TI 2002-098/1
Tinbergen Institute Discussion Paper

Tax Evasion and the Source of Income

Klarita Gërxhani[1]
Arthur Schram[1,2]

[1] *Faculty of Economics and Econometrics, University of Amsterdam,* [2] *Tinbergen Institute.*

Tinbergen Institute
The Tinbergen Institute is the institute for economic research of the Erasmus Universiteit Rotterdam, Universiteit van Amsterdam and Vrije Universiteit Amsterdam.

Tinbergen Institute Amsterdam
Keizersgracht 482
1017 EG Amsterdam
The Netherlands
Tel.: +31.(0)20.5513500
Fax: +31.(0)20.5513555

Tinbergen Institute Rotterdam
Burg. Oudlaan 50
3062 PA Rotterdam
The Netherlands
Tel.: +31.(0)10.4088900
Fax: +31.(0)10.4089031

Tax Evasion and the Source of Income: An Experimental Study in Albania and the Netherlands

by

Klarita Gërxhani*

and

Arthur Schram**

Abstract

A series of experiments in Albania and the Netherlands give us the opportunity to compare behavioral patterns related to tax evasion. Subjects have to decide between a random 'registered' income, the realization of which will be known to the experimenter for sure, and a random 'unregistered' income that will only be known to the experimenter with some (audit) probability. After the actual income has been determined, subjects have to report it and pay taxes accordingly. If they are audited, there is a fine for underreporting. This experiment was organized (separately) among high school pupils, high school teachers, university students, university teachers and university (non-academic) staff, in Albania and the Netherlands. The results show that (i) when tax evasion is possible, subject choose unregistered income more frequently; (ii) many subjects are willing to choose an income that allows for tax evasion but report their income honestly anyway; (iii) compliance increases with the audit probability; (iv) individual decisions to choose a type of income (registered vs. unregistered) and to evade taxes are made simultaneously; (v) Albanians evade taxes less than the Dutch do; (vi) pupils and students evade taxes more than teachers and personnel do; (vii) the differences across groups within a country are at least as large as the differences between the two countries. Finally, we argue that the distinct levels of tax evasion outside of the laboratory in the two countries are not attributable to different tax attitudes or cultures, but to different tax institutions and the way individuals have learned to deal with them.

JEL codes: C91, H26, O17, O57

Key words: Tax evasion, Cross-country studies, Experiments

October 2001

Comments by participants at the World Meetings of Economic Science Association (ESA) 1999 in Lake Tahoe, Nevada and 2000 in New York, and the European Public Choice Meetings 2000 in Sienna, Italy are gratefully acknowledged. We also thank Jordi Brandts, Jim Cox, Michael Ellman, Claude Montmarquette, Theo Offerman for useful suggestions.

*Corresponding author, Tinbergen Institute and University of Amsterdam
address:
Department of General Economics
Roetersstraat 11
1018 WB Amsterdam
the Netherlands
email: klarita@fee.uva.nl
CREED, Department of Economics and Econometrics, University of Amsterdam

**CREED, Department of Economics and Econometrics, University of Amsterdam

"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)

1. Introduction

The informal sector is attracting more and more attention. Though initially observed in the Third World (Hart 1971, 1973), it is now considered to be an important phenomenon everywhere. Numerous studies have emerged, analyzing the informal sector in underdeveloped and developed countries, as well as in countries in transition from socialist to market oriented economies (for recent surveys, see Schneider & Enste, 2000 and Gërxhani, 2002). The term 'informal sector' refers to a broad phenomenon, including a variety of activities related to the labor market: tax evasion, activities against government regulation, criminal activities and more. Research usually focuses on one of these in isolation, even though they are often directly or indirectly related to each other. In this study, we focus on one of the phenomena related to the informal sector: tax evasion.

Although the literature provides various definitions of tax evasion, there is a common idea underlying them. The term tax evasion simply refers to the underreporting of taxable money income with the intention to escape taxes.

There are two main reasons for focusing on this aspect of the informal sector in this paper. First of all, tax evasion is directly linked to large budget deficits and hence to lower investments in public goods. Besides being of general interest from an economic point of view, this means that the effects may differ significantly depending on the level of development of a country. This is one of the matters we are interested in. Second, studying tax evasion creates the opportunity to study the decision-making process related to the informal sector at the individual level. Though tax evasion has been studied extensively (Tanzi, 1982, Feige, 1989, and many others), little research has been undertaken to get a better understanding of the individual decision whether or not to evade taxes. Most of the existing literature has emphasized the 'goods' and 'bads' of tax evasion. Some studies provide an international comparison, based on empirical estimates (e.g., Schneider and Enste, 2000). Gërxhani (2002) argues that especially a distinction between developed countries, developing countries and nations in transition is important to understand the informal sector.

When studying tax evasion, it is important to distinguish between different types of income. In almost every country, one can make a distinction between income that is officially registered

and unregistered income. Registered income is generally observed in jobs within the public sector and in private sector employment with contracts. Typically, taxes are withheld from the regular wage payments when income is registered. Unregistered income can occur in cases where there is no job contract and in case of self-employment. In the latter case, income must be reported to the tax authorities in order to determine the income tax owed.

