The informal sector in transition: Tax evasion in an institutional vacuum

Gerxhani, K.

Citation for published version (APA):
Chapter 4

An Experimental Study of Tax Evasion in Albania and the Netherlands

"If we were to stop a person in the street and ask him or her why people evade tax, the answer would almost certainly be 'greed'"

- WEBLEY et al. (1991:8)

4.1. Introduction

This chapter focuses on the analysis of tax evasion behavior, using laboratory experiments in two European countries, Albania in the East and the Netherlands in the West. In this way, we hope to get a grip on the individual decision involved and a better understanding of the factors that cause differences in the level of tax evasion between a developed country and a country in transition.

Providing a comparative experimental study of tax evasion in a developed country and a country in transition is just one of the contributions of this chapter. Two other contributions, to be discussed more extensively below, are:

(i) it compares tax evasion across subject pools within each country. This allows us to compare differences across countries to differences within a country.

(ii) it presents a new laboratory environment distinguishing two possible sources of income: registered and unregistered. This novel feature of our design is important. In most countries, an employer in the 'official' labor market reports her/his own and her/his employee(s)' income to the tax authorities, whereas income from self-employment and 'unofficial' activities must be reported by the income earners themselves. Hence, individuals can affect their opportunities to evade taxes by the choice of their source
of income. Previous experimental designs did not allow for this endogeneity in the choice of income type.

The organization of this chapter is as follows. Section 4.2 discusses the existing evidence of tax evasion in Albania and the Netherlands, and reviews the relevant literature. Section 4.3 presents our experimental design. Section 4.4 discusses the experimental results. The implications of these results follow in section 4.5. Section 4.6 concludes.

4.2. Country evidence on tax evasion and the relevant literature

More studies have taken place about the Netherlands than about Albania. These typically do not refer to tax evasion per se. Instead, they refer to the ‘hidden’ or ‘shadow’ economy. We assume that tax evasion is a major element in the phenomena they refer to. Boeschoten and Fase (1984) quantify the ‘hidden’ economy in the Netherlands in 1965 – 1982. They suggest a rising trend up to about 20% of GDP in this period. On the other hand, Frey and Weck (1981) give an estimate of 9.6% for 1977-1981. Schneider and Enste (2000) report on the average size of the ‘shadow’ economy for OECD countries, over 1990-93; 1994-95 and 1996-97. Their estimate for the Netherlands is about 13-16%, 13.7% and 13.8% respectively.

There is no quantitative evidence on the existence of tax evasion in Albania during communism. Nevertheless, as stated in chapter 3, there was an informal sector. Taxes could be evaded, for example, by not working in official employment or by selling privately grown agricultural products. For the period of transition, evidence on tax evasion is not based on careful analysis either. However, anecdotal evidence suggests high levels of tax evasion in current Albania. There have been some attempts -mainly based on sample surveys- to quantify the informal sector in general. For example, the 1996 EBRD Transition Report claims that 70% of the households in Albania do not pay their utility bills (tax bills included). In addition, a study of the Albanian Center for Economic Research (ACER, 1999) reports that 73% of the surveyed enterprises (the sample unit) do not declare all of their profit. On average, this underreporting constitutes 20% of their profit. According to a recent study of UNDP (2000), the ‘black’ economy is thriving, accounting for an estimated 50% of GDP and

---

71 However, note that this evasion of taxes was less significant than in developed market economies, because personal income tax was less important.
depriving the government of much-needed tax revenue. All in all, in spite of a lack of statistical evidence, a common opinion appears to be that tax evasion is higher in Albania than in the Netherlands.

The studies on the Netherlands and Albania have applied the traditional tools in this field, focusing on the two main approaches: the direct and the indirect approach. As chapter 2 points out, the use of laboratory experiments is becoming a popular method, especially with respect to analysis of tax evasion. Though there are varieties to the basic design of this type of experiments (cf. 2.6.2), the results regarding tax compliance are quite robust. These results show that:

(i) tax evasion increases with the tax rate (Friedland et al., 1978 and Collins et al., 1990);
(ii) tax evasion decreases with the level of fines (Friedland et al., 1978 and Collins et al., 1990);
(iii) tax evasion decreases with the audit probability (Friedland et al., 1978);  
(iv) tax evasion is lower when the proceeds are used to provide a public good (Alm et al., 1991);
(v) the decision about tax evasion is made jointly with the labor supply decision (Collins et al., 1990);
(vi) a large subset of people never cheat, because they appear to believe that cheating is wrong (Baldrey, 1986);
(vii) tax evasion increases with income (Giese and Hoffman, 1999);
(viii) women evade taxes less than men do (Giese and Hoffman, 1999).

The experiments we ran allow us to compare behavior across countries when institutions are controlled. This is, in fact, one of the main reasons the experimental method was used in the first place. The choice to evade taxes may be affected by institutions, by culture, or by both. By controlling for differences in institutions, information about the effect of culture can be collected. If Albanians do indeed evade taxes more often than the Dutch do, our experiment will provide information as to whether Albanian culture is more receptive to tax evasion or whether the difference is more likely due to more primitive tax collection institutions. In addition, by conducting and analyzing the same experiment in two countries, this chapter
compares differences between two subject pools within the same country, to differences between subject pools across countries.

4.3. Experimental design and procedures

The experiments were run manually in order to make it possible to organize them at various locations without having to arrange computer facilities. In each session there were 12 participants, divided into 3 groups of 4. A session consisted of 8 independent rounds. The experiment itself generally lasted less than one hour. Calculating payoffs took some time, however, so the last subject usually left about 75 minutes after the start of the experiment. All experiments took place in 1999-2000.

Let us start with a general overview of the design. More details are given below. In the first part of a round, subjects' personal income for that round is determined. This income is private information, unknown to the experimenter or other participants. For rounds 1-3 this is all that happens. In rounds 4-8 there is a second part where subjects have to report their income to the experimenter. This reported income is taxed. In some cases there is an audit. In case of underreporting, the subject is fined. Tax proceeds are aggregated within a group of 4 and in most sessions divided equally among the group members. In two sessions the taxes were taken but the proceeds were not redistributed. Except for these two sessions, taxes may be seen as contributions to a public good with mpcr = 1 (see Isaac et al., 1984). Groups remain constant across rounds and know how taxes and the public good are determined, but no information is provided between rounds about the tax proceeds themselves. Hence, even within groups, subjects are not provided with any kind of information about other subjects' choices.

One important element of this experimental design is the distinct subject pools. These varied along two dimensions: country (the Netherlands, NL, and Albania, AL) and socio-economic category. For the latter, the following groups were distinguished: (i) high school students (HS); (ii) university students (US); (iii) high school teachers (HT); (iv) university non-academic personnel (UP); and (v) university teachers (UT).  

---

72 In one session (high school teachers in the Netherlands) there were only eight participants (two groups of four).
73 We did not run sessions with university teachers in Amsterdam because we feared that the personal contacts both authors have with most of the faculty could cause serious experimenter effects.
Sessions in Albania were run at the ‘Harry Fultz’ high school in Tirana and at the Economics Faculty of the University of Tirana. The Dutch sessions were run at the ‘Vossius Gymnasium’ high school in Amsterdam and at the University of Amsterdam.\textsuperscript{74}

A second treatment concerns the audit probability. For all subject pools, we used a (high) probability of $1/2$ of auditing the reported income (as described below). For the students (NL-US and AL-US) additional sessions were run, where the audit probability was equal to $1/6$. For the students in Albania (AL-US), we also ran sessions with no audits (probability equal to 0). Finally, as a third treatment, sessions with Albanian students (AL-US) were run, where the tax proceeds were not redistributed as a public good.\textsuperscript{75} Table 4.1 summarizes the sessions.

