DISSERTATION

INCOME TAX EVASION: THEORETICAL MODELING AND EMPIRICAL EVIDENCE

Submitted by

Emin F. Gahramanov

Department of Economics

In partial fulfillment of the requirements

For the degree of Doctor of Philosophy

Colorado State University

Fort Collins, Colorado

Summer 2007

COLORADO STATE UNIVERSITY

May 24, 2007

WE HEREBY RECOMMEND THAT THE DISSERTATION PREPARED UNDER OUR SUPERVISION BY EMIN F. GAHRAMANOV ENTITLED "INCOME TAX EVASION: THEORETICAL MODELING AND EMPIRICAL EVIDENCE" BE ACCEPTED AS FULFILLING IN PART REQUIREMENTS FOR THE DEGREE OF DOCTOR OF PHILOSOPHY.

Committee on Graduate Work

Robert W. Kling

Charles F. Revier

Frank Caliendo

W. Marshall Frasier

Advisor: Chuen-Mei Fan

Department Head: Steven Shulman

ABSTRACT OF DISSERTATION

"INCOME TAX EVASION: THEORETICAL MODELING AND EMPIRICAL EVIDENCE"

Income tax evasion is a very important problem faced by most of the countries around the world. The phenomenon interferes with economic efficiency, socially desirable income distribution, long-run economic growth, and might even negatively affect the price stability.

The intent of this study is to contribute to the economic theory of income tax evasion by demonstrating the ways to resolve the paradoxical relationship between the tax rate and compliance and to conduct various cross-model and cross-country comparisons, relying both on the theoretical and applied analysis.

The study considers the intergenerational welfare implications of the recent dramatic decline in the income tax audit rate in the United States, which has been a source of big concern for many politicians, economists, and general public. It has been demonstrated that the wide-spread evasion can worsen the welfare of the generation working during the fall in the audit rate. Other issues, such as tax compliance costs and revenue-maximizing taxation have also been analyzed.

Emin F. Gahramanov
Economics Department
Colorado State University
Fort Collins, CO 80523
Summer 2007

ACKNOWLEDGMENTS

First, I would like to thank my academic advisor Dr. Chuen-mei Fan for not only leading me throughout my graduate study and directing my dissertation, but being like a mother to me while I was away from my real parents. Dr. Fan would be the first person I would rush to express my joy or even frustration on virtually any manner, and she would always be ready to help, support and lead me with her encouragements and wisdom. Dr. Fan constantly amazed me with her impeccable personality. I would always remember her as a fine, honest, knowledgeable and hard-working person.

Second, I am grateful to Dr. Liang-Shing Fan, who was my academic advisor while I was in the master's program. Although he did not have any formal obligation to me while I was pursuing my doctorate degree, Dr. L-S Fan would continue to be a tremendous source of moral support for me.

Third, I would like to thank Dr. Robert Kling. I sincerely considered Dr. Kling my informal advisor, who would stand ready to help, whenever I needed it. Dr. Kling is an excellent teacher and intelligent individual. It is a real pleasure to know Dr. Kling.

Fourth, I deeply thank my teachers and committee members Dr. Charles Revier and Dr. Marshall Frasier. Dr. Revier is a perfect teacher and a true gentleman. Dr. Frasier impressed me all the time with his sense of humor and dedication to students, which is greatly appreciated.

Fifth, I would like to express my gratitude to Dr. Frank Caliendo. Although I have known Dr. Caliendo for a much shorter time than most of my teachers, I learned and benefited a lot from him. His rigor, dedication, and intelligence will serve as an example to me.

Now, it is my real pleasure to sincerely thank my family. This includes my parents, grandparents, sister, my father and mother-in-law, as well as my brothers-in-law. Thanks immensely for keeping me in your prayers and for playing a major role in all my successes. Of course, I am very fortunate to have my lovely wife Konul standing by me at all times, and being a perfect wife and invaluable friend. Enormous thanks go to my beautiful little son Fikrat, whose smile gave a new meaning and joy to my life.

Last, but not the least, I thank God for letting me to be where I am today, and always fulfilling my dreams. I owe Him for all the good things I have in my life, including the successful completion of this dissertation. Needless to mention, any errors and imperfections in this study are solely mine and no one else's responsibility.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

"Like mothers, taxes are often misunderstood, but seldom forgotten."
Lord Bramwell, 19th Century English jurist.

Chapter One: Introduction.

1.1. Problem Statement.

Income tax evasion is one of the important and highly-debated topics in economics literature. It is difficult to overstate the impact of this phenomenon on modern economies. On the one hand, tax evasion means lower revenue generated for the government, *ceteris paribus*. On the other hand, evaded taxes might become important sources of private capital accumulation, which in turn, can be channeled into investment and might become essential for the long-run economic growth. In addition, various sectors traditionally are open to different possibilities for successful evasion and thus, economic agents' decisions can be distorted, possibly causing the misallocation of resources.

Gary Becker (1968) pioneered the analysis of illegal behavior, which was first applied to the problem of income tax evasion by Allingham and Sandmo (1972). Since then, there have been numerous tax evasion models, that frequently reach controversial conclusions. For instance, Yitzhaki (1974) showed that an increase in the tax rate makes people more honest in declaring their incomes. This result has been puzzling to many. Even though not all the empirical results contradict Yitzhaki's conclusion, some prominent authors, in analyzing the phenomenon, have proposed the complete abandonment of the benchmark portfolio choice models of tax evasion (i.e., Allingham-Sandmo (1972) and Yitzhaki (1974) specifications) and developed several new models, ranging from static to dynamic ones.

Basically, income tax evasion models can be divided into two categories. In the first category there are frameworks which elaborate on existing tax evasion models at the micro theory level and try to develop ones leading to more consistent results from the viewpoint of economic intuition and real-world evidence. These are mainly static models. Sometimes they are tested empirically. In the second category, macroeconomic implications of tax evasion are emphasized. These models are predominantly dynamic. Although analytically rich, the latter have not carefully considered many important questions, such as the level of capital accumulation, *dynamic efficiency* (whether the capital stock on the balanced growth path maximizes per capita consumption), and the consequent implications for saving and consumption in the economy.

As far as cross-country analysis is concerned, an important point to consider is that different economies face different tax evasion "gambles". For example, in modern developed economies (such as the U.S.A. or Canada) taxpayers generally do not have control over the amount of their income reported to the government. Their earnings are reported to tax authorities by the third parties, be it an employer or a local financial intermediary. However, in many developing countries the reporting system is not that advanced, meaning that income declaration is bilateral, between a taxpayer and the government. In the absence of other counter-acting incentives, the bilateral reporting (possibly leading to weaker enforcement parameters) opens up more possibilities for evasion, and empirical studies should take into account that distinction between countries' tax systems. One very simple way to address it is by assuming different values for the taxpayers' subjective evaluation of getting caught. The latter is traditionally approximated by the tax audit rate, and we will rely on such an approach, too.

The main focus of this study will be to analyze the aforementioned theoretical questions by developing tax evasion models, based on microeconomic decision making that have macroeconomic repercussions. Not only will we focus on theoretical details, but also try to conduct some applied analysis.

1.2. The Importance of the Study.

The models of tax evasion are considered to be important from several aspects.

The main one, in our view, is a theoretical aspect wherein tax evasion imposes a considerable burden on the government budget, distorts economic decision-making, redistributes wealth arbitrarily, and may have an important impact on saving and capital accumulation. It is, therefore, important to tackle at least some of these issues on theoretical grounds. In addition, as was mentioned above, some traditional static models of tax evasion conclude that more aggressive taxation encourages compliance. As we said, this result was first obtained by Yitzhaki (1974) and caused considerable debates, which still continue. "Despite the fact that Yitzhaki's result accords with neither common sense nor evidence, it has dominated the tax evasion literature for more than quarter of a century" (Al-Nowaihi and Pyle 2000, p. 249). The authors continue: "... [T]he prediction of the theoretical model concerning the effect of a change in the rate of income tax does not accord with empirical evidence concerning its effects" (Ibid. p. 257). However, *not* all the empirical studies conclude that more aggressive taxation encourages evasion. But most studies do.

Hence, many realized the importance of developing a framework which will not produce allegedly counterintuitive and empirically invalid results. Although this goal was more or less accomplished in various frameworks of tax evasion models, none had attempted to resolve the issue while staying as "close" as possible to the original benchmark tax evasion model of Allingham and Sandmo (1972), which was the factual cause of controversy. In fact, some suggested that the subjective expected utility framework, adopted by Allingham and Sandmo, ought to be abolished (Yaniv 1999). However, the expected utility framework has important advantages, such as being analytically plausible and relatively simple, i.e., we believe that its complete abandonment may not be necessary.