Some efficient distribution of labor across jobs with registered and unregistered income will exist. A well functioning labor market will help achieve this distribution. However, it is easier to evade taxes in case of unregistered income than when income is registered. This may provide individual incentives to choose a job where income is not registered. This incentive may distort the way the labor market works and cause inefficiencies in the distribution across types of jobs.[1] The experimental environment we will discuss below allows us to make the distinction between the two types of income. This allows us to study the choice of income type simultaneously with the decision whether or not to evade taxes. Hence, we can discuss the extent to which the evasion possibilities affect the choice of income type.[2]

Gërxhani's assertion that a distinction according to the level of a country's development is important for understanding tax evasion is the initial motivation for the study presented here. We study tax evasive behavior in two European countries, Albania (a country in transition) and the Netherlands (a developed country). The reason to focus on these two countries is as follows. First of all, Albania was the most isolated and the latest country in Eastern Europe to open itself up to democratic and economic changes. Therefore, when Albania entered the process of transition, it was struck by a crisis that was even deeper than in other transition countries. In addition, for a long time Albania was one of the poorest countries in Europe. This situation created suitable conditions for the informal sector to be the prevalent economy in Albania. It appeared everywhere. Initially starting as an emergency exit from the numerous problems of the formal sector, it later became an inevitable part of the Albanian society (Gërxhani, 2002). Given our wish to compare tax evasion in a country in transition to tax evasion in a developed country, Albania appears to be an appropriate choice for the former. The choice of the Netherlands as the developed country is based first of all on the observation that it is indeed a country with a very high standard of living and is (contrary to Albania) well developed and with stable economic and political institutions. Moreover, its size is

[1] Of course, there may be inefficiencies as a consequence of the income tax *per se* as well.
[2] Collins and Plumlee (1991) analyze a different, but related, issue. They present an experiment where subjects' income (to be reported to the tax authorities) depends on an effort made and observe more effort when tax evasion is possible.

comparable to Albania. As a practical matter, the facilities and experience of running experiments are readily available in the Netherlands. These two countries create a sharp contrast, which will allow a fertile basis for testing the difference in tax evasive behavior between a developed country and a country in transition.[3]

In our study, we apply the experimental method. Although it has sporadically been used to study tax evasion before, this only holds for developed countries. An important motivation is that it provides an opportunity to directly test existing hypotheses on individual tax evasion. Moreover, the laboratory provides an opportunity to obtain comparative results from two types of countries with respect to individuals' attitudes towards tax compliance, while controlling for the numerous differences between the countries. In this way, we hope to get a grip on the individual decision involved and a better understanding of the factors that cause distinct levels of tax evasion between a developed country and a country in transition.

More specifically, the experiments allow us to compare behavior across countries when *institutions* are controlled for. The choice to evade taxes may be affected by institutions, by culture, or by both. By controlling for differences in institutions, we collect information about the effect of culture. The existing empirical evidence seems to show that Albanians evade taxes more often than the Dutch do (cf. section 2). Our experiments will provide us with information as to whether Albanian culture is more open to tax evasion or whether the difference is more likely due to more primitive tax collection institutions.

When running experiments in different countries, one needs to be careful when drawing conclusions from observed differences. If one observes differences in behavior of one subject pool in country A and another in country B, it might be true that behavior in both countries differs. It might also be the case that behavior differs across subject pools in general. One needs a comparison of differences in behavior across countries to differences within countries before one can truly attribute differences across subject pools to country differences. In our setup we distinguish various subject pools within each of the two countries.

In summary, the contribution of this paper is threefold:

(1) we provide a comparative experimental study of tax evasion in a developed country and a country in transition (the Netherlands and Albania);

[3] Given Gërxhani's (2002) theory, it would be interesting to also study tax evasion in a developing country. Unfortunately, we did not have this opportunity. Nevertheless, the distinction between a developed and transition country remains of interest.

(2) we compare tax evasion across subject pools (i.e., socio-economic groups) within each country. This allows us to compare differences across countries to differences within a country;
(3) we present a new laboratory environment in which subjects can choose between two possible sources of income: 'registered' and 'unregistered'.

The organization of this paper is as follows. In Section 2, we discuss the existing evidence of tax evasion in Albania and the Netherlands and review the relevant literature. Section 3 presents our experimental design and conjectures about the treatment effects to be expected. Section 4 presents the experimental results, which are discussed in section 5. Section 6 concludes.

2. A brief review of the literature

There are three areas in the literature that are of interest for our study: empirical evidence on the informal sector in the Netherlands and Albania, studies on the informal sector using various tools, and cross-national experimental research. We will briefly discuss the literature on each of these.

Empirical evidence of tax evasion in the Netherlands and Albania

There are more studies about the Netherlands than about Albania. These typically do not refer to tax evasion per se. Instead, they use terms like the 'hidden' or 'shadow' economy when referring to the informal sector. 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 everywhere, there was an informal sector. Taxes could be evaded, for example, by not working in official employment or by selling privately grown agricultural products.[4] For the period of transition, existing evidence on tax evasion is not based on careful analysis either. Anecdotal evidence suggests high levels of tax evasion in current Albania,

[4] This evasion of taxes was less significant than in developed market economies, because personal income tax was less important.

however. 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 profits. On average, this underreporting constitutes 20% of their profits. According to a recent study of UNDP (2000), the informal economy is thriving, accounting for an estimated 50% of GDP and 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.

Tools used to study the informal sector

Previous studies on the Netherlands and Albania have applied the traditional empirical tools in this field, focusing on two main approaches for analyzing the informal sector: the direct and the indirect approach. The direct approach uses sample surveys to gather information about individuals' behavior, whereas the indirect approach derives conclusions about the informal sector from an analysis of macro-economic statistics.[5]

Both methods have serious shortcomings. The indirect approach does not give any information about the specifics of the informal sector (Thomas, 1992). In the direct method there is a potential respondent jeopardy[6] (Lee, 1985). In any survey, there is a problem of respondent's motivation to answer seriously, but this problem appears to be especially relevant in studies of the informal sector and tax evasion.