Note from table 4.1 that the average earnings in experimental francs (the currency used in the experiment) across subject pools did not differ much. A possible exception is that earnings were higher in sessions with a low audit probability. Furthermore, earnings were low in the sessions without public goods. This simply reflects the fact that tax proceeds were not redistributed amongst the participants. Note that the exchange rate was varied from francs to local currency in order to account for differences in purchasing power, both across countries and across pools within a country.

Following Roth et al. (1991) and Brandts et al. (2000) three aspects of the design were considered, which require special attention when conducting a multi-national experiment: experimenter effects, language effects, and currency effects.

The term \textit{experimenter effect} refers to the possibility that different sessions of the same experimental treatment may yield different results, due to possible effects of uncontrolled procedural differences across locations. In our case, the two experimenters ran all sessions together. Therefore, in principle, this kind of experimenter effects is not expected. In addition, the experiments were ‘double blind’ as long as no audit took place. When there was no audit, the experimenters had no way of knowing whether or not the subjects had truthfully declared their income. When there was an audit, the experimenters did discover whether or not the participant had reported truthfully.

\textsuperscript{74} We are grateful to the deans of both high schools and of the Faculty of Economics in Tirana for giving us the opportunity to run our experiments at their schools.

\textsuperscript{75} The sessions without audit and those without public good were designed after the results from the other sessions had been analyzed.
### Table 4.1: Experimental sessions

<table>
<thead>
<tr>
<th>Subject Pool</th>
<th># sessions/ #subjects</th>
<th>Audit Probability</th>
<th>Average Earnings</th>
<th>Public good?</th>
<th>Exchange rate</th>
</tr>
</thead>
<tbody>
<tr>
<td>AL-HS</td>
<td>2/24</td>
<td>1/2</td>
<td>3490</td>
<td>Yes</td>
<td>100 fr = 33 lek</td>
</tr>
<tr>
<td>AL-US</td>
<td>2/24</td>
<td>1/6</td>
<td>4284</td>
<td>Yes</td>
<td>100 fr = 33 lek</td>
</tr>
<tr>
<td>AL-US</td>
<td>2/24</td>
<td>1/2</td>
<td>3160</td>
<td>Yes</td>
<td>100 fr = 33 lek</td>
</tr>
<tr>
<td>AL-US</td>
<td>2/24</td>
<td>1/2</td>
<td>3583</td>
<td>Yes</td>
<td>100 fr = 33 lek</td>
</tr>
<tr>
<td>AL-US</td>
<td>2/24</td>
<td>0</td>
<td>2898</td>
<td>No</td>
<td>100 fr = 33 lek</td>
</tr>
<tr>
<td>AL-HT</td>
<td>2/24</td>
<td>1/2</td>
<td>3447</td>
<td>Yes</td>
<td>100 fr = 50 lek</td>
</tr>
<tr>
<td>AL-UP</td>
<td>2/24</td>
<td>1/2</td>
<td>3334</td>
<td>Yes</td>
<td>100 fr = 50 lek</td>
</tr>
<tr>
<td>AL-UT</td>
<td>2/24</td>
<td>1/2</td>
<td>3177</td>
<td>Yes</td>
<td>100 fr = 50 lek</td>
</tr>
<tr>
<td>NL-HS</td>
<td>2/24</td>
<td>1/2</td>
<td>3315</td>
<td>Yes</td>
<td>100 fr = fl 1.00</td>
</tr>
<tr>
<td>NL-US</td>
<td>2/24</td>
<td>1/6</td>
<td>3414</td>
<td>Yes</td>
<td>100 fr = fl 1.00</td>
</tr>
<tr>
<td>NL-US</td>
<td>2/24</td>
<td>1/2</td>
<td>3372</td>
<td>Yes</td>
<td>100 fr = fl 1.00</td>
</tr>
<tr>
<td>NL-HT</td>
<td>1/8</td>
<td>1/2</td>
<td>3501</td>
<td>Yes</td>
<td>100 fr = fl 2.00</td>
</tr>
<tr>
<td>NL-UP</td>
<td>1/12</td>
<td>1/2</td>
<td>3076</td>
<td>Yes</td>
<td>100 fr = fl 2.00</td>
</tr>
</tbody>
</table>

*Note:* labels are defined in the main text; average earnings are in experimental francs (fr). The exchange rate is from francs to the local currency, Lek (Albania) and Guilder (NL, denoted by fl.). The official exchange rates were $1 = 141 Lek and $1 = 2,20 fl. at the time of the experiments.

To control for unwanted *language effects* the instructions for the experiment were initially written in English, and then translated into Albanian and Dutch.

With respect to *currency effects*, it should be noted that the differences in wealth and purchasing power between the two countries and across groups within a country are large. The exchange rate was varied from experimental francs to the local currency as described in table 4.1 in order to maintain sufficient and comparable saliency across groups. The actual rates used were determined by an educated guess. It was hard to develop a more objective method to determine them.

Let us now turn to a more detailed description of the experimental design. For each subject in every round, income is determined by a random draw from an independent distribution. The distribution is chosen separately by each subject in each round. This is done by distinguishing two sets of envelopes. These are called the X-envelopes and the Y-envelopes. At the beginning of each round, subjects indicate on a form whether they want to choose an envelope from the X-set or the Y-set. They are then asked to pick one of the six envelopes in that set for that round. In this envelope, they find a note with their income for that round. They open the envelope privately, so the realized income remains unknown to the experimenter. Note that draws are independent: there are 6 envelopes of each type for each subject in every round. The subjects know the distributions of incomes in the two sets. The distinct income values in the
two sets are given in table 4.2.

Table 4.2: Income distributions

<table>
<thead>
<tr>
<th></th>
<th>Average</th>
<th>Standard deviation</th>
<th>Audit probability</th>
</tr>
</thead>
<tbody>
<tr>
<td>X-set</td>
<td>700</td>
<td>216.03</td>
<td>1</td>
</tr>
<tr>
<td>Y-set</td>
<td>400</td>
<td>322.49</td>
<td>0, 1/6 or 1/2</td>
</tr>
</tbody>
</table>

Note that the X-set has a higher average income and a lower standard deviation. The distribution of X-values stochastically dominates the distribution of Y values. Hence, X represents the risk neutral and risk averse choice, whereas a choice of Y is an indication of risk loving behavior. In the first three rounds, the choice between X and Y is the only choice subjects have to make. The number of X- and Y-choices in these rounds provides an indication of a subject's attitude towards risk.

In rounds 5-8 subjects have to report their income and pay taxes. Here, a second difference between X and Y is added. Subjects are informed that X envelopes will always be audited. Y envelopes are audited with a probability of either 0, 1/6 or 1/2, depending on the treatment (see table 4.1). This distinction represents a difference between registered (X) and unregistered (Y) income. Subjects always report their own income, but they know they will certainly be audited in case of an X-choice. If they choose Y, a die is thrown (independently per subject and round) to determine whether or not an audit will take place.76

Subjects have to pay 25% of the reported income as a tax. In case of an audit, they have to pay 25% of their actual income. As mentioned above, the aggregate tax proceeds from a group of 4 are divided equally across the group members (except in the two sessions without public goods, see table 4.1). This is done after completion of all 8 rounds. If an audit reveals that a subject has underreported income, a fine of 25% of the real income is imposed. The proceeds of the fine are not added to the public good. Note that the fine is not dependent on the level of underreporting. Therefore, if a subject decides to evade taxes, expected earnings are maximized by reporting the minimum possible income (i.e., 0). If an audit revealed that a subject has overreported income, no fine is imposed.