As far as the arbitrary redistribution of the wealth is concerned, for instance, one may ask: are some taxpayers who get away with cheating much better-off than the rest of us, who were not so lucky? That question has not been properly addressed by now, to our best knowledge. In fact, is it possible that if all people cheat on their taxes, then everyone is worse-off to some extent? That sounds like an interesting question to address.

Finally, it is reasonable to assume that evaded taxes are often and easily converted into interest-bearing assets, thus, contributing to private capital accumulation. The problem of capital accumulation raises another question: whether the economy over-accumulates capital, becoming dynamically inefficient. Thus, it is interesting to discuss the level of capital accumulation, after allowing illegal funds to flow into people's financial asset accounts.

1.3. The Objectives of the Study.

In this study we pursue four main objectives. The first objective is to critically assess Yitzhaki's finding within the benchmark tax evasion model, without abandoning the expected utility approach. We will try to accomplish this within a static framework. Note that we are not targeting in this work to resolve the theoretical contradiction of the benchmark tax evasion model, since it has already been resolved in previous studies and after all, not everyone unanimously agrees that the real-world evidence actually contradicts what Yitzhaki concluded. Further, the model may resolve a certain puzzle but fail to empirically confirm other patterns of observed economic behavior. In addition, the authors who were able to theoretically resolve the puzzle accomplished it by introducing new set of assumptions and by designing frameworks which are significantly different from the original one. Thus, we are trying to understand whether it is possible to accomplish the same goal rather by staying "very close" to the original model formulation or if not, then why so. The importance of this is evident at least because some prominent studies proclaimed the benchmark tax evasion models completely useless. In this study we will attempt to introduce a progressive tax rate structure, the absence of which is allegedly one of the main shortcomings of the benchmark models.

The second objective is to develop a new dynamic model of tax evasion, flexible enough to consider the Yitzhaki puzzle, and empirically assess various policy changes. Some other dynamic model will also be analyzed. In the models under consideration we will also pay attention to the implications arising from the introduction of tax compliance costs. Note that not only is discussing various models interesting for the sake of analytical clarity, but also useful for choosing the one with more sensible and intuitive conclusions.

The third objective is to simulate the models, by relying on parameters for a number of OECD and developing countries. The cross-model and cross country-comparison may provide additional grounds for the criticism regarding various theoretical underpinnings of the models. In the empirical section, the main focus is the U.S. economy but other countries will also be briefly analyzed. We will also closely focus on intergenerational welfare comparisons. Finally, obtained qualitative and quantitative results will be used to suggest some modifications and extensions of theoretical models.

1.4. The Organization of the Study.

The rest of this dissertation is organized as follows. Chapter Two will be devoted to a review of the theoretical literature on tax evasion. Static and dynamic models will be reviewed wherein implications and findings will be discussed. In Chapter Three we will revisit the benchmark model of income tax evasion. Namely, we will extend the benchmark model of tax evasion by adding an endogenous tax rate structure. In Chapter Four, a new dynamic model will be designed, which will be a two-period overlapping-generations (OLG) framework with tax cheaters as decision-makers. In the dynamic framework we will try to adopt as few assumptions as possible. Implications for evasion, consumption, saving, speed of convergence to the long-run equilibrium, capital accumulation, dynamic efficiency, and compliance issues will also be discussed.

In Chapter Five we will turn our attention to the problem of tax evasion in the context of households with infinite life span. As some prominent studies claim, moving from a static to continuous-time framework is sufficient to obtain more plausible results (Lin and Yang 2001). We will revisit, assess, and modify the existing benchmark continuous-time framework of the income tax evasion problem.

The presented models will be calibrated in Chapter Six, where we will mainly focus on the U.S. economy. We will look at main macroeconomic indicators and undertake some cross-model and cross-country comparisons. Finally, the last chapter will briefly summarize the findings of the study, critically evaluate them, and suggest possible theoretical extensions.

Chapter Two: Literature Review.

2.1. Static Models of Tax Evasion.

Tax evasion, non-compliance and illegal economic activities are very important problems for virtually any economy, no matter how advanced and developed it is. According to the report of *Internal Revenue Service* (2006), the estimated gross *tax gap* (the difference between true and actually paid tax liabilities) for the year 2001 in the U.S.A. was $345 billion, which corresponds to about 16.3% of the true overall tax liability. In Greece and U.K., for example, the amount of taxes evaded is estimated to comprise 22.5 and 11.5 percents of taxes collected, respectively (Gupta 2004, p. 3).

Economists have tried to analyze the behavior of economic agents who engage in illegal activities and the incentives encouraging them to do so by developing formal models and applying them to a variety of socioeconomic problems, such as tax evasion. Following Becker's (1968) classic paper on the economics of criminal activity, Allingham and Sandmo (1972) pioneered the analysis of income tax evasion (the A-S model) from theoretical perspectives. The analysis assumes that dishonest taxpayers (trying to determine the amount of income to declare) are *von Neuman-Morgenstern* (VNM) expected utility maximizers given the opportunity locus between "being caught" and "not being caught" (Tresch 2002, p. 513). That is, private decisions about evasion are similar to decisions about investment in a risky asset. Consequently, an income tax evasion model is often time referred to as a "portfolio choice model of tax evasion".

Formally, the taxpayer's problem is to maximize the following expected utility function:

$$E[U] = (1-q)U(W-\theta X) + qU(W-\theta X - F(W-X)),$$

where θ is the income tax rate, q is the probability of detection, X is the amount of declared income, F is the fine rate levied on the amount of concealed income and

W is the true income, i.e., X plus the concealed income.[1] The optimum solution will largely depend upon the likelihood of detection and the penalty size. The basic conclusion is that, if the utility function is continuous and unbounded from below, a sufficiently high penalty rate and/or probability of detection will eliminate all the incentives to hide income. The key comparative-static result is that an increase in the tax rate has an *ambiguous* effect on the incentives to cheat due to the competing income and substitution effects. The substitution effect arises since a higher tax rate means greater marginal benefit of cheating. The substitution effect alone leads to more evasion since the opportunity cost of honesty goes up. In addition, an increase in the tax rate has a negative income effect since taxpayers feel less wealthy, therefore, provided decreasing absolute risk aversion this effect alone would simply reduce cheating. Both the substitution and income effects compete against each other and thus, it is impossible to say *a priori* whether higher taxes encourage or discourage dishonesty. However, given real-world evidence, it is tempting to assume that the substitution effect dominates.

Kolm (1973) approached the income tax evasion problem by trying to distinguish the 'public' problem from the 'private' one wherein the former refers to the tax collector and the latter refers to the taxpayer. The taxpayer's problem is to maximize her expected utility (as in the A-S model) by deciding how much out of her income to declare, given the parameters of the model, i.e., the income tax rate, the probability of detection, and the penalty size. Even though the total tax yield is used to provide public goods, which, in turn, affects the utility of the taxpayer, the latter, being a "small" agent, ignores the public goods-driven utility in her decision-making. On the contrary, the tax collector chooses to maximize the taxpayer's expected utility from both the private goods (income), and the public goods, by choosing the above-mentioned model parameters. Thus, in the social planners' view, the agent's expected

[1] Unless otherwise specified, all the similar up-coming notations will carry the same meaning as here.

utility can be expressed as

$$S = (1 - q)U(W - \theta X) + qU(W - \theta X - F(W - X)) + V(T),$$

where V is the utility of the public goods, T is identically equal to expected tax revenue ($= \theta X + qF(W - X)$) minus the costs of detecting evasion expressed per person, i.e., T is "[T]he net average revenue of the whole operation" (Ibid. p. 267). A single taxpayer takes T as given and therefore maximizes the first two sums of the expression above (call it $E[U]$ as in Allingham and Sandmo 1972). However, the government, given the taxpayer's decision, maximizes S by choosing θ, q, and F, i.e., taking $\frac{\partial E(U)}{\partial X} = 0$ which is the taxpayer's optimum condition. For simplicity it is assumed that the taxpayers are identical in their preferences and earning capabilities, meaning that S times the number of citizens gives the sum of the citizen's utilities. So, by separately solving the taxpayer's and the tax collector's problems, *assuming* that $\frac{\partial X}{\partial \theta} < 0$, and then comparing the resulting conditions simultaneously, Kolm comes to the conclusion: "for *ex ante* public choices, a public pound has a greater social value than a private pound...This value of public funds must be used for the optimal choices of other public finance instruments (public expenditures, other taxes, public prices, etc.)" (Ibid. p. 269).

