, hence yielding identical tax revenues. In such a case, we obtain the standard equivalence result presumed by the literature in the absence of any misperception (Table B1).

### Appendix C. Instructions for participant

This is a decision-making experiment in time and resource allocation. Funds for this experiment have been provided by an external research foundation. Take the time to read carefully the instructions. A good understanding of the instructions and well thought out decisions during the experiment can earn you a considerable remuneration. Earnings from the experiment will be paid at the end of the experiment to you in cash and in vouchers exchangeable for goods. The experiment consists of three parts. Below are the instructions for the first of three parts.

#### Part One

During the first part of this experiment, you will be given 3 min to solve as many two-digit by two-digit multiplication questions as possible. This preliminary exercise will allow you to familiarize yourself with the software in preparation for the second and third parts of the experiment. For each correctly answered question in this part, you will earn 0.5 shekel to be paid in cash at the end of the experiment.

This completes the instructions for Part One of the experiment. If you have any questions, please raise your hand and a monitor will come to assist you.
Below is an example of a screenshot from Part One of the experiment.

![Screenshot](image)

**Part Two**

This second part of the experiment consists of two stages, time and salary allocation stages.

**Time-allocation stage**

In the first time-allocation stage, you are given 3 min. You need to decide how much of this time you wish to devote to answering multiplication questions. For each correctly answered question, you earn two points. You will be able to redeem

---

**Table B1**

Simulating the role of the income effect.

<table>
<thead>
<tr>
<th>(k)</th>
<th>(\alpha)</th>
<th>Revenues under consumption tax</th>
<th>Revenues under income tax</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.1</td>
<td>0.37</td>
<td>0.25</td>
</tr>
<tr>
<td>0.8</td>
<td>0.1</td>
<td>0.42</td>
<td>0.31</td>
</tr>
<tr>
<td>0.6</td>
<td>0.1</td>
<td>0.45</td>
<td>0.37</td>
</tr>
<tr>
<td>0.4</td>
<td>0.1</td>
<td>0.48</td>
<td>0.42</td>
</tr>
<tr>
<td>0.2</td>
<td>0.1</td>
<td>0.49</td>
<td>0.46</td>
</tr>
<tr>
<td>1</td>
<td>0.3</td>
<td>0.36</td>
<td>0.25</td>
</tr>
<tr>
<td>0.8</td>
<td>0.3</td>
<td>0.4</td>
<td>0.31</td>
</tr>
<tr>
<td>0.6</td>
<td>0.3</td>
<td>0.44</td>
<td>0.37</td>
</tr>
<tr>
<td>0.4</td>
<td>0.3</td>
<td>0.46</td>
<td>0.42</td>
</tr>
<tr>
<td>0.2</td>
<td>0.3</td>
<td>0.48</td>
<td>0.46</td>
</tr>
<tr>
<td>1</td>
<td>0.5</td>
<td>0.33</td>
<td>0.25</td>
</tr>
<tr>
<td>0.8</td>
<td>0.5</td>
<td>0.38</td>
<td>0.31</td>
</tr>
<tr>
<td>0.6</td>
<td>0.5</td>
<td>0.42</td>
<td>0.37</td>
</tr>
<tr>
<td>0.4</td>
<td>0.5</td>
<td>0.45</td>
<td>0.42</td>
</tr>
<tr>
<td>0.2</td>
<td>0.5</td>
<td>0.48</td>
<td>0.46</td>
</tr>
<tr>
<td>1</td>
<td>0.7</td>
<td>0.29</td>
<td>0.25</td>
</tr>
<tr>
<td>0.8</td>
<td>0.7</td>
<td>0.35</td>
<td>0.31</td>
</tr>
<tr>
<td>0.6</td>
<td>0.7</td>
<td>0.4</td>
<td>0.37</td>
</tr>
<tr>
<td>0.4</td>
<td>0.7</td>
<td>0.44</td>
<td>0.42</td>
</tr>
<tr>
<td>0.2</td>
<td>0.7</td>
<td>0.47</td>
<td>0.46</td>
</tr>
<tr>
<td>1</td>
<td>0.9</td>
<td>0.26</td>
<td>0.25</td>
</tr>
<tr>
<td>0.8</td>
<td>0.9</td>
<td>0.32</td>
<td>0.31</td>
</tr>
<tr>
<td>0.6</td>
<td>0.9</td>
<td>0.38</td>
<td>0.37</td>
</tr>
<tr>
<td>0.4</td>
<td>0.9</td>
<td>0.43</td>
<td>0.42</td>
</tr>
<tr>
<td>0.2</td>
<td>0.9</td>
<td>0.47</td>
<td>0.46</td>
</tr>
</tbody>
</table>
your total earned points for falafel sandwiches and pizza slices, as explained below. At any time during these 3 min, you may choose to stop answering questions by pressing “Cancel” instead of answering the question and then by pressing the “Stop” button. Whatever time remains from the initial 3 min is converted to bottled soft drinks according to the following exchange rate: every 15 s remaining on the clock is worth one bottled soft drink. Below is an example of a screenshot from the time-allocation stage. Please take a look at this now.