Recently, a third method is becoming more popular: laboratory experiments. Whereas other methods (direct and indirect) are used to analyze almost every aspect of the informal sector, the experimental method is mostly restricted to tax evasion. The first experimental study on tax evasion is a collaboration of social psychologists and economists (Friedland, Maital and Rutenberg, 1978). They study individuals' behavior when confronted with changes in tax rates, penalties and audit probabilities. This and the following experimental studies on tax evasion are closely related. The basic design of these experiments is described in a survey by Alm (1991). It is simple: participants declare taxes based on an income determined in the experiment. They are provided with information regarding tax rates, audit probabilities and fines for cheating

[5] There are several applications of the indirect approach, such as monetary methods and discrepancy methods. Hansson (1989), for example, applied the discrepancy method to estimate tax evasion as the difference between households' total expenditures on consumption, savings and taxes and total income reported on tax returns.
[6] This refers to the respondents' impression of being threatened by questions.

(underreporting income). Though there are varieties to this basic design, the results regarding tax compliance are quite robust. These results show that:

(i) tax evasion increases with the tax rate (Friedland et al., 1978);
(ii) tax evasion decreases with the level of fines (Friedland et al., 1978);
(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 decision on how much effort to put in income earning (Collins and Plumlee, 1991);
(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).

As mentioned above, the main advantage of laboratory experimentation is the control it provides. A disadvantage is that the laboratory setting might not be applicable to the real world setting it aims at studying. The extent of external validity depends crucially on the experimental design, of course. An appropriately designed experiment complements research that applies other methods (such as the direct and indirect methods of studying the informal sector). For further discussion of the role of experiments in studying tax evasion, see Webley et al., 1991.

Cross-national experimental studies

By conducting and analyzing the same experiment in two countries, this paper is related to a variety of cross-country studies that have recently been undertaken (see Brandts, Saijo and Schram, 2000 for references). In many cases, differences in behavior are observed across countries (e.g., Roth, Prasnikar, Okuwo-Fujiwara and Zamir, 1991; Saijo and Nakamura, 1995). Given that the experiments are conducted in exactly the same way in every country, this may cause some concern to economists because economic analysis is traditionally based on the premise that economic behavior is guided everywhere by common principles.

In Brandts et al. (2000), however, a public good experiment is conducted and no differences in behavior are found between subjects in Japan, the Netherlands, Spain and the Unites States. The question is then, why differences are observed in some studies and not in others. This might have to do with the subject pools used. As far as we know, no one has compared differences between

subject pools within the same country to differences between subject pools across countries. Note that it is implicitly assumed that within country differences are relatively small if cross-country differences are attributed to cultural differences (as is often done). In this paper, we do include different subject pools in each country to get a more complete picture of behavior across countries.

3. Experimental design and expected treatment effects

General features of the procedures and design

The experiments are run manually in order to make it possible to organize them at various locations without having to arrange computer facilities. In each session there are 12 participants, divided into 3 groups of 4.[7] A session consists of 8 independent rounds. The experiment itself generally lasts less than one hour. Calculating payoffs takes some time, however, so the last subject usually leaves about 75 minutes after the start of the experiment. All experiments reported here took place in 1999-2000.

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. Besides allowing participants to get acquainted with the decision at hand, these three rounds allow us to measure the individuals' risk attitudes. We will use this to test the influence of risk attitude on the choice of income and tax evasion. In rounds 4-8 there is a second part of each round 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.

Our design distinguishes three treatments where we varied subject pools, audit probability and what we did with tax proceeds. We will discuss each in turn. A summary of our sessions is presented in table 1, below.

The first important treatment in our experimental design is the distinction of subject pools. For reasons discussed above, these varied along two dimensions: country (the Netherlands, NL, and Albania, AL) and socio-economic category. For the latter we distinguished the following groups: (i) high school pupils (HS); (ii) university students (US); (iii) high school teachers (HT); (iv) university non-academic personnel (UP); and (v) university teachers (UT).[8] Sessions in

[7] In one session (high school teachers in the Netherlands), there were only eight participants (two groups of four).

[8] 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. One could argue that the subject pools

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.[9]

Second, we varied the audit probability for some subject pools. This allows us to test previous results on the effect of audit probability (Friedland et al., 1978 report decreasing tax evasion as the probability rises). In addition, it allows us to investigate whether the effect of the probability of being 'caught' is different in the Netherlands than in Albania. For all subject pools, we used a (high) probability of 1/2 of auditing the reported income (as described below). For the students in each country (NL-US and AL-US), we ran additional sessions where the audit probability was equal to 1/6. It will be shown below, that a risk-neutral subject will evade taxes when the audit probability is equal to 1/6 but not when it is 1/2. For the students in Albania (AL-US), we also ran sessions with no audits (probability equal to 0). We decided to do so after analyzing the results of the original (1/2 and 1/6) sessions, because this group appeared to be insensitive to the audit probability. Sessions without audit allow us to test the robustness of this result.

As a final treatment, we ran sessions with and without a public good. In our first series of experiments, tax proceeds were aggregated within a group of 4 and divided equally among the group members. In this setup, taxes may be seen as contributions to a public good with mpcr = 1 (Isaac et al., 1984). After observing a relatively high level of compliance in the Albanian sessions, we considered the possibility that Albanian participants might be complying to taxes because they wish to contribute to the public good. In order to test this, we ran two sessions for AL-US where no public good was provided: tax proceeds were not returned to the participants in any way.

In all sessions, groups remain constant across rounds and know how taxes and (if applicable) 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.

Note from table 1 that the average earnings in experimental francs (the currency used in the experiment) across subject pools did not differ much. An exception is that earnings were higher in

---

used are quite similar to each other in many ways. Our results will show significant differences across some of these groups, however. With other groups these differences might even grow.

[9] 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.

Table 1: Experimental sessions

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

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.

sessions with a low audit probability. This is a consequence of fewer fines being administered. 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 we varied the exchange rate from francs to local currency in order to account for differences in purchasing power, both across countries and across pools within a country.