Thus far, everywhere in the literature the tax rate is assumed constant. The first attempt to relax that assumption was made by Srinivasan (1973). The total tax to be paid, $\theta(W)$, can be specified as follows: $\theta(W) > 0$, $\frac{d\theta(W)}{dW} > 0$, $\frac{d^2\theta(W)}{dW^2} \geqslant 0$, $\forall W$. Note that when $\frac{d^2\theta(W)}{dW^2}$ is strictly greater than zero, we have a progressive marginal income tax rate structure. The taxpayer tries to determine the proportion, λ, of her true income to understate. The penalty parameter is taken to be endogenous and is identically equal to $F(\lambda)$, the so-called "penalty multiplier". The penalties are assumed to be imposed on evaded income, *not* on evaded taxes only, i.e., we have

$F(\lambda)$ times the evaded income, λW.

The individual is assumed to choose λ to maximize her expected *income*, which is shown below:

$$E[I] \equiv q[W - \theta(W) - F(\lambda)\lambda W] + (1-q)[W - \theta\{(1-\lambda)W\}].$$

The main results are 1) an increase in q decreases the optimal proportion, λ^*, by which income is understated; 2) given a progressive tax function and constant q, the more income a person has, the larger the optimal proportion by which she will understate her income; 3) if the marginal tax rate is constant and q increases with income, then λ^* declines as income rises. Interestingly, Srinivasan is not concerned about the effect of a change in the tax rate on the fraction of evaded income.

Yitzhaki (1974) was the first one to reformulate the maximization problem, which had important implications regarding the relationship between evasion and tax rate. The expected utility is in the form of:

$$E[U] = (1-q)U(W - \theta X) + qU(W - \theta X - F\theta(W - X)).$$

Note that here the penalty is imposed on the amount of evaded taxes (as it is under most current tax laws), instead of on the amount of concealed income. This is, certainly, a more realistic scenario. But if so, the substitution effect vanishes when the tax rate goes up. This is because an increase in the tax rate increases the expected penalty payment along with an increase in the marginal benefit of cheating. At the optimum, these effects exactly cancel each other out and the substitution effect disappears, and only the income effect remains. Consequently, assuming decreasing absolute risk-aversion, higher tax rate makes people more honest in reporting their incomes. This finding is sometimes called in the literature the "Yitzhaki puzzle". The discussion over the paradoxical result of an inverse relationship between the size of evasion and the tax rate still continues in the modern income tax evasion literature. "The absence of ambiguity in theoretical models is often considered to be a good thing, but there is a puzzle involved in the Yitzhaki analysis. The ambiguity of the original A-S model is removed, but what is left is a result that goes directly against most people's intuition about the connection between the marginal tax rate and the amount of evasion, and also much empirical evidence" (Sandmo 2005, p. 647).

Pencavel (1979) undertakes an approach of making the tax rate structure endogenous *and* incorporating the hours of work into the utility function. The key is the tax payments system, which is a linear continuous function in the form $T = -s + \theta X^{\sigma}$. Here T is the tax payment, s stands for the welfare payment to the individual with zero income, σ captures the relationship between the reported income and changes in tax payments. Progressiveness is achieved by letting σ be greater than 1, because the marginal tax rate and the ratio of net tax payments to income would rise with the latter. Consequently, the individual's problem is to maximize

$$E[U] = qU[I^c, h] + (1-q)U[I^o, h],$$

where I^c is net income of the taxpayer caught cheating and thus penalized (with the penalty imposed on evaded taxes), I^o is the net income of the evader who is not caught cheating and h is the number of hours worked. So, not only does the tax penalty affect people's decision to declare income, it also affects people's labor-leisure choice in the first place. Further, $I^o = W(h) + s - \theta X^{\sigma}$. As before, when the tax parameters of the model increase but true income is held fixed, the individual with decreasing absolute risk aversion utility evades less, due to the previously defined income effect. "However, when true income becomes an endogenous variable, an increase in... [the tax parameters]... may reduce the hours of work and thus may reduce true income which may encourage the individual to report less income to the tax authorities"

(Ibid. p. 121). Thus, Pencavel generated an obvious ambiguity.

Christiansen (1980) compared fines to the probability of detection in trying to determine the corresponding effectiveness to deter cheating incentives. The expected utility to be maximized is of the form

$$E[U] = (1-q)U(\omega + i) + qU(\omega - Fi),$$

where ω is the post-tax income when there is no evasion and i is the amount the economic agent reduces her tax payment via evasion. So, $\omega + i$ is the amount of disposable income. Further, it is assumed that the probability of being caught is a function of the fine rate. Specifically, it is assumed that a higher probability of detection and fine rate are inversely related by "some tie" (Ibid p. 390). That is, Christiansen treats both q and F as policy factors that can be designed, and establishes two alternative relationships:

$$\begin{aligned} 1 - q - qF &= \text{constant}, \\ qF &= \text{constant}. \end{aligned}$$

If the first relationship is preserved, the expected gain from tax evasion ($= (1-q)i - qFi$) is unaffected, when the fine rate declines, for example. By determining that $\frac{di}{dF}$ is less than zero from the first order conditions, and utilizing the above relationships between q and F , the author establishes that "[i]f the fine rate is increased, but the efforts to detect tax evaders are adjusted so as to keep the expected gain from tax evasion unaltered, risk-averters will always reduce their tax evasion" (Ibid p. 391). It clearly follows that larger fines are more effective in deterring cheating than larger probability of being caught. Then the question arises: why does not the government simply raise the penalty parameter, instead of committing itself to the costly verification? According to Christiansen, the punishment should be an adequate response to the committed crime or violation. It would be politically insensible to confiscate the taxpayer's entire monetary and non-monetary assets for a small tax fraud.

Cowell (1985) focused on the individual who cheats the government not by directly hiding her income, but by taking two or more different jobs: the combined income from them is difficult to observe, and therefore, to tax. Assume that h_0 and h_1 are the proportions of the individual's time in legal and illegal employments, respectively. The time spent on leisure is $1 - H$ $(H \equiv h_0 + h_1)$. The true wage rate from the legal employment is W_0 and that from the illegal one is W_1. The tax schedule is progressive and linear in the form $T = \theta i_0 - s$. Here i_0 is taxable income. The two states of the world can be represented as "not being caught" (with probability $1 - q$) and "being caught" (with probability q). Thus, the corresponding disposable incomes are, respectively, $s + w_0 h_0 + W_1 h_1$ and $s + w_0 h_0 + w_1 h_1$, where $w_0 = (1 - \theta)W_0$, $w_1 = (1 - t_1)W_1$ and t_1 is the penal tax rate (imposed on evaded *income*). These states of the world would enter the decision-maker's well-behaved expected utility function (depending on leisure and disposable income) to be maximized with respect to h_0 and h_1. Again, in Cowell's world the agent does not decide how much out of her income to declare, but how much leisure to consume and how to allocate her remaining time between illegal and legal employments. With labor supply decision endogenous, there is no straightforward comparative statics conclusion, unless the utility and the labor supply functions are assumed to be in a specific form. Cowell assumed that the individual labor supply is backward-bending and the utility function is additively-separable. When the tax rate goes up, the taxpayer is inclined to be more honest due to the income effect. However, when the average tax rate goes up and the labor supply is backward bending, the individual will decide to work more. Thus, the total effect on 'evasion effort' is ambiguous.

In reality not all people evade, perhaps because there are some psychic or stigma costs to evasion. In other words, people may vary in accordance with their honesty characteristics. This was noted by Gordon (1989). Formally, let the consumption in two states of the world be, $C = (1-\theta)W + \theta E$ with probability $1-q$, and $C = (1-\theta)W - F\theta E$ with probability q. Here E is the concealed income. The utility of the taxpayer is in the form $U \equiv U(C, -E)$, where $\frac{\partial U}{\partial C} > 0$, $\frac{\partial U}{\partial(-E)} > 0$. The marginal relations due to the individual maximization problem lead to some interesting implications. In particular, when $1-q > qF$, sufficiently dishonest people will evade, and others will not. Interestingly, under decreasing absolute risk aversion and a constant marginal disutility assumption, an increase in the tax rate has an ambiguous effect on the incentives to hide income. "... [T]he reduction in wealth caused by the higher tax rate has two competing events: the conventional desire caused under decreasing absolute risk aversion to reduce the size of the gamble (tending to decrease evasion) and the response to the lower relative psychic costs of being dishonest (tending to increase it)" (Ibid p. 800). Note that the psychic costs become relatively lower because of the conventional substitution effect: an increase in the tax rate increases marginal benefit of cheating. If the marginal disutility of evasion is high enough, the original amount of undeclared income is relatively small, implying a small income effect. Hence, the substitution effect will dominate the income effect and evasion will go up with the tax rate.