---

76 To avoid that the throw of a die reveals a subject's choice, we also throw a die in case an X-envelope is chosen. In this case, an audit takes place irrespective of the outcome.
Chapter 4

Summarizing, rounds 4-8 proceed as follows. First, subjects choose a source of income: 'registered' (X) or 'unregistered' (Y). Then, a random draw takes place to determine the realization of the income. Next, subjects report their income. It is then determined whether or not an audit will take place. If there is no audit, the tax is determined on the basis of the reported income. If an audit takes place and no underreporting is observed, actual income determines the tax and no fine is administered. If underreporting is observed, tax and fine are determined by the actual income. If overreporting is observed, actual income determines the tax and no fine is administered.

At the end of a session we first determined the total tax revenue and public good payoff per group. Then subjects were called privately, so their payoff could be determined. They gave us the 8 envelopes with the actual incomes in each of them. Note that for the cases without audit, this is the first time we observed the actual income. At this stage, however, it could not be determined how much they had declared. Hence, we could still not observe whether or not they had evaded taxes. Then, the earnings were determined as the sum of realized incomes plus the public good payoff and minus the taxes and fines paid.

4.4. Results

The presentation of the results is split into three parts. After an overview of the choice of income type (X or Y), we will present the general results on tax compliance. This is followed by a more detailed regression analysis of both decisions.

4.4.1. Choice of income

The choice of income is represented by the fraction of subjects choosing income type Y. Recall that there is a difference between the first three rounds of the experiment and the last five. In the last five rounds, the choice may be influenced by the fact that X envelopes are always audited whereas the Y envelopes are audited with a probability smaller than 1. In these rounds, subjects might choose income Y because it opens the possibility of tax evasion. In addition, with the exception of one treatment, a public good is provided in rounds 4-8. This also might be a reason to switch from one type of income to another. When the audit probability (p) was 1/2 and a public good was provided, Y was chosen 42.6% of the cases in the first three rounds and 58.7% of the time in rounds 4-8. This difference is significant at the
1%-level (paired sample t=6.10). When the probability was 1/2 and no public good was provided, Y was chosen 33% of the time in rounds 1-3 and 45% of the time in rounds 4-8. This difference is significant at the 5%-level (paired sample t=2.57). When the audit probability was 1/6, Y was chosen 31.9% of the time in the first three rounds and 70.0% of the time in rounds 4-8. This difference is statistically significant at the 1%-level (paired sample t=6.21). Finally, when there was no audit for Y envelopes (p=0), Y was chosen 36% of the time in rounds 1-3 and 48% of the time in rounds 4-8. This difference is not statistically significant (paired sample t=1.57).

Overall, it can be concluded that the possibility of tax evasion significantly increases the frequency of choosing unregistered income. Table 4.3 displays the Y choices in more detail. The choices are split up for the first three rounds (1-3) and the last five rounds (4-8).

Table 4.3: Fraction of Y-choices per subject pool in each country

<table>
<thead>
<tr>
<th>Subject pools per country</th>
<th>Choice of Y</th>
<th># sessions/ # subjects</th>
<th>choice 1-3</th>
<th>choice 4-8</th>
</tr>
</thead>
<tbody>
<tr>
<td>AL - US (prob.=1/6)</td>
<td>2/24</td>
<td>0.47</td>
<td>0.61*</td>
<td></td>
</tr>
<tr>
<td>AL - US (prob.=0)</td>
<td>2/24</td>
<td>0.36</td>
<td>0.48</td>
<td></td>
</tr>
<tr>
<td>AL - US (no public good, prob. = 1/2)</td>
<td>2/24</td>
<td>0.33</td>
<td>0.45*</td>
<td></td>
</tr>
<tr>
<td>AL - US</td>
<td>2/24</td>
<td>0.44</td>
<td>0.62*</td>
<td></td>
</tr>
<tr>
<td>AL - HS</td>
<td>2/24</td>
<td>0.46</td>
<td>0.63*</td>
<td></td>
</tr>
<tr>
<td>AL - HT</td>
<td>2/24</td>
<td>0.43</td>
<td>0.56**</td>
<td></td>
</tr>
<tr>
<td>AL - UT</td>
<td>2/24</td>
<td>0.56</td>
<td>0.63</td>
<td></td>
</tr>
<tr>
<td>AL - UP</td>
<td>2/24</td>
<td>0.39</td>
<td>0.43</td>
<td></td>
</tr>
<tr>
<td>NL - US (prob.=1/6)</td>
<td>2/24</td>
<td>0.17</td>
<td>0.79*</td>
<td></td>
</tr>
<tr>
<td>NL - US</td>
<td>2/24</td>
<td>0.22</td>
<td>0.68*</td>
<td></td>
</tr>
<tr>
<td>NL - HS</td>
<td>2/24</td>
<td>0.46</td>
<td>0.57</td>
<td></td>
</tr>
<tr>
<td>NL - HT</td>
<td>1/8</td>
<td>0.25</td>
<td>0.58**</td>
<td></td>
</tr>
<tr>
<td>NL - UP</td>
<td>1/12</td>
<td>0.58</td>
<td>0.62</td>
<td></td>
</tr>
<tr>
<td>ALBANIA (p=1/2, with public good)</td>
<td>5/120</td>
<td>0.46</td>
<td>0.57*</td>
<td></td>
</tr>
<tr>
<td>NETHERLANDS (p=1/2)</td>
<td>4/68</td>
<td>0.37</td>
<td>0.62*</td>
<td></td>
</tr>
<tr>
<td>HS + US (both countries &amp; p=1/2, with public good)</td>
<td>4/96</td>
<td>0.40</td>
<td>0.62*</td>
<td></td>
</tr>
<tr>
<td>HT + UT + UP (both countries &amp; p=1/2)</td>
<td>5/92</td>
<td>0.46</td>
<td>0.55*</td>
<td></td>
</tr>
</tbody>
</table>

(*) indicates that the difference between choice 4-8 and choice 1-3 is statistically significant at the 5% level.

(**) indicates that the difference between choice 4-8 and choice 1-3 is statistically significant at the 10% level.

Note: Numbers represent the fraction of Y-choices. The dark shaded rows indicate the sessions where the audit probability p=1/6. The light shaded rows represent the additional sessions described in the main text. All other rows represent sessions with p=1/2.
The following conclusions can be derived from table 4.3. First of all, the aggregate conclusion that more subjects choose an unregistered income (Y) when tax evasion is added as an option and a public good is introduced holds for all treatments. In all cases, choice 4-8 > choice 1-3. In 8 out of 13 treatments, this difference is statistically significant at the 10%-level or better.

Second, changing the audit probability does not effect the variables choice 1-3 or choice 4-8 significantly. Compare the fractions for US (p = 1/6) with those for US (p = 1/2). When testing for each country and variable separately (independent samples t-tests), none of the differences are statistically significant at the 10%-level, even though Y is chosen 11%-points more often by Dutch students in rounds 4-8 when p=1/6 than when p=1/2. When comparing the fraction of Y choices by Albanian students without audit (p=0) with the fractions for p=1/6 or p=1/2, it appears once again that none of the differences is statistically significant at the 10% level. The fact that the choice in rounds 1-3 is not affected by the audit probability is not surprising. In these rounds, subjects do not even know that there will be taxes, audits and public goods later on. Hence, differences in audit probability cannot affect behavior. The choices in rounds 4-8 may vary with audit probability if subjects' choice of income type is dependent on the probability of getting caught at evading taxes. However, this does not appear in the data.