Design issues related to cross-country experiments

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

The term *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 authors of this paper ran all sessions. Therefore, in principle, this kind of experimenter effects is not expected. On the other hand, a priori we were afraid that Albanian participants might be excessively impressed by the presence of a western professor 'handing out money' (one of the authors is Dutch). To avoid this type of experimenter effect, the other author, an Albanian national working in the Netherlands, addressed the participants in Albania. The Dutch experimenter stayed in the background. In addition, the experiments are 'double blind' as long as no audit takes place. When there is no audit, the experimenters have no way of knowing whether or not the subjects have truthfully declared their

income. When there is an audit, the experimenters do, however, discover whether or not the participants have reported truthfully.

Second, to control for unwanted *language effects* the instructions for the experiment were initially written in English, and then translated into Albanian and Dutch. In addition, one of the authors speaks both Dutch and Albanian and is able to check for differences that might have occurred because of the translation.[10]

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. We varied the exchange rate from experimental francs to the local currency as described in table 1 in order to maintain sufficient and comparable saliency across groups. The actual rates used were determined by an educated guess. It is hard to conceive more objective method to determine them.

Detailed description of the 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 written 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 and the others. Note that draws are independent: we prepared 6 envelopes of each type for each subject in every round. Subjects know the distributions of incomes in the two sets. The distinct income values in the two sets are given in table 2.

Table 2: Income distributions

							Average	**Standard deviation**	**Audit probability**
X-set	100	300	400	500	600	700	433.33	216.03	1
Y-set	0	100	300	500	700	800	400	322.49	0, 1/6 or 1/2

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

[10] See Appendix 1 for the English version of the instructions.

three rounds, the choice between X and Y is the only decision 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 4-8 subjects have to report their income and pay taxes. Here, we add a second difference between X and Y. Subjects are informed that X envelopes will always be audited. Instead, Y envelopes are audited with a probability of either 0, 1/6 or 1/2, depending on the treatment (see table 1). Subjects always report their income, and 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.[11] This distinction between the two types of envelopes represents a difference between registered (X) and unregistered (Y) income.

Subjects have to pay 25% of the reported income as a tax. In case of an audit, the tax consists of 25% of their actual income. In the public good sessions, the aggregate tax proceeds from a group of 4 are divided equally across the group members. This is done after completion of all 8 rounds. In the sessions without a public good, the tax proceeds are not returned to the participants.

If an audit reveals that a subject has underreported income, a fine of 25% of the actual income is imposed (on top of the tax payment). 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.c., 0). If an audit reveals that a subject has overreported income, no fine is imposed.

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 honest reporting is observed, actual income (which in this case equals the reported 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 determine the total tax revenue and public good payoff per group. Then subjects are called privately, so we can determine their payoff. They give 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 observe the actual income. At this stage, however, we cannot determine how much

[11] 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.

they have declared. Hence, we can still not observe whether or not they have evaded taxes. We then determine the earnings as the sum of realized incomes plus the public good payoff and minus the taxes and fines paid.

Theoretical analysis and expected treatment effects

Before presenting the results, we provide a brief theoretical discussion of the subjects' decision and conjectures about the treatment effects that we expect. Assume that an individual is only interested in the own earnings.[12] Recall that our setup only makes tax evasion potentially profitable in case a Y envelope is chosen. In this case, as argued above, our fining system 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. Taking account of an audit with probability p, 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 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, after a Y choice, a risk neutral subject is indifferent between evading or not in the case where there is no public good (and p=1/2). The expected payoff is $0,75y^*=300$. This subject will choose an X envelope (with expected payoff $0,75x^*=325$.

Summarizing, in case there are no taxes, a risk neutral subject will choose an X envelope. In case of 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. Of course, a risk

[12] The consequences in case a subject attributes utility to the earnings of others are discussed in section 5.

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.

Given our experimental design, we can conjecture on the treatment effects to be expected. First, our subject pool treatment was set up to test the distinct levels of tax evasion (and their causes) in Albania and the Netherlands. If one assumes that distinct attitudes or cultures cause this difference, our first conjecture is that we will observe higher levels of evasion in any group in Albania than in any Dutch subject pool. If we do not observe these differences, our results will allow us to question this assumption. As for distinct subject pools within a country, the implicit assumption in many cross country experiments is that the differences within a country are smaller than those between countries. We will take this as our second conjecture.

In our second treatment, we vary the audit probability. Assuming a similar distribution of risk attitudes across countries, this yields our third conjecture that evasion will decrease with this probability in both countries. In addition, we can test the influence of risk attitude on evasion because choices in the first three rounds provide a measure of this attitude.

Finally, for Albanian students, we ran sessions without a public good. As mentioned above this was done to test the idea that Albanians do not evade because they want to voluntarily contribute to the public good. This gives our fourth and final conjecture that evasion will be higher in the sessions without a public good.

4. Results

The presentation of the results is split in 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. A more general discussion of our results and the conjectures on treatment effects will follow in section 5.

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 too might be a reason to switch from one type of income to another.

Table 3 displays the Y choices for each subject pool. The choices are split up for the first three rounds (choice 1-3) and the last 5 rounds (choice 4-8).

Table 3: Fraction of Y-choices per subject pool

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

(∗) 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: See table 1 for an overview of the treatments distinguished per subject pool. 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 about the choice between the two types of income can be derived from table 3. These conclusions are presented as observations 1-5.

Observation 1: *More subjects choose an unregistered income (Y) when tax evasion is added as an option.*

Support: In all cases, the fraction of Y choices in rounds 4-8 > the fraction of Y choices in rounds 1-3. In 8 out of 13 subject pools, this difference is statistically significant at the 10%-level or better.

Aggregating per treatment gives the following results. 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). 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).

When aggregating per country (p=1/2, with public good), the increase in Y choices is significant (for Albania: an increase from 46% to 57%, paired sample t=4.21; for the Netherlands: an increase from 37% to 62%, paired sample t=4.52).