Not all of the models assume that the likelihood of getting caught is given exogenously. Indeed, taxpayers can influence the probability of getting caught via their expenditures on "concealment technologies". This point was made clear by Cremer and Gahvari (1994). They assumed that individuals have well-behaved quasi-linear preferences and derive their utilities from consumption and leisure. The average tax rate is progressive due to the existence of a guaranteed minimum income and the marginal tax rate is constant. An individual's likelihood of being audited, q, is exogenous. Assuming that the taxpayer spends on the concealment technology to affect the probability of being caught when audited, the probability (representing concealment technology) is given as $\delta \equiv \delta(m, \gamma, \gamma wL)$, where m is the expenditure on concealment, γ is the fraction of income concealed, L is the labor supply, the income (if not caught) is wL, where w is the wage rate. Further, $\frac{\partial \delta}{\partial \gamma}$ and $\frac{\partial \delta}{\partial(\gamma wL)}$ are nonnegative and at least one is strictly greater than zero, and $\frac{\partial \delta}{\partial m} < 0$. Assuming independence, clearly, the consumption in the "not being caught" state of the world occurs with probability $1 - \delta q$. Then consumption in the "not being caught" and "being caught" states are $wL - \theta(wL - \gamma wL) - m + s$ (as before, s here denotes the guaranteed income) and $wL(1-\theta) - F\theta\gamma wL - m + s$. These consumptions (along with leisure) enter the taxpayer's expected utility function to be maximized with respect to m, γwL and L. Note that by changing m the taxpayer can affect her marginal tax rate, too, making it more or less progressive. So, in essence, the relevant tax rate is an "expected" tax rate. The implication is "...in the presence of tax evasion, the burden of taxation on individuals is not just the taxes and fines paid to the government, but also the concealment cost associated with evading some of the taxes... [T]he concealment technology... plays a crucial role in determining the redistribution impact of tax evasion" (Ibid. p. 225-227). The effect of the tax rate on the evasion effort is not analyzed.

Yaniv (1999) took a completely different approach to analyzing the problem of tax evasion. He argued that the expected utility theory has to be abandoned in favor of the *prospect theory.* That is, the objective function should be the so-called *value function,* which allows the taxpayer to attach a specific value to a certain state of the world.[2] An important innovation is the tax advance scheme, in which the taxpayer pays a certain amount prior to filing the return that is used to offset later when the taxes are due. "[T]he tax advance is assumed to consist of two components: an endogenous one, which is a positive function of *estimated* tax liability...and an exogenous one determined by the tax agency as a part of its tax enforcement policy"

[2]The value function is assumed to be convex for losses and concave for gains.

(Ibid. p. 755-756). If one uses the expected utility theory, the tax advance payment will get cancelled out, and the problem becomes the one formulated by Allingham and Sandmo (1972). But in prospect theory approach, there is a value attached by the individual to each possible outcome. Each outcome is described by the *change* in net income, starting from an arbitrarily chosen reference point. For the reference point Yaniv takes the income before filing the return and after the tax advance is paid back to the tax collector, i.e., $W - D$, where D is the size of the tax advance payment. So, the changes in the net incomes in two possible states of the world are $D - \theta X$ and $D - \theta X - F\theta(W - X)$.

Note from above that the evader expects a certain net refund, $D - \theta X$, no matter whether she is caught or not, and expected penalty payments, $-F\theta(W - X)$, if she is caught. Formally, the agent tries to maximize the value function

$$V = v(D - \theta X) + qv(-F\theta(W - X)),$$

and v represents the value attached to the different outcomes and the first component is weighted by a probability equal to 1. The maximization problem leads to the comparative static results consistent with traditional findings. However, there is one important difference: as long as declaration is sufficiently high, an increase in the tax rate encourages evasion, unlike the Yitzhaki (1974) case. In the expected utility framework, the only remaining income effect was causing a counter-intuitive comparative statics result. However, in the prospect theory framework, the income effect actually consists of two parts: "...the certain refund component...and the uncertain penalty component...The latter component affects the 'value' in the loss domain where the taxpayer is risk-seeker, and where a tax rate increase, which increases the loss expected at a given level of evasion, compensatively discourages declaration" (Ibid. p. 761).

Prospect theory is not new to economics and as we have seen, can produce important results for the subject we are interested in. In general, *Behavioral Economics* is on the rise, especially within the field of Public Finance (see, e.g., McCaffery and Slemrod 2004). Importantly, many would agree that even in countries with a strong third-party reporting system, tax evasion is a winning proposition for many, who actually choose not to evade. "Why? Proposed solutions to this puzzle by definition involve pushing beyond the standard economic model, either by enriching it, in the utility functions, or by trying something else altogether" (Ibid. p. 17). The arguments in favor of Behavioral Economics sound rather compelling: after all, few would argue against the observation that many economic agents take into account moral sentiments, the subjective evaluation of the fairness of the tax code and even more than that, e.g., whether a certain government has a "moral" right to tax its fellow citizens. The latter seems especially relevant in countries with inefficient government and much antagonism between the government and the society. Thus, behavioral aspects and institutional settings become very important in the analysis of income tax evasion. And some behavioral aspects have been introduced to analyze the theory of tax evasion. Gordon (1989), Yaniv (1999) are just few examples.[3] Another way to model tax evaders is by making the assumption that they are acting in a boundedly rational way. That, in turn, can be modeled in a variety of different ways. However, this would go beyond the scope of our research interests in this study.

2.2. Dynamic Models of Tax Evasion.

Compared to static models, there are few dynamic models of income tax evasion. Predominantly, important dynamic tax evasion models started emerging relatively recently.

Often time researchers are interested in how the taxpayer plans her evasion profile

[3] An interested reader can also consult Al-Nowaihi and Pyle (in MacDonald and Pyle, 2000). We have not covered here their analysis because their model is just a generalization of Gordon's (1989) case.

over time, thus, one needs to move from a static to a dynamic environment. In this section we will describe a few major contributions to the literature. The dynamic aspect of tax evasion was actually analyzed by Allingham and Sandmo (1972) in the same paper which pioneered the theoretical static model of income tax evasion. The framework is a discrete multi-year one. The individual's income (fixed in each period) is normalized to unity. The authors expect that X_t, the amount of income declared in period t, is a positive fraction between zero and one. It is also assumed that if caught cheating in period t, the taxpayer's previous cheating, if any, will also be discovered. Therefore, in "not caught" and "caught" states of the world, the agent's income is $1 - \theta X_t$ and $1 - \theta X_t - \Pi \sum_{\tau=1}^{t} (1 - X_\tau)$, respectively. Note that the latter implies a fairly arbitrary penalty structure on all the previous undeclared incomes.[4] The expected utility (of the myopic agent, ignoring the future and taking past as given) to be maximized is as follows:

$$E[U_t] = (1-q)U(1-\theta X_t) + qU(1-\theta X_t - \Pi \sum_{\tau=1}^{t} (1-X_\tau)).$$

Interesting questions to answer are whether in some arbitrary period the taxpayer declares only a fraction of her true income and whether there exists a period, T, when she becomes absolutely honest in filing her tax returns. In addition, it is interesting to see whether evasion (if any) increases or decreases over time. Note that at $t = 1$ the problem reduces to the static case and thus, one can conclude that initial partial evasion to whatever extent is made. Simple algebraic manipulations can demonstrate that X_t does not converge to 1 asymptotically, that is, the taxpayer does not decide to declare all of her income at any time. The most interesting finding is that declaration increases over time. This is found by showing first that higher penalty rate encourages honesty. "The relevance of this to our problem is immediate, for the passage of time

[4]For additional details about that penalty rule consult p. 334 in Allingham and Sandmo (1972).

is equivalent to the increase of a fixed penalty" (Ibid. p. 335-336).[5] The analysis up to this point is applied to the myopic individual. However, it is also shown that the forward-looking individual, who realizes that by cheating today she places herself in unfavorable position in the future, will maximize the *lifetime* utility and always decide to declare more than the myopic taxpayer.

Russel et al. (1982) developed a discrete model where the taxpayer evades over the course of several years. Penalties are considered to be retroactive. This is because, once again, if evasion is detected in one year, it is detected in other years and therefore, the penalties will be imposed on the past tax fraud as well. The taxpayer maximizes her cumulative (over the years) expected monetary net gain from evasion. Assume that the taxpayer decides either not to evade (d_1 alternative) in each year, or evade (d_2 alternative). Monetary gain in year t due to the evasion in that year is g_t, and F_t is the corresponding penalty parameter. Expected values of d_1 and d_2 are thus $E_A(d_1) = 0$ (no evasion, no payoff, no penalty) and $E_A(d_2) = (1-q)g_t - qF_t$. Assuming risk-neutrality, the taxpayer evades if the former is greater than the latter which reduces to $g_t > \frac{qF_t}{1-q}$.