Third, consider the effect of the provision of a public good. Because the sessions without a public good were only run for Albanian students with an audit probability of 1/2, their choices were compared with those in the AL-US (p=1/2) sessions with public good. The public good does not affect the choices in rounds 1-3 significantly. As above, this was expected, because subjects knew nothing about the situation in later rounds. However, the choices in rounds 4-8 are affected by the provision of a public good. Subjects in sessions with a public good choose Y significantly more often than subjects in sessions without a public good (44% and 33% of the time, respectively; independent samples t=2.02).

Fourth, when aggregating per country (for p=1/2, with public good), the increase in choosing Y after round 3 is statistically significant within each country at the 1%-level. When testing for differences between the two countries (independent sample t-tests), neither choice 1-3 nor choice 4-8 showed differences that were statistically significant at the 10%-level.
Finally, we can aggregate the data according to subjects' labor market position. Someone is denoted as being on the labor market if she/he has a job. The groups HS and US are (aside from small part time jobs) therefore not on the labor market -and consequently have little, if any, experience with paying direct taxes- while the other groups are. The last two rows of table 4.3 show the results of this aggregation (for p=1/2, with public good). Again, the increase in Y-choices after round 4 is statistically significant at the 1%-level for the aggregated observations in both groups (paired sample t=5.65 for pupils/students; t=2.83 for teachers/personnel). The differences between pupils/students on one hand and teachers/personnel on the other are not statistically significant at the 10%-level for choice 1-3 or choice 4-8.

Summarizing, our evidence shows a clear and significant shift towards unregistered income (Y), when tax evasion is made possible. Furthermore, there are no significant differences across countries or subject pools in the extent to which they choose Y. The choice of a Y envelope is not affected by audit probability either. However, the choice of unregistered income is significantly lower when there is no public good than when there is. Let us now turn to the tax compliance itself.

4.4.2. Tax compliance

Section 4.5 presents a brief theoretical discussion of the decision at hand. At this stage, it suffices to note that in our design a rational subject only interested in private earnings will either evade taxes completely (i.e., report an income of 0) or report income truthfully. The choice depends on her/his risk attitude. Moreover, it does not make sense for a selfish, rational subject to evade taxes after choosing an X envelope, given that the audit probability is equal to 1. Therefore, we start by simply analyzing the extent to which income was reported truthfully in various treatments.

First consider the subjects that chose X. Table 4.4 presents an overview of the number of X choices and the number of times that the income reported after choosing X was less than (<), equal to (=) or more than (>) the actual income. The numbers are aggregated across rounds 4-8.
Table 4.4: Distribution of income declaration after X choice

<table>
<thead>
<tr>
<th></th>
<th>P=1/2, public good</th>
<th>P=1/6, public good</th>
<th>P=0, public good</th>
<th>P=1/2, no public good</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>&lt;</td>
<td>=</td>
<td>&gt;</td>
<td>&lt;</td>
</tr>
<tr>
<td>Albania</td>
<td>1</td>
<td>251</td>
<td>6</td>
<td>0</td>
</tr>
<tr>
<td>Netherlands</td>
<td>1</td>
<td>126</td>
<td>3</td>
<td>0</td>
</tr>
<tr>
<td>Total</td>
<td>2</td>
<td>377</td>
<td>9</td>
<td>0</td>
</tr>
</tbody>
</table>

Note: The audit probabilities in the table refer to Y choices

This table shows, for example, that there were 388 X-choices in rounds 4-8 when the audit probability was 1/2 and there was a public good. After choosing X, in two cases less than the actual income was declared and in nine more was declared. In all other (377) cases subjects reported their actual income. In aggregate, only 3 out of 588 X decisions (0.5%) were followed by the subject declaring less than the actual income. Declaring more than the actual income was observed slightly more often: 18 times (3.2%). This result does not appear to be different in Albania than in the Netherlands. A possible reason for declaring more than the actual income will be discussed in section 4.5.

Next, consider the Y-choices. The distribution of decisions to declare less than (<), equal (=) or more than (>) the actual income is shown in table 4.5, separately for each subject pool. In case where less than the real income was declared, the table also reports the number of times that this declared income was non-zero. In addition, this table provides information about the average fraction of rounds that subjects evaded taxes. To calculate this, per individual we first determined the fraction of times they underreported income, treating overreporting as missing values. Then, we calculated the mean of these fractions across individuals in a subject pool. This mean is reported in the table in the column ‘evasion’.
Table 4.5: Distribution of income declaration

<table>
<thead>
<tr>
<th></th>
<th>P=1/2, public good</th>
<th></th>
<th></th>
<th></th>
<th>Evasion*</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>&lt;</td>
<td>=</td>
<td>&gt;</td>
<td></td>
<td></td>
</tr>
<tr>
<td>AL-HS</td>
<td>22 (13)</td>
<td>50</td>
<td>3</td>
<td></td>
<td>0.19</td>
</tr>
<tr>
<td>AL-US</td>
<td>12 (9)</td>
<td>50</td>
<td>12</td>
<td></td>
<td>0.10</td>
</tr>
<tr>
<td>AL-HT</td>
<td>5 (5)</td>
<td>54</td>
<td>8</td>
<td></td>
<td>0.05</td>
</tr>
<tr>
<td>AL-UP</td>
<td>2 (0)</td>
<td>45</td>
<td>4</td>
<td></td>
<td>0.02</td>
</tr>
<tr>
<td>AL-UT</td>
<td>10 (2)</td>
<td>64</td>
<td>1</td>
<td></td>
<td>0.08</td>
</tr>
<tr>
<td>NL-HS</td>
<td>27 (6)</td>
<td>39</td>
<td>2</td>
<td></td>
<td>0.24</td>
</tr>
<tr>
<td>NL-US</td>
<td>46 (3)</td>
<td>35</td>
<td>1</td>
<td></td>
<td>0.39</td>
</tr>
<tr>
<td>NL-HT</td>
<td>1 (1)</td>
<td>18</td>
<td>4</td>
<td></td>
<td>0.03</td>
</tr>
<tr>
<td>NL-UP</td>
<td>5 (1)</td>
<td>30</td>
<td>2</td>
<td></td>
<td>0.09</td>
</tr>
<tr>
<td>Total</td>
<td>130 (40)</td>
<td>385</td>
<td>37</td>
<td></td>
<td>0.14</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>P=1/6, public good</th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>&lt;</td>
<td>=</td>
<td>&gt;</td>
<td></td>
<td></td>
</tr>
<tr>
<td>AL-US</td>
<td>11 (5)</td>
<td>45</td>
<td>17</td>
<td></td>
<td>0.11</td>
</tr>
<tr>
<td>NL-US</td>
<td>65 (10)</td>
<td>20</td>
<td>10</td>
<td></td>
<td>0.58</td>
</tr>
<tr>
<td>Total</td>
<td>76 (15)</td>
<td>65</td>
<td>27</td>
<td></td>
<td>0.34</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>P=0, public good</th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>&lt;</td>
<td>=</td>
<td>&gt;</td>
<td></td>
<td></td>
</tr>
<tr>
<td>AL-US</td>
<td>15 (3)</td>
<td>34</td>
<td>9</td>
<td></td>
<td>0.13</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>P=1/2, no public good</th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>&lt;</td>
<td>=</td>
<td>&gt;</td>
<td></td>
<td></td>
</tr>
<tr>
<td>AL-US</td>
<td>8 (2)</td>
<td>40</td>
<td>6</td>
<td></td>
<td>0.08</td>
</tr>
</tbody>
</table>

**Note:** Numbers represent the number of times the reported income was less than (<) equal to (=) or more than (>) the actual income, given that Y was chosen. The numbers in parentheses report the number of times that underreporting did not involve reporting income equal to 0.