Finally, we can aggregate the data according to subjects' labor market position. We denote someone as being on the labor market if they have a job. The groups HS and US are therefore not on the labor market (aside from small part time jobs).[13] As a consequence, they have little, if any, experience with paying direct taxes, while the other groups are. The increase in Y-choices after round 3 is statistically significant at the 1%-level for the aggregated observations in each group (paired sample t=5.65 for pupils/students; t=2.83 for teachers/personnel).

Observation 2: *There are no differences across countries in the fraction of Y choices.*
Support: When testing for differences between the two countries at the aggregate level (p=1/2, with public good), neither choice 1-3 (independent sample t=1.62) nor choice 4-8 (independent sample t=0.97) showed differences that were statistically significant at the 10%-level.

Observation 3: *The fraction of Y choices is not affected by labor market position.*
Support: The last two rows of table 3 show the results of this aggregation (for p=1/2, with public good). The differences between pupils/students on one hand and teachers/personnel on the other are not statistically significant at the 10%-level for either choice 1-3 (independent sample t=1.31) or choice 4-8 (independent sample t=1.64).

Observation 4: *Changing the audit probability does not effect the fraction of Y choices in rounds 4-8 significantly.*
Support: 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

[13] This is supported empirically by the answers to the questionnaire distributed at the end of the experiment.

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, we once again find that none of the differences is statistically significant at the 10% level.

Observation 5: *There are more Y choices in rounds 4-8 when a public good is provided than when no public good is provided.* [14]

Support: Because the sessions without a public good were only run for Albanian students with an audit probability of 1/2, we compare their choices with those in the AL-US (p=1/2) sessions with public good. Subjects in sessions with a public good choose Y significantly more often than subjects in sessions without a public good (62% and 45% of the time, respectively; independent samples t=2.02).

Tax evasion

First consider the subjects that chose X. Observation 6 summarizes our results on tax evasion for these cases.

Observation 6: *In almost all cases, income was reported truthfully after choosing an X envelope.*

Support: Across all sessions we observed 588 X choices in rounds 4-8. After choosing X, in 3 cases (0.5%) less than the actual income was declared and in 18 (3.1%) more was declared. In all other (567) cases subjects reported their actual income truthfully. This result does not appear to be different in Albania than in the Netherlands.[15] A possible reason for declaring more than the actual income will be discussed in section 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, separately for each subject pool. In case less than the actual income was declared, the table also reports the number of times that this declared income was non-zero. In addition, it gives the average fraction of rounds that subjects evaded taxes. To calculate this, for each individual we first determined the fraction of times (s)he underreported income, treating overreporting as a missing value. Then, we calculated the mean of

[14] Choices in rounds 1-3 are not affected by the audit probability or public good treatments. This is not surprising, considering that in these rounds subjects do not even know that there will be taxes, audits and public goods later on.
[15] More details are available from the authors upon request.

Table 4: Distribution of income declaration

	P=1/2, public good			
	<	=	>	Evasion∗
AL-HS	22 (13)	50	3	0.19
AL-US	12 (9)	50	12	0.10
AL-HT	5 (5)	54	8	0.05
AL-UP	2 (0)	45	4	0.02
AL-UT	10 (2)	64	1	0.08
NL-HS	27 (6)	39	2	0.24
NL-US	46 (3)	35	1	0.39
NL-HT	1 (1)	18	4	0.03
NL-UP	5 (1)	30	2	0.09
Total	*130 (40)*	*385*	*37*	*0.14*
	P=1/6, public good			
AL-US	11 (5)	45	17	0.11
NL-US	65 (10)	20	10	0.58
Total	*76 (15)*	*65*	*27*	*0.34*
	P=0, public good			
AL-US	15 (3)	34	9	0.13
	P=1/2, no public good			
AL-US	8 (2)	40	6	0.08
TOTAL	**229 (60)**	**524**	**79**	**0.17**

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.

these fractions across individuals in a subject pool. This mean is reported in the table in the column 'evasion'. Our conclusions on tax evasion after choosing unregistered income (Y-envelope) are presented as observations 7-14.

Observation 7: *Not everyone switching to Y after round 3 evades taxes.*

Support: This observation is drawn from a comparison of tables 3 and 4. We can determine the increase in Y-choices after round 3 in table 3 and compare this to the extent of tax evasion in table 4. 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 even stronger when considering the fact 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%.

Observation 8: *In aggregate, underreporting of income occurs far more often than overreporting. This asymmetry is observed in about half of the subject pools. Overreporting is observed more in Albania than in the Netherlands.*

Support: In aggregate, the number of times income is underreported (229) is much higher than the number of times subjects report too much (79). 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 (p=0), tax evasion (underreporting) 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.[16] Contrary to what we observed for X choices, 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 χ^2 tests per round. Considering only the Y-choices where income was not reported honestly, we compare 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 was 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. In 5 of these tests it is significant at the 5%-level.

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. We will return to this point in the following section.

Observation 9: *It occurs quite often that subjects decide to evade taxes but do not evade them completely. This is more common in Albania than in the Netherlands.*

Support: In 60 out of 229 cases (26.2%) subjects reported an income higher than zero but lower than their actual income. For p=1/2 with public good, Albanians evaded taxes 51 times. In 29 cases (56.9%) they declared more than 0, however. 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

[16] 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.

p=1/6. As noted above, a selfish, rational individual evading taxes will report 0. We shall return to this phenomenon in section 5.

To get a better grip on tax evasion in our experiments, we now continue the analysis by considering the *extent* of tax evasion (in other words, the extent of underreporting per individual) while treating the cases where too much income was reported as missing values. In addition, we will not make a distinction with respect to how much is evaded: any reported income lower than the actual income is considered as tax evasion. This analysis is based on the numbers presented in the column 'evasion' in table 4.

Observation 10: *Tax evasion is higher in the Netherlands than in Albania.*
Support: Overall, the extent of tax evasion 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).