Now, in the second formulation, the taxpayer tries to maximize the cumulative net gain after t years. Then, the expected cumulative gains from d_1(evasion does not occur in year t only) and d_2 are $E_B(d_1) = (1+\varrho)G_{t-1}$ and $E_B(d_2) = (1-q)G_t + q\{(1+\varrho)G_{t-1} - F_t\}$, respectively. Here G_t is the cumulative gain after t years of undetected tax fraud, ϱ is the rate at which the illegal gains accrue. It is assumed that if evasion does not take place in year t, then any previous tax frauds are undetected. It can be shown that "...the decision, maximizing expected gain in year t is equivalent to the decision to maximize cumulative gain at the end of year t in the case of risk-neutral taxpayers" (Ibid. p. 380). Wage growth is also introduced into the model and the growth rate is r. Further, it is assumed that the agent can evade different amount at

[5]Recall that the agent, when discovered, has to pay the penalty also on all the previously evaded incomes.

each year; however, as a percentage of earned income it is always the same amount. And the model tries to find that optimal proportion, λ^*. Consequently, the gain becomes $g_t = \theta\lambda W(1+r)^{t-1}$. It follows then that $G_t = (1+\varrho)G_{t-1} + g_t$, $\forall t \geqslant 2$. Then the author proceeds by deducing F_t from the real-world evidences (a fraction of total evaded tax liabilities plus return of evaded taxes), assumes the linear tax schedule and obtains an explicit expression for $E_A(d_2)$, which is to be maximized with respect to λ. The main result is that if the taxpayer decides to evade in the first place, she will conceal her true income completely. Finally, a sufficiently high penalty rate equal to $\frac{1-q}{q}$ will render the expected gain from evasion negative.

Sengupta (1998) assumes no third-party reporting from investment income and thus, the individual, who cannot hide her wage income, can conceal only her investment income. The behavior of the agent is analyzed via a two-period overlapping-generations model and the utility function is additively-separable. Thus, the expected utility to be maximized is:

$$E[U] = U(c_t) + \eta[qV(e_{t+1}^d) + (1-q)V(e_{t+1}^{nd})],$$

where η is the positive rate of discount, being less than one, $c_t = (1-\theta)W - k_{t+1}$, is the consumption of a generation t young, e_{t+1}^d and e_{t+1}^{nd} are the consumptions of a generation t old individuals, when caught and not caught cheating, respectively. Here k_{t+1} is the total saving of the young in period t, which becomes the capital in the next period. Period t saving yields profit in the next period equal to output minus wages in the *next* period. The utility function is well-behaved and the individual is risk-averse. The individual decides in the first period how much to save and how much of profit to report. The production function is in the form $Y_t = f(K_t, L_t)$, where the arguments stand for capital and labor. A perfectly competitive environment is assumed and capital depreciation is ruled out. Note that the individual cannot cheat in the first period when she supplies her labor, but can cheat in the second period when she decides to consume out of her saving. The main results of the study are that the agent in the beginning period decides to report more of her profit to the tax authorities when the penalty parameter and the probability of getting caught are higher. Interestingly, an increase in the tax rate on wage income actually lowers saving. "The result could be explained as follows. The wage income tax is the tax that cannot be avoided... [The] after tax income in the first period is reduced, saving is reduced to compensate for the reduction in the income available for consumption" (Ibid. p. 428). Furthermore, an increase in the wage income tax rate increases cheating *only if* the individual's measure of absolute risk aversion is non-declining in second-period consumption. It is worth noting, though, that Sengupta does not provide any explicit analytical solution to his model since all the functions are in general form. In fact, to solve the model, one has to solve for eight endogenous variables from the system of eight equations, which makes it fairly complicated.

Lin and Yang (2001) modify the portfolio choice model of tax evasion by assuming perfect foresight and infinite time horizon. The dynamic problem is to maximize the household's instantaneous utility

$$U = \int_{t=0}^{+\infty} \exp(-\rho t)[(1-\phi)\ln c + \phi \ln G]dt,$$

where $0 < \phi < 1$, c is the consumption per capita, G is the level of public good, and ρ is the rate of time preference, used as the discount rate. With tax evasion, the level of public good is $\theta Ak - \overline{r}eAk = \theta(1-\overline{r}e)Ak$. Here A is the technology parameter, k is the per capita capital net of depreciation, and the production function (per capita output) is Ak, $\overline{r}$ is the expected rate of return on a dollar of evaded *taxes* $(= 1-q-qF))$, e is the portion of true income concealed.[6] The problem can be

[6]This is because the return is 1 with probability $1-q$ and $-F$ with probability q (Ibid, p. 1829).

solved via stochastic *Bellman equation.* By recognizing that capital accumulation is random due to the audit and using a standard *Brownian motion* (or *white noise*), it is shown that the individual's capital accumulation evolves according to the stochastic differential equation $dk = ([(1-\theta) + \overline{r}\theta e]Ak - c)dt + (eAk\sigma)dz$, where $\sigma > 0$ and z stands for the *Brownian motion* (Ibid. p. 1831 and pp. 1838-1839). Initial conditions at time $t = 0$ are also given. The authors present one of solutions of the model in terms of e with $\frac{\partial e}{\partial \theta} > 0$, meaning that the evasion increases with the tax rate. "The key here, of course, lies in extending the model from statics to dynamics... At any time t the agent is assumed to maximize the sum of the immediate payoff... plus the continuation value... By contrast there is no continuation value in the static setting and hence the agent's problem is to maximize the immediate payoff...Note that the positive expected... return on a dollar of concealed income... implies a favorable gamble faced by individuals. It is well known that a risk-averter takes no part of an unfavorable or barely fair gamble, but always takes some part of a favorable gamble... Facing a favorable gamble with infinite continuation as in our model, it is only reasonable that the degree to which an individual will take part in the favorable gamble varies directly from the expected return from the gamble" (Ibid. pp. 1832-1833).[7] Certainly, the expected return on evaded income is positively affected by higher tax rate.

It is worth noting that Lin and Yang (2001) were first to analyze income tax evasion in the context of the Stochastic Optimal Control. Their contributions are widely cited and their analytical results are compelling. Several modifications have also been suggested. Later we will return to their study and allow for some modifications.

Niepelt (2005) proposes an important modification (in the dynamic setting) to the standard benchmark model of tax evasion. He assumes first that penalties are increasing in the duration of an evasion spell and benefits of evasion increase linearly.

[7] The "continuation value" can be thought of as the expected present discounted value of future payoffs.

In addition, in the standard models of tax evasion it is implicitly assumed that evasion is either fully detected or not detected at all. In reality, though, the detection risk can spread for some sources of income, assuming, of course, that the taxpayer's income vector incorporates various sources. Even if the taxpayer is audited and has only one source of income, tax evasion may still not be detected entirely.

Further, the unit of analysis is not the taxpayer, but the unit of capital or the source of income. The unit of capital is analyzed in two states: when the income from a particular source is reported (state v_0) and when it is not (state v_1). Another interesting assumption is that there is a fixed cost to the taxpayer for switching between two states: "[s]uch a cost may arise... because an advisor has to be hired who knows how to convincingly make a case vis-à-vis the tax authority. Or it may arise because hiding a capital and making it reappear involves some transactions that temporarily reduce the return" (Ibid. p. 1616). If evasion is profitable, the capital will switch between two states. A unit of capital switched from a state to another state and remains there for a minimum time, $\overline{T}$. For example, the value of one unit of capital in state (v_0, t) is the the present discounted value of the payoffs from the unit:

$$k(t) = \int_{t}^{\overline{T}} \exp(-\rho(x-t))r(1-\theta)dx + e^{-\rho(\overline{T}-t)}k(\overline{T}),$$

where t stands for the time that has passed since the unit of the capital was last switched between states v_1 and v_0, r is the dividend yield for the units of undeclared and undetected capital, $r(1-\theta)$ stands for the dividend yield for honestly reported ones and x is the dummy of integration. The author then proceeds by deriving the value of the unit of capital in state v_1 and determines the so-called "optimal stoppage time" (using the *smooth pasting* condition), the time span after which it is optimal to switch between states v_1 and v_0. One of the main results of the model is that higher tax rate increases the optimal stoppage time. "...A higher statutory tax rate induces households to wait longer, and face higher expected fines before switching to reporting accrued income" (Ibid. p. 1619). Of course, this can be interpreted as the direct relationship between more aggressive taxation and evasion incentives, as income tax rate goes up. Thus, this continuous time dynamic model is also capable of resolving the Yitzhaki puzzle.

Conclusions.