(*) This column reports the average fraction of times income was underreported (including X-choices). The way in which this was calculated is described in the main text.

The following conclusions can be drawn from table 4.5. First of all, in aggregate, the number of times income is underreported (229) is much higher than the number of times subjects report too much (79).

Second, however, there are various groups where the number of times income was underreported is similar to the number of times it was overreported. For the groups AL-HS, AL-UT, NL-HS, NL-US and AL-US when p=0, tax evasion appears to be more systematic than overreporting of income. For the other groups, the differences in the number of times income is under- or overreported are small. Not shown in the table is the observation that overreporting was quite constant across rounds. For p=1/2 with public good, for example, the 37 times that too much income was reported occurred 9, 4, 8, 11, and 5 times in rounds 4-8, respectively. Contrary to what we observed for X choices (table 4.4), overreporting appears to be more prevalent in Albania than in the Netherlands. For p=1/2, with public good, Albanians
reported more than the actual income in 28 out of 342 decisions (8.2%) and the Dutch did so in 9 out of 210 decisions (4.3%). For \( p=1/6 \), Albanians overreported on 17 out of 73 occasions (23.3%) and the Dutch did so 10 out of 95 times (10.5%). This is tested formally using \( \chi^2 \) tests per round. Considering only the Y-choices where income was not reported honestly, we compared the distribution of choices across under- and overreporting per country. In both treatments where a comparison between countries is possible (\( p=1/2 \) and \( p=1/6 \)), the extent of overreporting relative to underreporting is higher in Albania than in the Netherlands. In 7 of the 10 tests (two audit probabilities, 5 rounds) this effect is statistically significant at the 10%-level, while in 5 of these tests it is significant at the 5%-level. However, at first sight, reporting too much income does not make sense. If an audit takes place, nothing is gained or lost compared to truthful reporting. If no audit takes place, more taxes are paid (based on reported income) than necessary. Note that the payoff is determined by the actual income, not by the reported income. This point will be further discussed in the following section.

Third, it occurs quite often that subjects decide to evade taxes but do not evade them completely. In 60 out of 229 cases (26.2%) subjects reported an income higher than zero but lower than their actual income. This is more common in Albania than in the Netherlands. For \( p=1/2 \) with public good, Albanians evaded taxes 51 times. In 29 cases (56.9%) however, they declared more than 0. Only 11 out of 79 cases (13.9%) with tax evasion in the Netherlands showed a reported income higher than 0. Similar numbers hold for \( p=1/6 \). As noted above, a selfish, rational individual evading taxes will report 0. This phenomenon will be also discussed in section 4.5.

To get a better grip on the issue of tax compliance in our experiment, let us now continue the analysis by considering the extent of non-compliance (in other words, the extent of underreporting) while treating the cases where too much income was reported as missing values. In addition, no distinction will be made with respect to the extent of underreporting: any reported income lower than the real income is considered as tax evasion. This analysis is based on the numbers presented in the columns ‘evasion’, in table 4.5.

Overall, the extent of underreporting is larger in the Netherlands (0.58 for \( p=1/6 \) and 0.24 for \( p=1/2 \) with public good) than in Albania (0.11 and 0.09, respectively). This country
difference is statistically significant at the 1%-level in both cases (independent sample \( t=5.93 \) for \( p=1/6 \) and \( t=3.58 \) for \( p=1/2 \)).

Second, for \( p=1/2 \) with public good, the data are aggregated according to labor market position (HS/US versus HT/UT/UP). The students have a higher level of non-compliance (0.23 for both countries together; 0.14 in Albania and 0.31 in the Netherlands) than the pool of teachers/personnel (0.05 in aggregate, 0.05 in Albania and 0.06 in the Netherlands). This difference is statistically different at the 1%-level (independent sample \( t=5.11 \)).

What follows from these first two results is that the aggregated group of Dutch students and pupils evades taxes more often than this group in Albania does. Both groups evade more often than non-students in either country. Non-students do not differ across the two countries. The difference between students/pupils in the two countries shows up when comparing the two nations in aggregate. On the other hand the difference when aggregating per country (0.24-0.09) is smaller than when aggregating per labor market position (0.23-0.05). Hence, subject pool differences within a country are at least as important as country differences.

Another conclusion about the extent of tax compliance is concerned with the effect of the audit probability. An increase in this probability from 0 to 1/6 to 1/2 only marginally decreases evasion amongst Albanian (university) students, from 0.13 to 0.11 to 0.10 (cf. table 4.5). None of the pair wise differences is statistically significant at the 10% level. Dutch students, however, have the tendency to underreport more when the audit probability is 1/6 (0.58) than when it is 1/2 (0.39). This difference is statistically significant at the 10%-level (independent sample \( t=1.80 \)).

Next, consider the effect of supplying a public good. Albanian students (\( p=1/2 \)) show a marginally higher level of evasion if a public good is provided (0.10) than if it is not (0.08). However, the difference is not statistically significant at the 10% level.

A final conclusion from table 4.5 is drawn from a comparison with table 4.3. It is possible to determine the increase in Y-choices after round 3 in table 4.3 and compare this to the extent of tax evasion in table 4.5. For example, there is a 14 %-point increase for AL-US (\( p=1/6 \)) and the average evasion in this group is 11%. Hence, not everyone switching to Y decides to evade taxes (note that this conclusion is much stronger when considering the fact

---

Note that this comparison could be only made for \( p=1/2 \), with public good, because subject pools were only varied for these parameters (cf. table 4.1).
that some of the 11% were already choosing Y). This conclusion holds for the majority of groups. In aggregate, the increase in Y choices after round 3 is 19 %-points and the average evasion is 17%.

Summarizing the results on tax compliance, we observe that almost all subjects truthfully report their income after choosing a registered income. The few that do not report honestly tend to report too much rather than too little. Tax evasion does occur after unregistered income is chosen. In aggregate, many more cases of underreporting than overreporting are observed in this case. However, two observations about the Albanian subjects compared to the Dutch are noteworthy: they tend to overreport their income more and they tend to report more than zero when evading taxes. Dutch subjects evade more than Albanians do and students/pupils evade more than others. Dutch students evade most and (contrary to their Albanian counterparts) are sensitive to the audit probability. Evasion by Albanian students is not significantly affected by the presence of a public good. Finally, for most groups, the extent of evasion is less than one would expect, given the increase in Y choices after round 3.

4.4.3. Regression analysis

In order to analyze subjects' decisions in more detail, we ran regressions. The first choice to be explained is the decision whether or not to evade taxes. Denote the variable describing the extent of evasion (the column 'evasion' in table 4.5) by \( E \). We estimated coefficients of two models where the dependent variable was derived from this variable. In model A, the variable to be explained is the decision whether or not to evade taxes at least once. In model B, the dependent variable is (a function of) the extent of tax evasion. For model A, we transformed \( E \) by defining a dummy variable \( E' \): \( E'=0 \), if \( E=0 \) and \( E'=1 \), otherwise. In other words, \( E' \) is equal to 1 if a subject underreported income at least once in rounds 4-8. Model A is a logit regression with dependent variable \( E' \). In model B, we transformed \( E \) to correct for the fact that its values are constrained to the interval \([0,1]\). In this model, the dependent variable \( \hat{E} = \ln((E+0.001)/(1-E+0.001)) \).\(^78\) Model B is a linear regression with \( \hat{E} \) as dependent variable.\(^79\)

\(^78\) The constant 0.001 was added to the numerator and denominator because \( \hat{E} \) would otherwise not be defined for \( E=0 \) or \( E=1 \). The results are not affected by adding a dummy variable to the independent variables selecting the cases where \( E=0 \) or 1.