Observation 11: *Pupils and students evade taxes more than other groups do.*
Support: For p=1/2 with public good, we aggregate the data according to labor market position (HS/US versus HT/UT/UP). The pupils and 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).

Observation 12: *Differences across groups within a country are at least as important as differences between countries.*
Support: This follows from the analysis underlying observations 10 and 11. The aggregated group of Dutch pupils and students evades taxes more often than this group in Albania does (0.31 versus 0.14). Both groups evade more often than teachers/personnel in either country (0.06 versus 0.05). Non-pupils/students do not differ across the two countries. The difference between pupils/students in the two countries appears when comparing the two nations in aggregate. The difference when

aggregating per country (0.24-0.09) is smaller than when aggregating per labor market position (0.23-0.05).[17]

Observation 13: *Tax evasion in Albania is not affected by the audit probability; Dutch subjects evade more when the audit probability is lower.*
Support: An increase in the probability from 0 to 1/6 to 1/2 only marginally decreased evasion amongst Albanian (university) students, from 0.13 to 0.11 to 0.10 (cf. table 4). None of the pair wise differences is statistically significant at the 10% level. Dutch students, however, had the tendency to comply less when the audit probability was 1/6 (0.58) than when it was 1/2 (0.39). This difference is statistically significant at the 10%-level (independent sample t=1.80).

Observation 14: *The provision of a public good does not affect the extent to which Albanian students evade taxes.*
Support: The sessions without a public good were only run for Albanian students with p=1/2. The results show a marginally higher level of evasion if a public good is provided (0.10) than if it is not (0.08). The difference is not statistically significant at the 10% level, however. Note that the difference is even in the wrong direction if we consider the conjecture that Albanians comply more when a public good is provided.

Regression Analysis

In order to analyze subjects' decisions in more detail, we ran regressions. The first choice we want to explain is the decision whether or not to evade taxes. Denote the variable describing the extent of evasion (the column 'evasion' in table 4) 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

[17] Note that we can only make this comparison for p=1/2, with public good, because we only varied subject pools for these parameters (cf. table 1).

the interval [0,1]. In this model, the dependent variable $\hat{E} = \ln[(E+0.001)/(1-E+0.001)]$.[18] Model B is a linear regression with $\hat{E}$ as dependent variable.[19]

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 10,000.[20]
- 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. For all teachers and personnel, the value is 1. Because these groups are represented by LABORMARKET, JOB distinguishes pupils/students with a parttime job from those without.
- GENDER is a dummy variable equal to 0 for men and 1 for women.

Table 5 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 evade taxes at least once (model A) than men (a similar result is reported in Giese and Hoffman, 1999). The difference between men and women just misses statistical significance in explaining the extent of tax evasion (model B).

---

[18] The constant 0.001 is 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.

[19] 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.

Table 6: Regression results

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

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

Second, the risk attitude of 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. The positive coefficient implies that 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.

Third, three variables related to our treatments: country, labor market position, and audit probability, have significant effects on the decision whether or not to evade taxes. The Dutch evade more than the Albanians do and pupils/students evade more than teachers/personnel. With an audit probability of 1/6, subjects are more inclined to evade taxes at least once and they also evade more often (compared to an audit probability of 1/2). Surprisingly, we find no significant differences in tax evasion between p=0 and p=1/2. These results are in line with observations 10, 11, and 13. In line with observation 14, there is no significant effect of the no-public good treatment.

Finally, we analyze the decisions of income choice (X or Y) and tax evasion jointly. To do so, we need data 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

[20] Obviously, there is a possible endogeneity problem related to the inclusion of earn 4-8 as an independent variable. It is the only direct way to check the effect of earnings on evasion in a regression context, however. The conclusions for the other variables do not change if earn 4-8 is dropped from the regressions.

independent observations. As a consequence, we only consider these choices in one of the rounds 4-8, to wit, round 4.

The model is a two step logit for the choice of income source and tax evasion. The method is explained in Maddala (1983) and applied in Schram (1992). First (step 1), a logit regression is estimated for the binary choice whether or not to evade taxes in round 4. This is done for the set of subjects that chose Y. The parameters estimated (β_e) are then used to determine the so-called 'inclusive value' I_i:

(1) $I_i = \log[1+\exp(X_i'\beta_e)]$,

where X_i denotes the vector of independent variables for individual i.

In step 2, a logit regression is estimated for the binary choice between income source X or Y. For this regression, the inclusive value of step 1 is added as a regressor. For the probability of choosing Y, this yields:

(2) $Pr_{iy} = 1/\{1+\exp(X'\beta_x - \sigma I_i)\}$,

where β_x and σ are parameters to be estimated.

To understand (2), consider two special cases. For $\sigma = 0$, (2) reduces to:

(3) $Pr_{iy} = 1/\{1+\exp(X'\beta_x)\}$,

and for $\sigma = 1$ we have:

(4) $Pr_{iy} = \{1+\exp(X_i'\beta_e)\}/\{ 1+\exp(X_i'\beta_e) + \exp(X'\beta_x)\}$.

Equations (3) and (4) both describe the probability of choosing Y for special cases. (3) is a binomial logit model for the choice of X or Y. When deciding on X or Y, the individual does not consider the possibilities of evading after choosing Y (Pr_{iy} is independent of β_e). Hence, for $\sigma = 0$ the income and evasion decisions are made *sequentially*.

In (4), we have a multinomial logit probability for a model with three choices: choose X (denoted by x), choose Y and evade (ye), or choose Y and comply (c). The corresponding probabilities are, respectively:

(5a) $Pr_{ix} = \exp(X_i'\beta_y)\}/\{ 1+\exp(X_i'\beta_e) + \exp(X_i'\beta_y)\}$

(5b) $Pr_{iye} = \exp(X_i'\beta_e)\}/\{ 1+\exp(X_i'\beta_e) + \exp(X_i'\beta_y)\}$

(5c) $Pr_{iyc} = 1/\{ 1+\exp(X_i'\beta_e) + \exp(X_i'\beta_y)\}$.