**Point-allocation stage**

You will now be asked to allocate the points you earned from the first time-allocation stage (at a rate of 2 points per correctly answered question) between half falafel sandwiches and pizza slices. Each point can be exchanged for either a two half falafel sandwiches or two pizza slices. You must spend all of your earned points on any combination you like of half falafel sandwiches and pizza slices. Note that the bottled soft drinks you earned in the first stage cannot be exchanged for falafel sandwiches or pizza slices. Below is an example of a screenshot from the point-allocation stage of the experiment. Please take a look at this now.

After you have decided how to allocate your earned points between half falafel sandwiches and pizza slices, press the “Confirm Choice” button, followed by the “Proceed to the Next Round” button. The time-allocation and point-allocation stages will be repeated once more.

**Payment**

At the end of the experiment, you will be paid the money you earned from correctly solving multiplication questions in the first part. Of the two rounds of the second part, one round will be randomly chosen and you will be given vouchers for the number of bottled soft drinks, half falafel sandwiches and pizza slices that you earned in the randomly chosen round. The falafel vouchers are valid for redemption for up to 1 year at Falafel El HaNegev in Mercaz Oren. The pizza vouchers are valid for redemption for up to 1 year at American Pizza also in Mercaz Oren. The soft drink vouchers are redeemable for up to 1 year at either Falafel El HaNegev or American Pizza in Mercaz Oren.

This completes the instructions for Part Two of the experiment. If you have any questions, please raise your hand and a monitor will come to assist you.

Below are two examples of screenshots from Part Two, the first from the time-allocation stage and the second from the point-allocation stage.
Part Three—Tax on prices of goods

Time-allocation stage

The time-allocation stage of the third part is identical to that of the second part. That is, you are given 3 min during which time you must decide how much of the 3 min to devote to answering multiplication questions in exchange for points and how much of the 3 min to preserve in exchange for bottled soft drinks. Each correctly answered question earns you two points redeemable for falafel sandwiches and pizza slices, as explained below. For every 15 s you choose to preserve on the clock you earn one bottled soft drink. Below is an example of a screenshot from the time-allocation stage of the experiment. Please take a look at this now.

Point-allocation stage

Like the point-allocation stage of the second part of the experiment, you will now be asked to allocate the points you earned from the first time-allocation stage (at a rate of 2 points per correctly answered question) between half falafel sandwiches and pizza slices. However, there is now a tax on the prices of these two goods such that the price of each good including the tax is double the price paid in the second part. In other words, each two points can be exchanged for either half a falafel sandwich or one pizza slice. You must spend all of your earned points on any combination you like of falafel sandwiches and pizza slices. Note that the bottled soft drinks you earned in the first stage are not subject to the tax and cannot be exchanged for falafel sandwiches or pizza slices. Below is an example of a screenshot from the point-allocation stage of the experiment. Please take a look at this now.

After you have decided how to allocate your earned points between half falafel sandwiches and pizza slices, press the “Confirm Choice” button, followed by the “Proceed to the Next Round” button. The time-allocation and point-allocation stages will be repeated twice more.