\(^79\) Note that it is not reasonable to treat the individual choices in distinct rounds as independent observations. This is why we consider the extent of evasion as the variable to be explained in model B.
As independent variables, we used the following.

- **COUNTRY** is a dummy variable with value 0 for the Netherlands and 1 for Albania.
- **LABORMARKET** is a dummy variable equal to 0 for teachers/personnel and 1 for pupils/students.
- **AUDPROB0** is a dummy representing the audit probability with value 1 for sessions where \( p = 0 \) and 0 otherwise.
- **AUDPROB1/6** is a dummy representing the audit probability with value 1 for sessions where \( p = 1/6 \) and 0 otherwise.
- **NOPUBGOOD** is a dummy variable with value 1 for sessions without a public good and 0 otherwise.
- **CHOICE 13** represents the fraction of choices in the first three rounds where the subjects chose Y. As explained above, this is a measure of risk aversion with higher values indicating more risk seeking behavior.
- **EARN 48** is equal to total actual earnings (in francs, before taxes and public good) divided by 10000.\(^{80}\)
- **AGE** is the subject’s age divided by 100.
- **JOB** is a dummy variable with value 1 if the subject has a (part time) job and 0 otherwise. Note that for all teachers and personnel, the value is 1. Because these groups are represented by **LABORMARKET**, **JOB** distinguishes students/pupils with a part time job from those without.
- **GENDER** is a dummy variable equal to 0 for men and 1 for women.

Table 4.6 presents the regression results. The results are quite similar for the two models. First of all, the background variables **AGE** and **JOB** do not affect tax evasion in either model. Gender does appear to have an influence: women are less inclined to avoid taxes at least once (model A) than men. The difference between men and women just misses statistical significance in explaining the extent of tax evasion (model B). Second, the risk attitude of

\(^{80}\) Obviously, there is a possible endogeneity problem related to the inclusion of earn 4-8 as an independent variable. However, it is the only direct way to check the effect of earnings on evasion in a regression context. The conclusions for the other variables do not change if earn 4-8 is dropped from the regressions.
subjects, as measured by their choices in the first three rounds, affects the decision to evade at least once but not the extent of tax evasion. Risk seeking subjects are more likely to evade taxes at least once. Third, the earnings in round 4-8 do not affect the decision whether or not to evade at least once but do affect the extent of tax evasion. Subjects who earn more evade more. This is in line with results by Giese and Hoffman (1999) who also find a positive effect of income on tax evasion.

Table 4.6: Regression results

<table>
<thead>
<tr>
<th>Dependent variable</th>
<th>MODEL A - Logit E'</th>
<th>MODEL B - OLS E</th>
</tr>
</thead>
<tbody>
<tr>
<td>Constant</td>
<td>-1.61 (0.12)</td>
<td>-5.73 (0.00)*</td>
</tr>
<tr>
<td>COUNTRY</td>
<td>-1.34 (0.00)*</td>
<td>-1.64 (0.00)*</td>
</tr>
<tr>
<td>LABORMARKET</td>
<td>1.01 (0.06)**</td>
<td>1.06 (0.06)**</td>
</tr>
<tr>
<td>AUDPROB0</td>
<td>-0.16 (0.77)</td>
<td>-0.18 (0.76)</td>
</tr>
<tr>
<td>AUDPROB1/6</td>
<td>0.73 (0.06)**</td>
<td>1.02 (0.02)*</td>
</tr>
<tr>
<td>NOPUBGOOD</td>
<td>-0.52 (0.41)</td>
<td>-0.51 (0.39)</td>
</tr>
<tr>
<td>CHOICE 1-3</td>
<td>0.97 (0.04)*</td>
<td>0.65 (0.18)</td>
</tr>
<tr>
<td>EARN 4-8</td>
<td>3.91 (0.12)</td>
<td>4.63 (0.08)**</td>
</tr>
<tr>
<td>AGE</td>
<td>0.30 (0.96)</td>
<td>0.22 (0.91)</td>
</tr>
<tr>
<td>JOB</td>
<td>-0.02 (0.96)</td>
<td>0.11 (0.80)</td>
</tr>
<tr>
<td>GENDER</td>
<td>-0.53 (0.09)**</td>
<td>-0.54(0.11)</td>
</tr>
</tbody>
</table>

Note: numbers represent the regression coefficient; p-values in parentheses.
* indicates statistical significance at the 5% level; **indicates statistical significance at the 10% level.

Finally, the three variables related to our treatments: audit probability, country and labor market position all have significant effects on the decision whether or not to evade taxes. With an audit probability of 1/6 subjects are more inclined to evade taxes at least once and they also evade more (compared to an audit probability of 1/2). Surprisingly, we find no significant differences in tax evasion between p=0 and p=1/2. The Dutch evade more often than the Albanians do and pupils/students evade more than teachers/personnel.

Next, we analyze the decisions of income choice (X or Y) and tax evasion jointly. To do so, data are needed on both choices simultaneously. Therefore, we cannot aggregate across rounds, as before. On the other hand, we cannot treat an individual's choices in distinct rounds as independent observations. As a consequence, only these choices in round 4 are considered. The regression results provide evidence that the decisions about income source
and tax evasion are indeed made simultaneously in round 4. For a formal presentation and detailed results of this evidence, see Gërxhani and Schram (2001).

4.5. A discussion of the results

Let us start with a brief theoretical discussion of the decision making process. Assume that an individual is only interested in their own earnings. Recall that the setup of this experiment only makes tax evasion potentially profitable in case a Y envelope is chosen. In this case, as argued above, the fining system of this experiment makes reporting an income of 0 optimal once one has decided to evade taxes. Consider the case with public good. Given that a Y has been chosen with realization y, in case of an audit (probability p) the payoff from reporting 0 consists of y minus fine and tax plus 1/4 of the tax paid (from the public good). In case of no audit, the income is y. The expected income from evasion is therefore p(y-0.25y-0.25y+0.0625y) + (1-p)y = y-0.4375py. The income in case of honest reporting is equal to y-0.25y+0.0625y= 0.8125y. Hence, a risk neutral subject will evade taxes (by reporting y=0) when y-0.4375py > 0.8125y, or p < 0.43. Hence, given a choice of Y, a risk neutral subject will evade taxes in our p=0 and p=1/6 treatments but will report honestly in case p=1/2.

A priori, the expected income, x*, from choosing X is 433.33 and from Y, y*, is 400 (cf. table 4.2). Taking into account the possibility of tax evasion implies that a risk neutral subject should choose X for the p=1/2 sessions. For p=1/6, the expected payoff from choosing Y and evading taxes is y*-0.4375y*/6 = 370.83. The expected payoff from choosing X is equal to x*-0.25x*+0.0625x* = 0.8125x* = 325.08, which is lower than the expected payoff from choosing a Y envelope. For p=0 the expected payoff of choosing Y and evading is even higher: 400.

A similar calculation shows that a risk averse subject is indifferent between evading or not, in the case where there is no public good (and p=1/2). This subject will choose an X envelope.