In this case, i is considering the choice problem as if it consists of three distinct and equally weighted options: X, Y-evade and Y-comply. Hence, when deciding which envelope to choose, this individual takes the possibility of evading into account. The probability of choosing Y (4) is the sum of the probabilities of choosing Y and evading (5b) and choosing Y and not evading (5c). Hence, for $\sigma = 1$ the income and evasion decisions are made *simultaneously*.

As a consequence, an estimate of σ provides information about the way in which the decisions to choose X or Y and whether or not to evade taxes are related. A coefficient close to 0 implies that the decision between X and Y is made independent of the subsequent tax evasion decision. A coefficient close to 1 is evidence that the decisions about income source and evasion are made simultaneously.

Observation 15: T*he decisions about income source and tax evasion are made simultaneously in round 4.*

Support: We estimated the coefficients of eqs. (1)-(2) using the same regressors as in table 5 for step 1 and the treatment variables for country, labor market position and audit probability as independent variables in step 2.[21] We estimate a coefficient $\sigma = 2.21$, with a 95% confidence interval for σ of (0.80, 3.62). Hence, we reject $\sigma=0$ but do not reject $\sigma=1$.

5. A discussion of the results

First, consider the three novel features of our experimental design: it compares tax evasive 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. All three have proven to be empirically important. Though we found differences across countries, they were opposite to what we observe outside of the laboratory. The

[21] The results of the regressions are not reported here because these regressions use only one round of data, whereas the results in table 5 are based on all the data we have on evasion. Most of the results from table 5 are also found when considering only round 4, however. Details are available upon request.

implications of these differences are discussed below. We have also shown that significant differences across groups exist that are sometimes larger in magnitude than the differences between countries. The distinction in type of income also turned out to be important. When tax evasion was possible, there was a shift towards unregistered income. This shift did not differ significantly across subject pools. As discussed in section 1, this may distort the allocation of labor across different types of income.

Next, consider the conjectures on our treatment effects as discussed in section 3. The first was that evasion would be higher in Albania. We observed the opposite. Second, we expected differences within a country to be smaller that differences across countries. Again, we observed the opposite. The third conjecture was that evasion would decrease with audit probability. This was observed in the Netherlands but not in Albania. Finally, the conjecture that more evasion would take place if there was no public good was not supported. We did observe a decrease in choices of unregistered income when there was no public good, however. All in all, it appears that our conjectures (which were mainly based on the existing literature) are not capturing what is happening in our experiments. In the conclusions, we will discuss the implications of our results.

Finally, there were patterns of choices that cannot be explained by the simple theoretical framework: (i) there was some overreporting of income in all subject pools (cf. observation 8); (ii) in many cases, tax evasion was incomplete, i.e., did not involve reporting an income of zero (cf. observation 9); (iii) there was a significant move towards Y-choices that could not be completely attributed to the possibility of tax evasion (cf. observation 7); (iv) in sessions without a public good, subjects chose the unregistered income less often (cf. observation 5). (i) and (ii) were observed more often in Albania than in the Netherlands, whereas (iv) is based on data that were only gathered in Albania.

These observations might be explained by the role of the public good in combination with non-selfish preferences. The theoretical discussion presented in section 3 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 and incomplete evasion can thus be interpreted as voluntary contributions to the public good. Subjects

might also be willing to switch to the (more risky) Y envelopes in case their taxes contribute to a public good. In that case, Y incomes become less attractive when there is no public good.[22]

An alternative explanation is that we are simply observing erratic behavior. There are a few conclusions from our data that seem to point against simple error as an explanation, however. First of all, all these 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 Albanian high school pupils. 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.

6. Conclusion

The literature shows fine examples of both non-experimental and experimental research on tax evasion. The latter offers the opportunity to study individual behavior in a controlled environment. This is what the study in this paper does. It allows us 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, our study puts the experimental environment closer to reality than previous experiments on tax evasion have done 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 evasion of taxes are made simultaneously.

Our evidence on the choice of income type 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, we do find a significantly lower choice of unregistered income when there is no public good than when there is. When analyzing tax compliance, we observe that almost all subjects report their income truthfully after choosing a registered income. The few that do not report honestly tend to report too much rather than too

[22] Table 4 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.

little. Tax evasion does occur after unregistered income is chosen. When the latter happens, many more cases of underreporting than overreporting are observed in aggregate.

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 pupils/students 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.

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 we controlled for differences across institutions, this country difference was the other way around. If it is true that tax evasion outside the laboratory in Albania is higher than in the Netherlands, than our experiments show that this is not a consequence of differences in attitudes or cultures. If it were, we would have found higher evasion in the Albanian sessions. Apparently, differences in tax institutions between the two countries –which do not appear in the experiments- matter. In ongoing research, one of the authors (Gërxhani) has analyzed data from a large household survey in Tirana, which collects more information about tax evasive behavior. A combination of this evidence with the experimental results of this paper indicates that the distinct formal institutions and their 'informal' use by individuals cause the different levels of tax evasion in the two countries. For a detailed analysis see Gërxhani (2002).

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, we would have concluded that evasion is higher in Albania. This provides a *caveat* for the interpretation of distinct experimental results that have been observed across countries.

References

Albanian Center for Economic Research (ACER), 1999: Problems of Economic Growth in Albania, Tirana, Albania.

Alm, J., 1991: "A Perspective on the Experimental Analysis of Taxpayer Reporting". The Accounting Review 66/3: 577-593

Alm, J., Jackson, B.R. and McKee, M., 1991: "Estimating the Determinants of Taxpayer compliance with Experimental Data", National Tax Journal, XLV, no.1.