In this chapter we have concisely reviewed some fundamental works on the theory of income tax evasion. They were classified between static and dynamic ones. As we see, the discussion over the inverse relationship between the tax rate and evasion is an important one in tax evasion literature. To say that the Yitzhaki puzzle dominates the literature is far from the truth. We saw that many works on income tax evasion mainly pursued the goal of removing the early paradoxical finding of Yitzhaki. And they indeed succeeded (e.g., Gordon 1989, Yaniv 1999, etc.). Most of them adopted significantly different assumptions compared to the original ones of Allingham, Sandmo, and Yitzhaki. Few others (we did not discuss) conducted some experimental studies to gain insights into the private response to higher tax rate, penalty, and probability parameters, and so on.[8]

However, not many studies were interested in the macroeconomic implications of an individual tax evasion. Evaded taxes are often easily converted into interest-earning assets, and thus, contribute to the long-run economic growth. Higher growth rate of per capita capital can be translated into higher growth rate of income per capita.[9] Thus, on the one hand, a point of interest, along with other things, could be the analysis of the paradoxical result in the framework close to the original benchmark one. This is what we will try to do later in this study. Further, we can try to make

[8]See, for example, Pudney et. al. (in MacDonald and Pyle, 2000, pp. 267-289). The tax experiment was conducted for 270 individuals in Turkey. The main results are: 1) higher tax rates encourage evasion; 2) younger people are more likely to engage in evasion.

[9]Lin and Yang (2001, pp. 1833-1835) is an interesting source of reference for this type of discussion.

inferences about national saving and other macroeconomic indicators in the presence of tax evasion.

Chapter Three: A Static Model of Income Tax Evasion.

3.1. The Benchmark Tax Evasion Model: Revisited.

One of the criticisms of the benchmark tax evasion models is that the tax rate schedule is proportional. "The assumption of a constant tax rate is questionable. A progressive income tax schedule needs to be incorporated in the analysis. A progressive income tax system will act as a further disincentive to the declaration of income in the absence of any counteracting incentive (e.g., in terms of an increased likelihood of punishment)" (Cullis and Jones 1998, p. 200). But it is well-known that the most important questionable theoretical conclusion of the benchmark tax evasion model is the inverse relationship between an incentive to cheat and the tax rate. All other results are theoretically intuitive (e.g., a higher fine rate and/or probability of being caught encourage compliance). Thus, we start this chapter by trying to asses the Yitzhaki puzzle by deviating from the original framework of Allingham and Sandmo (1972) as little as possible but incorporating a progressive tax system in the analysis.

Introduction.

Recall from section 2.1. that the benchmark model of tax evasion was developed by Allingham and Sandmo (1972), where the taxpayer's problem is to maximize the following VNM expected utility function:

$$E[U] = (1-q)U(W-\theta X) + qU(W-\theta X - F(W-X)).$$

Once again, here θ is the income tax rate, q is the probability of detection, X is the amount of declared income, F is the fine parameter, W is the true income. The key comparative-static result is that an increase in the tax rate has an ambiguous effect on the incentives to cheat due to competing income and substitution effects.

However, Yitzhaki (1974) noted that when the penalty is imposed on the amount of evaded taxes, rather than on evaded income, the substitution effect vanishes and only the income effect remains. Consequently, assuming decreasing absolute risk-aversion, higher tax rate makes people more honest in reporting their incomes. Note, in Yitzhaki's model the taxpayer's problem is to maximize

$$E[U] = (1-q)U(W-\theta X) + qU[W - \theta X - F\theta(W-X)].$$

Due to Yitzhaki's counter-intuitive finding, some have suggested complete abandonment of the state-preference approach to the tax evasion problem. Others have tried different tax evasion models to resolve the puzzle (e.g., Lin and Yang 2001).

The Model: a Simple Scenario.

In the state-preference framework, though, a more realistic model should treat the income tax rate as progressive, which will create further disincentives for compliance, *ceteris paribus*. To show the implications, we will derive the basic results of the model by adopting Yitzhaki's (1974) notations. Now the taxpayer's problem is to maximize (3.1) below

$$E[U] = (1-q)\,U(W - \theta\,(X)\,X) + qU(W - \theta\,(X)\,X - F\theta\,(W)\,(W-X)), \quad (3.1)$$

where the average income tax rate, θ, is a function of declared income, X, defined $\forall X \in (0, W]$, and $\frac{d\theta(X)}{dX}$ is assumed to be strictly positive and $\frac{d^2\theta(X)}{dX^2}$ is at least equal to zero.[1] That is, the tax schedule is progressive in terms of the average tax rate. The total tax to be paid is a positive, increasing, convex function of income.

[1]Note that $\theta(X)$ is the average tax rate. The assumption that $\frac{d\theta(X)}{\theta(X)}$ is positive just comes from the definition of the progressive income tax schedule. However, whether $\frac{d^2\theta(X)}{dX^2}$ is at least zero is an empirical question. Furthermore, the evaded taxes are $\theta(W)W - \theta(X)X = \theta(W)(W - \frac{\theta(X)}{\theta(W)}X)$. But for the sake of simplicity we are using a simpler penalty structure. We will later consider a more realistic setup to show that our results will be essentially the same.

The first-order condition is:

$$\begin{aligned}\Phi \equiv \frac{\partial E[U]}{\partial X} &= -(1-q)\,U'(Y)\,[\frac{d\theta(X)}{dX}X+\theta(X)] \\ &-qU'(Z)[\frac{d\theta(X)}{dX}X+\theta(X)-F\theta(W)]=0, \qquad (3.2)\end{aligned}$$

where $Y \equiv W-\theta(X)X$ and $Z \equiv W-\theta(X)X - F\theta(W)(W-X)$, representing the after-tax income if evasion is not detected and detected, respectively. The second order condition is

$$\begin{aligned}D \equiv \frac{\partial^2 E[U]}{\partial X^2} &= -(1-q)\,[-U''(Y)\left(\frac{d\theta(X)}{dX}X+\theta(X)\right)^2 \\ &+U'(Y)\left(\frac{d^2\theta(X)}{dX^2}X+2\frac{d\theta(X)}{dX}\right)] \\ &-q[-U''(Z)\left(\frac{d\theta(X)}{dX}X+\theta(X)-F\theta(W)\right)^2 \\ &+U'(Z)\left(\frac{d^2\theta(X)}{dX^2}X+2\frac{d\theta(X)}{dX}\right)]<0. \qquad (3.2')\end{aligned}$$

Equation (3.2$'$) implies a unique solution for utility maximization. The conditions for an interior solution are

$$\frac{\partial E[U]}{\partial X}|_{X=0} = -(1-q)U'(W)\theta(X) > q(\theta(X)-F\theta(W))U'(W(1-F\theta(W))), \quad (3.3)$$

and

$$\frac{\partial E[U]}{\partial X}|_{X=W} = -(1-q)\,U'(W(1-\theta(W)))[\frac{d\theta(W)}{dX}W+\theta(W)]$$

$$-qU'(W(1-\theta(W)))[\frac{d\theta(W)}{dX}W+\theta(W)-F\theta(W)]<0. \quad (3.4)$$

And consequently,

$$\frac{U'(W)\theta(X)}{U'(W(1-F\theta(W)))} < \frac{q(\theta(X)-F\theta(W))}{1-q}, \quad (3.3')$$

and

$$qF\theta(W) < \theta(W) + \frac{d\theta(W)}{dX}W. \tag{3.4'}$$

Logically, the average tax rate *evaluated at income equal to* W is unaffected by the change in X. Thus, the last term in $(3.4')$ vanishes, meaning that we have $qF < 1$ and this is exactly Yitzhaki's condition for an interior solution (Yitzhaki 1974, p. 201).

Comparative Statics.

By applying the *Implicit Function Theorem* (IFT) to (3.2), the following comparative static results can be obtained:

$$\begin{aligned}\frac{\partial X}{\partial F} &= \frac{q}{D}[-U''(Z)\,\theta(W)(W-X)\left(\frac{d\theta(X)}{dX}X + \theta(X) - F\theta(W)\right) \\ &\quad -U'(Z)\,\theta(W)].\end{aligned} \tag{3.5}$$

From (3.2) the expression $\frac{d\theta(X)}{dX}X + \theta(X) - F\theta(W) < 0$ (otherwise $\Phi \neq 0$), which makes (3.5) positive. Analogously,

$$\begin{aligned}\frac{\partial X}{\partial q} &= -\frac{1}{D}[U'(Y)\left(\frac{d\theta(X)}{dX}X + \theta(X)\right) \\ &\quad -U'(Z)\left(\frac{d\theta(X)}{dX}X + \theta(X) - F\theta(W)\right)].\end{aligned} \tag{3.6}$$

Clearly, (3.6) is positive as well. Equations (3.5) and (3.6) demonstrate that higher probability of detection and fine rate increase the amount of declared income.