Summarizing, when there are no taxes, a risk neutral subject will choose an X envelope. When there are taxes, a public good and an audit probability p=0 or p=1/6, this subject will choose a Y-envelope and report an income of 0, irrespective of the income received. In case of taxes and p=1/2, this subject will choose X. The same holds when there is no public good.

81 The consequences in case a subject attributes utility to the earnings of others are discussed below.
Of course, a risk seeking (averse) subject will tend to choose Y more (less) often in all cases. Hence, observed choices of Y when p=1/2 indicate risk seeking behavior. Recall that we observed many such choices.

This theoretical background can be used when interpreting our results. First note that this experiment uses the same basic design as most tax evasion experiments with respect to tax rates, audit probability and fine for cheating. Although varieties exist, these common elements serve as a basis of comparison for the results of experiments on tax evasion. Referring back to the experimental studies mentioned in section 4.2, we conclude the following:

(i) compliance decreases with the tax rate (Friedland et al., 1978 and Collins et al., 1990). This outcome has been sufficiently observed in previous experiments. Hence, there was no such treatment in this experiment. Instead, a fixed tax rate across sessions (25% of the reported income) was applied.

(ii) tax evasion decreases with the level of fines (Friedland et al., 1978 and Collins et al., 1990). Again, this was not one of the treatments in this experiment.

(iii) compliance increases with the audit probability (Friedland et al., 1978). The same is observed here.

(iv) tax evasion is lower when the proceeds are used to provide a public good (Alm et al., 1991). Our results do not support this conclusion.

(v) the decision about tax evasion is made jointly with the labor supply decision (Collins et al., 1990). We test this outcome through the provision of two sources of income, registered (X) versus unregistered (Y). The fact that there was a significant increase in Y-choices after tax evasion was possible provides support for this outcome. Even stronger support was obtained in the regression analysis for round 4, which showed that the decisions concerning the choice of envelope (i.e. income) and evasion were made simultaneously.

(vi) a large subset of people never cheat, because they appear to believe that cheating is wrong (Baldrey, 1986). Even though this hypothesis could not formally be tested, it is useful to note that even when p=0 or p=1/6, many subjects still did not evade taxes. In addition, recall that not every switch towards Y after round 3 was followed by tax evasion.

(vii) tax evasion increases with income (Giese and Hoffman, 1999). Our experimental results support this conclusion (cf. the results in tables 4.5 and 4.6).
(viii) women evade taxes less than men do (Giese and Hoffman, 1999). The same result is observed in our regression analysis.

Next, consider the three novel features in our experimental design: it compares tax evasion behavior in a developed country to that in a country in transition; it compares tax evasion across subject pools within countries; and it introduces a new laboratory environment consisting of two types of income. These make it possible to add some conclusions to the previous experimental findings on tax evasion.

1) When tax evasion is possible, there is a shift towards unregistered income. This shift does not differ significantly across subject pools. Many subjects consciously choose an income type that will allow them to evade taxes.

2) There are differences across groups and countries regarding the extent of non-compliance. Students and pupils evade more than other groups in both countries. The extent of evasion is larger in the Netherlands than in Albania. Country differences appear to be a bit smaller than differences across groups within each country.

3) There are patterns of choices that cannot be explained by the simple theoretical framework: (i) there is some overreporting of income in all subject pools; (ii) in many cases, tax evasion is incomplete, i.e., does not involve reporting an income of zero; (iii) there is a significant move towards Y-choices that cannot be completely attributed to the possibility of tax evasion; (iv) in sessions with a public good, subjects choose the unregistered income less often. The first two cases are more often observed in Albania than in the Netherlands. These observations might be explained by the role of the public good and non-selfish preferences.

The theoretical discussion presented above assumes that individuals attribute utility to personal earnings only. If they also attribute utility to the earnings of others (through altruism or considerations of fairness, for example) then they might want to give more (pay taxes) because it increases their utility. For a discussion of the literature in non-selfish preferences, see Bolton and Ockenfels (2000) or Schram (2000), for example. Overreporting, incomplete evasion and the switching to Y income can thus be interpreted as voluntary contributions to the public good. In case there is no public good, Y incomes become less attractive.82

82 Table 4.5 also shows that overreporting and evasion greater than 0 occur less in the no public good sessions than in the comparable sessions with public good.
An alternative explanation is that this can be simply attributed to an erratic behavior. However, there are a few conclusions from our data that seem to point against simple error as an explanation. First of all, all three phenomena are quite constant across rounds. Second, there are clear differences across groups. For example, Albanian university students overreport their income much more often than high school students. There is no reason why university students would err more often. If the public good explanation holds, we are left with the interesting conclusion that it is more prevalent in Albania than in the Netherlands. Future research will have to show the importance of and provide explanations for these results.

4.6. Conclusion

Research on the informal sector in general and tax evasion in particular has proven to be very important. The literature shows fine examples of both non-experimental and experimental research. The latter offers the opportunity to study individual behavior in a controlled environment. This is what the study reported in this chapter does. It provides a possibility to compare behavior across two countries, Albania and the Netherlands, as well as across various socio-economic groups within these countries, when institutions are controlled for. This makes it easier to understand differences in behavior across countries and cultures. In addition, this study puts the experimental environment closer to reality than previous experiments on tax evasion, by providing two sources of income: registered and unregistered. Our results show that this is a relevant distinction: the decisions about income source (i.e. labor supply) and tax evasion are made simultaneously.

At first sight (supported by the scarce empirical evidence that is available) Albania is characterized by a larger extent of tax evasion than the Netherlands. However, when differences across institutions are controlled for, this country difference was the other way around. When there is a possibility to evade taxes, all groups in both Albania and the Netherlands display an increasing preference for the unregistered income. Nevertheless, the level of evasion is lower in Albania than in the Netherlands. In addition, subject pool differences within a country are as important as country differences. Especially having a fulltime job or not appears to be important.
If it is true that tax evasion in Albania is higher than in the Netherlands, than our experiment shows that this is not a consequence of differences in attitudes or cultures. If it were, higher evasion would have been found in the Albanian sessions. Apparently, differences in tax institutions between the two countries —that do not appear in the experiments, of course— matter. Chapter 6 analyzes data from a large household survey in Tirana, which collects more information about tax evasion behavior. A combination of this evidence with the experimental results of this chapter provides further understanding of the individual motivations to evade taxes. We will return to this point in the concluding chapter of this thesis.

Finally, our cross subject pool design in both countries has shown that it is dangerous to draw conclusions about ‘cross cultural’ differences from experiments that use only one subject pool in each country. For example, if we had only compared high school teachers in both countries, the conclusion would have been that evasion is higher in Albania. This provides a caveat for the interpretation of experimental results that have been observed in distinct countries.
Appendix 4.1: English version of the Albanian and Dutch instructions used in the experiments

Instructions Part I

Introduction
Welcome to this experiment. Here you can earn money. The amount of money you earn will depend on your own decisions, on the decisions of other participants and on the outcome of a random event. The money will be paid to you personally and privately from the other participants. Your decisions are anonymous. They are not related to your name. In addition, there are no 'good' or 'bad' decisions in the experiment. This is not a test. We only want to study the decisions people make.

The currency we are going to operate with during the experiment is called experimental francs. When the experiment is finished, the amount in francs will be converted to guilders by dividing the total earnings by 50.\(^8^3\) Therefore, you can always calculate your earnings in guilders by dividing the amount in francs by 50.

The structure of the experiment
The 12 participants in this experiment are divided into 3 groups of 4. We will call them group 1, group 2 and group 3. The experiment consists of 8 rounds. You will stay in the same group in all rounds, but you will not know who else is in your group. Soon, we will let you know what group you are in, but this is not so important with respect to the decisions.