Baldrey, J.C., 1986: "Tax Evasion is not a Gamble: A Report on two Experiments". Economics Letters 22: 333-35.

Boeschoten, W.C. and Fase, M.M.G., 1984: The Volume of Payments and the Informal Economy in the Netherlands, 1965-1982, M. Nijhoff, Dordrecht.

Bolton, G., and Ockenfels, A., 2000: "ERC: A Theory of Equity, Reciprocity and Competition", American Economic Review 90 (1): 166-193.

Brandts J. , Saijo T. and Schram, A., 2000: "How Universal is Behavior? A Four Country Comparison of Spite, Cooperation and Errors in Voluntary ContributionMechanisms", mimeo, University of Amsterdam.

Collins, J.H. and Plumlee, R.D., 1990: "The Taxpayer's Labor and Reporting of Decision: The Effect of Audit Schemes"? The Accounting Review 66:3, 559-576.

EBRD, 1996: Transition Report, EBRD, London.

Feige, E.L., 1989: The Underground Economies: Tax Evasion and Information Distortion (ed.), Cambridge University Press.

Frey, B.S., and Weck, H., 1981: ‘’Bureaucracy and the Shadow Economy: A Macro Approach’’, in Anatomy of Government Deficiency, Hanusch, H., (ed.), Detroit: Wayne State University Press.

Friedland, N., Maital, S., and Rutenberg, A., 1978: “A Simulation Study of Income Tax Evasion”, Journal of Public Economics, 10: 107-116.

Geise, S.and Hoffman, G., 1999: “Tax Evasion and Risky Investments in an Intertemporal Context – An Experimental Study”, Discussion paper 153, Humboldt University.

Gërxhani, K., 2002: “The Informal Sector in Transition: Tax Evasion in an Institutional Vacuum”, Ph.D. thesis, University of Amsterdam, Thela Thesis publisher, Amsterdam.

Hansson, I., 1989: ”’The Underground Economy in Sweden”, in Feige (eds.), The Underground Economy: Tax Evasion and Information Distortion, Cambridge University Press, 219-37.

Hart, K., 1971: “Small Scale Entrepreneurs in Ghana and Development Planning”, Journal of Development Planning, July.

Hart, K., 1973: ”Informal Income Opportunities and Urban Employment in Ghana”, Journal of Modern African Studies 11/1: 61-89.

Isaac, R.M., Walker, J., and Thomas, S., 1984: “Divergent Evidence on Free Riding: An Experimental Examination of Possible Explanations”, Public Choice 43: 113-149.

Lee, R.M., 1985: “De-jeopardizing Techniques for Asking Sensitive Questions on Survey: Applications from the Study of the Black Economy”, unpublished paper, Department of Sociology and Anthropology, University College, Swansea.

Maddala, G., 1983: Limited-Dependent and Qualitative Variables in Econometrics, Cambridge University Press, Cambridge.

Roth A.E., V. Prasnikar, M. Okuwo-Fujiwara, and S. Zamir, 1991: "Bargaining and Market Behavior in Jerusalem, Ljubiljana, Pittsburgh and Tokyo: An Experimental Study". American Economic Review 81: 1068-1095

Saijo T. and H. Nakamura, 1995: "The 'Spite' Dilemma in Voluntary Contribution Mechanism Experiments", Journal of Conflict Resolution 39: 535-560.

Schneider, F. and Enste, D.H., 2000: "Shadow Economies: Size, Causes and Consequences", Journal of Economic Literature, vol. 38/1: 77-114.

Schram A., 1992: "Testing Economic Theories of Voter Behaviour Using Micro-Data", Applied Economics 24: 419-428.

Schram, A., 2000: "Sorting out the Seeking: The Economics of Individual Motivations", Public Choice 103: 231-258.

Tanzi, V., 1982: The Underground Economy in the United States and Abroad (ed.), D.C Heath and Co. Lexington, Massachusetts, Toronto.

Thomas, J.J., 1992: Informal Economic Activity, LSE, Handbooks in Economics, London: Harvester Wheatsheaf.

UNDP (United Nations Development Program), 2000: Albanian Human Development Report 2000, Tirana, Albania.

Webley, P., Robben, H., Elffers, H., and Hessing, D., 1991: Tax Evasion: An Experimental Approach, Cambridge University Press, Cambridge.

Appendix 1: English version of the Albanian and Dutch instructions used in the experiments[23]

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.[24] 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 soon come by every desk one by one. 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.

---

[23] This is the basic outline of all instructions used in the experiments. Other instructions, based on the various treatments described in the main text, are available upon request.

[24] 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 in that 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 in that round.

……...

Presenting your decision

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

participant: __

group _

	Choice X or Y
Round 1	X Y
Round 2	X Y
Round 3	X Y

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, please 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 actual amount of your earning or some other amount.

The audit

There is a probability that we will audit your reported income. If you have chosen an X-envelope, we will always audit whether 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 report 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 actual earnings. You will have to pay 25% of these earnings to the fund. In addition, you have to pay 25% of your actual earnings as a fine for declaring less. If you have reported more than your actual 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 actual earnings, that will be paid to you at the end, will be determined by the actual 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 actual 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 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 anonymous. 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.

participant: __

group __

	Choice X or Y	**Earnings** (reported by you)	Actual earnings (to be filled out by us)	**Contribution to the fund** (to be filled out by us)	**Fine** (to be filled out by us)
Round 4	X Y				
Round 5	X Y				
Round 6	X Y				
Round 7	X Y				
Round 8	X Y				

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 'actual earnings') 'no audit'. If we do audit the envelope, we will write the actual earnings in your table. This is your reported amount, if you reported the actual 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 actual 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 actual amount after been audited, we will write down a zero (0). If after being audited, your reported amount is smaller than the actual amount, we will write down 25% of the actual 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, to wit ……
How many hours per week do you work (please enter)?	……