However, the effect of the change in the tax rate on incentives to cheat is not that straightforward. Let us present the average tax rate as $\theta(X) + h$, $\forall X \in [0, W]$, where h is a shift parameter. For instance, the government may decide to increase the average tax rate across all income brackets by say, 2%. Thus, we can use IFT once again by first differentiating Φ with respect to h. Then, we evaluate the result at $h = 0$ and divide the resulting expression by the second-order condition, D. That is, as before we are using here the Implicit Function Theorem. Therefore[10],

$$\begin{aligned}
\frac{dX}{dh}|_{h=0} &= \frac{1}{D}[(1-q)\left(-U''(Y)X\left(\frac{d\theta(X)}{dX}X+\theta(X)\right)+U'(Y)\right) \\
&\quad +q\left(-U''(Z)(X+F(W-X))\left(\frac{d\theta(X)}{dX}X+\theta(X)-F\theta(W)\right)\right) \\
&\quad +(1-F)U'(Z)]. \qquad (3.7)
\end{aligned}$$

Let us have a closer look at condition (3.7). It is clear that we cannot sign it.

More compactly, (3.7) can be restated as follows:

$$\begin{aligned}
\frac{dX}{dh}|_{h=0} &= \frac{1}{D^*}[(1-q)U'(Y)\left(R_A(Y)X\left(\frac{d\theta(X)}{dX}X+\theta(X)\right)+1\right) \\
&\quad +qU'(Z)\left(R_A(Z)(X+F(W-X))\left(\frac{d\theta(X)}{dX}X+\theta(X)-F\theta(W)\right)\right) \\
&\quad +1-F]. \qquad (3.7')
\end{aligned}$$

where $R_A(Y) \equiv -\frac{U''(Y)}{U'(Y)}$ and $R_A(Z) \equiv -\frac{U''(Z)}{U'(Z)}$. These are the absolute risk aversion functions. Since $Y > Z$ and assuming decreasing absolute risk aversion, we can conclude that $R_A(Y) < R_A(Z)$. Assuming diminishing marginal utility of income, $U'(Y) < U'(Z)$. Observe also that $R_A(Z)(X + F(W - X)) > R_A(Y)X$. Nevertheless, we cannot get rid of the ambiguity in the overall expression. This differs from the Yitzhaki (1974) result and confirms the original finding of Allingham and Sandmo (1972).

3.2. A More Realistic Specification.

Note that the tax cheater's unpaid tax liability is calculated as $\theta(W)W - \theta(X)X$.

[10] Another clarification is worth making here. An increase in the shift parameter implies an increase in the *marginal* tax rate as well. Note that when we write the average tax rate function as $\theta(X) + h$ (and thus, the total tax liability $(\theta(X) + h)X$), then the marginal tax rate becomes $\frac{d\theta(X)}{dX}X + \theta(X) + h$. Clearly, if we shift h, we shift also the marginal tax rate. This is very important because the puzzle involves an inverse relationship between the *marginal* tax rate and the amount of evaded income.

Then, the problem is to maximize

$$E[U] = (1-q)\,U(W-\theta\,(X)\,X) + qU(W-\theta\,(X)\,X - F\,(\theta(W)W - \theta(X)X)), \quad (3.8)$$

where, as before, the income tax rate, θ, is a function of declared income, X, and $\frac{d\theta(X)}{dX}$ is assumed to be strictly positive.

The first-order condition is:

$$\begin{aligned} \overline{\Phi} \;&\equiv\; \frac{\partial E[U]}{\partial X} = 0 = -\,(1-q)\,U'\left(\overline{Y}\right)[\frac{d\theta\,(X)}{dX}X + \theta\,(X)] \\ &-qU'(\overline{Z})[(1-F)\left(\frac{d\theta(X)}{dX}X + \theta(X)\right)], \end{aligned} \quad (3.9)$$

where $\overline{Y} \equiv W - \theta\,(X)\,X$ and $\overline{Z} \equiv W - \theta\,(X)\,X - F\,(\theta(W)W - \theta(X)X)$. The second order condition is

$$\begin{aligned} \overline{D} \;&\equiv\; \frac{\partial^2 E[U]}{\partial X^2} = -\,(1-q)\,[-U''\left(\overline{Y}\right)\left(\frac{d\theta\,(X)}{dX}X + \theta\,(X)\right)^2 \\ &+U'\left(\overline{Y}\right)\left(\frac{d^2\theta\,(X)}{dX^2}X + 2\frac{d\theta\,(X)}{dX}\right)] \\ &-q[-U''\left(\overline{Z}\right)\left((1-F)\left(\frac{d\theta(X)}{dX}X + \theta(X)\right)\right)^2 \\ &+U'\left(\overline{Z}\right)\left((1-F)\left(\frac{d^2\theta\,(X)}{dX^2}X + 2\frac{d\theta\,(X)}{dX}\right)\right)] < 0. \end{aligned} \quad (3.9')$$

From the first-order condition it can be shown that (3.2$'$) is negative. Thus, the second-order condition holds. Consequently, the conditions for an interior solution are as follows

$$\frac{U'\,(W)}{U'(W(1-F\theta(W)))} < \frac{q(F-1)}{1-q} \quad (3.10)$$

and

$$qF < 1 \quad (3.11)$$

Both equations (3.10) and (3.11) are identical to the Yitzhaki conditions (1974,

p. 201).

As before, we can establish that

$$\begin{aligned} \frac{\partial X}{\partial F} &= \frac{q}{D}[-U''(Z)(\theta(W)W - \theta(X)X)(1-F)\left(\frac{d\theta(X)}{dX}X + \theta(X)\right) \\ &\quad -U'(Z)\left(\frac{d\theta(X)}{dX}X + \theta(X)\right)] > 0, \end{aligned} \tag{3.12}$$

and

$$\begin{aligned} \frac{\partial X}{\partial q} &= -\frac{1}{D}[U'(Y)\left(\frac{d\theta(X)}{dX}X + \theta(X)\right) \\ &\quad -U'(Z)(1-F)\left(\frac{d\theta(X)}{dX}X + \theta(X)\right)] > 0. \end{aligned} \tag{3.13}$$

Finally,

$$\begin{aligned} \frac{dX}{dh}|_{h=0} &= \frac{1}{D}[(1-q)\left(-U''(Y)X\left(\frac{d\theta(X)}{dX}X + \theta(X)\right) + U'(Y)\right) \\ &\quad +q\left(-U''(Z)(X + F(W-X))(1-F)\left(\frac{d\theta(X)}{dX}X + \theta(X)\right)\right) \\ &\quad +(1-F)U'(Z)]. \end{aligned} \tag{3.14}$$

We have just seen that most of the results of our simpler model specification hold here, too. Namely, we obtained a positive relationship between the amount of declared income and enforcement variables. The original ambiguity nature of the Allingham and Sandmo (1972) is also restored. One issue lies in making the model more explicit, and seeing which way the ambiguity in (3.14) will go once we experiment with a variety of conventional functional forms. An interesting question is how progressive a tax structure should be in order to remove the Yitzhaki puzzle once the model is formulated more explicitly.

Furthermore, it will be useful, perhaps, to see how the progressiveness in the income tax rate affects the relationship between the tax rate and compliance. Interestingly, for instance, based on the econometrics study of Belgian data (see Geeroms and Wilmots 1985), the Yitzhaki result holds, while most of the studies of the U.S. economy reveal the opposite conclusion (see, e.g., Clotfelter 1983). There have been several studies which attempted an international comparison of tax progressivity by constructing what is called *Kakwani* (1976) *Index* (based on the Lorenz curve of gross income and the concentration curve of net tax liabilities). According to Wagstaff and Doorslaer (2001, p. 310), net tax liabilities are less progressive in Nordic counties (due to their reliance on proportional local income tax) than in the United States. In general, major OECD countries can be presented in the following descending order of progressivity of net tax liabilities on taxable incomes: France, Netherlands, Spain, Canada, Australia, Belgium, Germany, Ireland, Italy, U.K., Finland, U.S., Norway, Sweden and Denmark. It follows from the list that the Belgian tax code is relatively more progressive than that of the U.S.

Conclusions.

It has been argued that the presence of a constant tax rate is one of the most important shortcomings of tax evasion models. We showed that a particular progressive tax rate schedule indeed resolves what is called the Yitzhaki puzzle, according to which a higher tax rate would unambiguously encourage compliance. We were able to ensure an ambiguous theoretical relationship between the tax rate and income compliance in the context of the original Allingham-Sandmo (1972) framework by simply making the tax code progressive in nature. The advantage of our approach is that it is capable of generating the ambiguous relationship between the tax rate and income declaration without appealing to stigma costs, labor-leisure decisions, loss-aversion, infinite-planning horizon or any other significant augmentations of the basic tax evasion model.