The decision made in each round is independent of the decisions made in other rounds. After the third round, we will change something in the decision making process. The exact information about this change will be explained to you when that moment comes.

The decision
In every round, you have to choose an envelope. There are two types of envelopes. We call them X-envelopes and Y-envelopes. There are six envelopes of each type. We will come by every desk one by one soon. Then you will have to inform us whether you want an X-envelope or a Y-envelope. We will take the six envelopes of the type you chose and you may pick one of the six. You will find an amount of money written inside the envelope. For the first three rounds, this amount represents your earnings per round.
Of course, you do not know beforehand the amount you will get. However, we will now tell you the amounts written in the envelopes.

\(^8^3\) This is the basic outline of all instructions used in the experiments. Other instructions, based on the various treatments described in the chapter, are available upon request.
\(^8^4\) In Albania, it is converted to Leks by dividing the total earnings by 3.
There are six different amounts in the X-envelopes. They are:

- 100 francs
- 300 francs
- 400 francs
- 500 francs
- 600 francs
- 700 francs

Hence, if you choose an X-envelope, you will get one of these amounts per round.

In the Y-envelopes, there are also six different amounts. They are:

- 0 francs
- 100 francs
- 300 francs
- 500 francs
- 700 francs
- 800 francs

Hence, if you choose a Y-envelope, you will get one of these amounts per round.

Presenting your decision

It is not allowed to talk during the experiment. This is the reason why you will inform us about your decision in writing. For this purpose, we have prepared a table, which will be distributed now.

<table>
<thead>
<tr>
<th>participant: ___</th>
<th>group ___</th>
</tr>
</thead>
<tbody>
<tr>
<td>Choice X or Y</td>
<td></td>
</tr>
<tr>
<td>Round 1</td>
<td>X</td>
</tr>
<tr>
<td>Round 2</td>
<td>X</td>
</tr>
<tr>
<td>Round 3</td>
<td>X</td>
</tr>
</tbody>
</table>

The procedure is easy for rounds 1, 2 and 3. When round 1 begins, we will ask you to circle a X or a Y. If you want an X-envelope, you should circle ‘X’, and if you want a Y-envelope, you should circle ‘Y’. When we come along, we will see what you have chosen. Then, we will let you choose from the six envelopes of that type. You can pick one of the envelopes. You may see for yourself how much you have earned. We do not need to know at that stage. You simply keep the envelope. When the experiment is finished, you take the envelope with you for the payment procedure. We will then give you the amount of money written inside the envelope in cash.
Now you will have a few of minutes to read these instructions again. If you have any questions, raise your hand, and one of us will come to you to answer that question. If you are finished, remain quiet until we start the experiment.

**Instructions Part II**

**The fund**

We will add something for the following rounds. Again, you have to choose an X- or a Y-envelope, which determines your earnings. Now, everybody must contribute part of her/his earnings to a fund. To be more precise, everybody should contribute 25% or a quarter of her/his earnings.

The money collected from a group (there are three groups in total) will then be distributed to all participants of the group. Hence, if one group contributes a total of 500 francs to the fund in a round, everybody of that group will receive 125 francs from the fund. This is independent of the amount contributed by each member of the group individually.

**Reporting your earnings**

Because we cannot see how much you earn per round (you pick an envelope and open it yourself), you must report your earning to us in each round. The amount reported will determine your contribution. It is up to you to report the real amount of your earning or some other amount.

**The audit**

There is a probability that we will audit your contribution. If you have chosen an X-envelope, we will always audit if you reported the correct amount. We will do this by looking inside your envelope. If you have picked a Y-envelope, we will throw a die to decide whether or not we will audit. If the die shows a 1 (one), 2 (two) or 3 (three), we will audit the amount reported. Otherwise, we will not. Hence, in the case of a Y-envelope, there is a probability of 1/2 that the reported amount will be audited. In the case of an X-envelope, the amount will always be audited.

There are two possibilities if we audit your reported amount.

- If the amount written inside the envelope is the same as the amount reported, we will write it down and you will contribute 25% of that amount to the fund.
- If the amount is not the same, we will write down your real earnings. You will have to pay 25% of these earnings to the fund. In addition, you have to pay 25% of your real earnings as a fine for declaring less. If you have reported more than your real earnings, you do not have to pay a fine. Collected fines are not deposited in the fund.
Again: if we do not audit your reported earnings, your contribution will be 25% of these reported earnings. The real earnings, that will be paid to you at the end, will be determined by the real amount in the envelope. You have to take this envelope with you when you go to be paid at the end of the experiment. If we do not audit you, we do not know during the experiment whether or not you have reported the correct amount.

**Your final earnings**

Your final earnings from round 4 to round 8 will be calculated when you come one by one at the end of the experiment and receive the money you have earned today.

1. We determine the real earnings (in francs) by adding up the amounts in your envelopes.
2. Then, we count the amount that your group deposited in the fund for the five rounds altogether. This total amount will be divided by 4 (the number of group's members) and the result (in francs) will be added to your earnings.
3. Further, we determine the total amount that you contributed to the fund. This amount will be subtracted from your earnings.
4. At the end, we determine whether you have to pay fines or not, and if yes, how much. This amount will also be subtracted from your earnings.
5. In this way, we know your total earnings in francs. These will be divided by 50 in order to determine your earnings in guilders. This is the amount you will receive from us.

**Registration**

We want your decisions to remain as anonymous as possible. This is the reason why the procedure is made such that the other participants cannot see from our behavior what you have chosen. Therefore, the procedure during rounds 4-8 is as follows.

First, we come along to give you a chance to choose an envelope. This is done in the same way as in the first three rounds. Then, you will have a chance to declare your earnings in the space provided in the table that will now be distributed.

<table>
<thead>
<tr>
<th>participant: __</th>
<th>group __</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Round 4</strong></td>
<td><strong>X Y</strong></td>
</tr>
<tr>
<td><strong>Choice X or Y</strong></td>
<td></td>
</tr>
</tbody>
</table>
The reported earnings can be the same as the amount in your envelope, but you can also report a different amount.

Then, we will come by again, and the following will happen.

(i) First, we throw a die.

(ii) Then, we see whether you have chosen an X- or a Y-envelope in that round.

(iii) If you have an X-envelope, we will audit the amount, independent of the number on the die. If you have a Y-envelope, we will audit the amount only if the die shows number 1 (one), 2 (two), or 3 (three).

(iv) If we do not audit the envelope, we will write in your table (in the space ‘real earnings’) ‘no audit’. If we do audit the envelope, we will write the real earnings in your table. This is your reported amount, if you reported the real amount.

(v) Then, we will write down your contribution to the fund of your group in the table. In case of an audit, this is 25% of the real amount. In case no audit took place, this is 25% of your reported amount.

(vi) Finally, we determine whether you should pay a fine. If you are not audited or if your reported amount appears to be equal to (or larger than) the real amount after being audited, we will write down a zero (0). If after being audited, your reported amount is smaller than the real amount, we will write down 25% of the real amount.

After we have been by everybody, we will start the next round. Note that during the experiment, we do not give any information about the total amount in the funds.

It may take a while to finish a round. It is important that you remain seated quietly and do not communicate with the other participants.

Now you can read these instructions again for a few minutes before we start the fourth round.

The questionnaire distributed at the end of the experiment

Your gender (please check): ... male ... female
Your age (please enter): ......
Your education (please check or enter): ....Economics ..........Other, e.g. ..................
How many hours per week do you work (please enter)? .........