Payment

At the end of the experiment, you will be paid the money you earned from correctly solving multiplication questions in Part One. You will also be paid the number of vouchers for bottled soft drinks, half falafel sandwiches and pizza slices that you earned from one randomly chosen round of Part Two. Finally, you will receive additional vouchers for bottled soft drinks, half falafel sandwiches and pizza slices according to your earnings from one randomly chosen round among the three in which you
participated in this Part Three of the experiment. The falafel vouchers may be redeemed at Falafel El Hanegev in Mercaz Oren. The pizza vouchers may be redeemed at American Pizza also in Mercaz Oren. The soft drink vouchers may be redeemed at either Falafel El Hanegev or American Pizza in Mercaz Oren. All vouchers are valid for up to 1 year.

This completes the instructions for this part of the experiment. If you have any questions, please raise your hand and a monitor will come to assist you. Below are examples of three screenshots from Part Three, the first during the time-allocation stage, the second at the end of this stage and the third from the point-allocation stage.

Below are two examples of screenshots from Part Two, the first from the time-allocation stage and the second from the point-allocation stage.
Participant ______

**Part Three—Tax on correctly answered questions**

**Time-allocation stage**

Like the time-allocation part of the second part, you are given 3 min during which time you must decide how much of the 3 min to devote to answering multiplication questions in exchange for points and how much of the 3 min to preserve in exchange for bottled soft drinks. However, there is now a tax on the points you earn. Each correctly answered question earns you two points. But with the tax on the points you will keep half of the total points you earn. In other words, your net income is 1 point for each correctly answered question. These points are redeemable for falafel sandwiches and pizza slices, as explained below. You must spend all of your earned points on any combination you like of falafel sandwiches and pizza slices. For every 15 s you choose to preserve on the clock you earn one bottled soft drink. Note that the bottled soft drinks you earn are not subject to the tax.

**Point allocation**

The point-allocation part of the third part is identical to that of the second part. That is, you will now be asked to allocate the points you earned from the first time-allocation part (at a rate of 2 points per correctly answered question) between half falafel sandwiches and pizza slices. Each point can be exchanged for either two half falafel sandwiches or two pizza slices. You must spend all of your earned points on any combination you like of half falafel sandwiches and pizza slices. Note that the bottled soft drinks you earned in the first part cannot be exchanged for falafel sandwiches or pizza slices.

After you have decided how to allocate your earned points between half falafel sandwiches and pizza slices, press the “Confirm Choice” button, followed by the “Proceed to the Next Round” button. The time-allocation and point-allocation stages will be repeated twice more.

**Payment**

At the end of the experiment, you will be paid the money you earned from correctly solving multiplication questions in Part One. You will also be paid the number of vouchers for bottled soft drinks, half falafel sandwiches and pizza slices that you earned from one randomly chosen round of Part Two. Finally, you will receive additional vouchers for bottled soft drinks, half falafel sandwiches and pizza slices according to your earnings from one randomly chosen round among the three in which you participated in this Part Three of the experiment. The falafel vouchers may be redeemed at Falafel El Hanegev in Mercaz Oren. The pizza vouchers may be redeemed at American Pizza also in Mercaz Oren. The soft drink vouchers may be redeemed at either Falafel El Hanegev or American Pizza in Mercaz Oren. All vouchers are valid for up to 1 year.

This completes the instructions for this part of the experiment. If you have any questions, please raise your hand and a monitor will come to assist you. Below are examples of three screenshots from Part Three, the first during the time-allocation stage, the second at the end of this stage and the third from the point-allocation stage.

Below are two examples of screenshots from Part Two, the first from the time-allocation stage and the second from the point-allocation stage.
Time Allocation

Timer: 01:41

Drinks Left: 6

Salary: 10 Points

Correct: 5

Wrong: 1

Consumption

Confirm Choice

Price of 1 Pizza = 1.00 Points

Price of 1/2 Falafel = 1.00 Points

Salary: 5

Basket chosen: (3,2)

1.00 x 3 + 1.00 x 2 = 5.00
References


Feldman, N.E., Katuscak, P. Effects of predictable tax liability variation on household labor income. Unpublished manuscript.