Chapter Four: A Dynamic Overlapping-generations Tax Evasion Model.

4.1. Tax Evasion and Overlapping Generations.

"Inter-generational transfers are today at the center of the economic policy debate. The reduction in public debt, the financing of social security (pensions), the taxation of capital and bequests, and the design of the education system all imply substantial inter-generational transfers. The tool economists provide to analyze these issues is the overlapping-generations [OLG] model... [I]t is the natural framework to study the allocation of resources across the different generations" (De La Croix and Michel 2002, p. xiii). OLG models are the popular ones in the field of Macroeconomics, Public Finance, Environmental Economics, Population Economics, and Development Economics, and so on. They are very flexible and certainly can be applied for the study of income tax evasion.

But why an OLG model of tax evasion? There are two main reasons for modeling tax evasion behavior in the context of overlapping generations. First, and foremost, is to demonstrate the fundamental difference between static and dynamic models. In basic static models, the state of the economy is always fixed. But this is not true in a dynamically evolving economic environment. We want to show that one promising way to solve the Yitzhaki puzzle is to move from a static to a dynamic set-up. After all, when the tax rate goes up, the taxpayer, according to Yitzhaki, tends to cheat less. But what if that behavior affects the state of the economy in a way which would encourage her to comply less? Static models do not allow such a feedback effect and not surprisingly, therefore, researchers prefer to introduce many new assumptions in the model, to tackle the Yitzhaki's problem. However, to keep the modeling simple, we will consider only a two-period dynamic OLG model. One might counterargue that two-period model does not produce enough "dynamism". After all, it is well-known that in a basic two-period OLG framework, everything is "fixed" per-period (which corresponds to half-a-lifetime!), and savings becomes productive only in the second period. "This unrealistic lag structure is an unfortunate by-product of overlapping-generations models with only two periods of life" (Barro and Sala-i-Martin 2004 p. 193). But as we have already mentioned, our main goal is not to resolve the Yitzhaki puzzle, but to analyze it in a tractable dynamic setting. Even a two-period model is sufficient to achieve our goal. We will try to analytically show that once dynamism is introduced, the Yitzhaki puzzle might disappear. Of course, it might be more likely so with "enough" dynamism, as recent more realistic *multi-period* OLG models would provide.

Second, researchers on tax evasion literature often complain that given realistic enforcement parameters, cheating sounds like a winning proposition for many who do not evade. After all, when the probability of getting caught is only 1% (for the average U.S. taxpayer), and the penalty rate is, say, 1.5, the expected return on cheating is close to 100%, not bad, as one may wonder. And here is the "power" of a two-period OLG model. In essence, a taxpayer in such a model does not make a decision about cheating at every instant, keeping in mind the audit rate of 1%. Rather, in a two-period discrete setting she asks herself a different question: given the enforcement parameters, what fraction of my *overall, per-period* income should I report to the government (assuming that I am not cheating when old)? But if we set the problem that way, 1% yearly likelihood of getting caught might become as high as 30%![11] Thus, no longer do we have to assume a "low" chance of getting caught, which, hopefully, may generate more plausible empirical conclusions.

To sum-up, we pursue two main objectives in this chapter. First, we will develop a new rigorous, but yet analytically tractable dynamic model of income tax evasion

[11] In a simple scenario, assume that there are 100 young individuals living and working, and each year 1 is caught cheating. Then in 30 years (assuming that none is caught twice), 30 people will be caught. So, if the taxpayer makes her decision "per-period", she should decide that there is 30% chance of apprehension.

and obtain the closed-form solutions. The model can be used to analyze the impact of various policy changes (such as tax rates) on the tax compliance decision. Second, we aim to incorporate the tax compliance costs, faced by the taxpayer and see how that impacts her consumption and evasion behavior. We make the model flexible enough to study some real-world problems, such as the impact of falling audit rate in the U.S. economy and its welfare implications.

Introduction.

We will develop a two-period OLG model in the subsequent sections, and obtain the closed-form analytical solutions. We will focus on the dynamics of the economy, and analyze theoretical underpinnings of the possibility of dynamic inefficiency. We will also briefly discuss the problem of taxation for the revenue-maximizing government, and introduce simple exogenous costs of compliance. As far as the latter is concerned, it can help us differentiate between various agents with various net income, since those with higher costs of compliance will end-up having lower net disposable earnings, *ceteris paribus.*

Later in this work we will also apply the model to study the welfare implication of the recent trend observed in the U.S., namely, the phenomenon of a dramatically falling audit rate, during 1996-2000 (U.S. General Accounting Office 2001, p. 2).[12] Namely, we will show that different generations will be affected differently due to a decline in the audit rate, and it might be even the case that some taxpayers, who are never caught cheating (and enjoy higher disposable income because of lower compliance), may become worse-off (in terms of life-cycle utility) than their otherwise identical "more honest" predecessors. Hence, we will focus on these and similar issues in the subsequent sections.

[12]Interestingly, the audit rate for both low-income and high-income earners decreased and the overall drop was about 70%, from 1.67% in year 1996 to 0.49% in year 2000. However, according to the IRS Commissioner Mark W. Everson, the situation has been slightly improving since then. For instance, in 2004, compared to years 2002 and 2003, there was 36 and 19 percents jump in the audit rates, respectively.

The Model.

Time is divided into discrete periods, t, where $t \in [0, +\infty)$. Assume that at each time period a certain number of people are born who live and work, plan up to the next period, when they eventually die. That means there is a turnover in population. The model specifications presented below will be in the manner of Samuelson (1958) and Diamond (1965) frameworks. The main difference is that we consider a dishonest economic agent, who, along with her consumption decision, also decides to what extent to understate her income.[13]

The economy's supply side is represented by a large number of firms and described by a *Harrod-neutral* (labor-augmenting) constant returns to scale and twice continuously differentiable production function, $F(K_t, A_t L_t)$, defined on $\mathbb{R}_{++}$. Here K and L stand for capital and labor, respectively, and A is the "effectiveness" of labor, describing the state of the technology. Assume that after the production process, the part of the capital which is not depreciated is identical to the good produced, meaning the total economy-wide production capacity is $F(K_t, A_t L_t) + (1 - \varkappa)K_t$, where $\varkappa$ is per-period depreciation rate.[14] Hereafter, we will assume that the capital stock depreciates fully during the production process, i.e., $\varkappa = 1$.[15]

Assume that A and L both grow at exogenous rates g and n, correspondingly.[16] Therefore, $L_{t+1} = (1+n)L_t$ and $A_{t+1} = (1+g)A_t$.

In a perfectly competitive environment inputs will earn their marginal contributions to the output and firms break-even. Then, labor's value marginal product (=real wage) is $W_t = (f(k_t) - k_t f'(k_t))A_t$, where $k_t \equiv K_t / A_t L_t$, i.e., the amount of capital per unit of effective worker. Here $f(k) > 0$, $\forall k$ is the production function in

[13]For a review of the Diamond model see, e.g., Romer (2001, pp. 75-90). De La Croix and Michel (2002) has comprehensive coverage of various OLG models as well.

[14]See De La Croix and Michel (2002, p. 4).

[15]If one period lasts 30 years, a small annual depreciation rate of 2.5% causes more than a half of the capital stock to depreciate after one period.

[16]It is important to keep in mind that these are not annual growth rates, but growth rates per generation. For all practical purposes we will assume that n and g are both positive.

intensive form, satisfying the *Inada conditions*: $\lim_{k \to 0} f'(k) = +\infty$ and $\lim_{k \to +\infty} f'(k) = 0$. Also, $f'(k) > 0$, $f''(k) < 0$, $\forall k$. If we denote the real wage per effective worker, $f(k_t) - k_t f'(k_t)$, as w_t, labor's wage per worker becomes $w_t A_t$.

Assume an individual (a taxpayer) has a well-behaved time additively-separable and twice continuously differentiable lifetime VNM expected utility function, $E[U]$, defined on $\mathbb{R}_{++}$. Individual's preferences satisfy the following property: $\frac{\partial U}{\partial C}(\equiv U'(C)) > 0$, $\frac{\partial^2 U}{\partial C^2} < 0$, $\forall C > 0$, where C stands for consumption in general. All individuals live for two periods, working only in the first. Denote the consumption in period t of young and old individuals as C_{1t} and C_{2t}, respectively. In period t the young earns the real wage, $w_{1t} A_t$, by inelastically supplying a unit of labor (which, along with the capital stock, is used to produce output) and consequently, faces the tax rate $\theta \in (0, 1)$.[17] The taxpayer at time t decides how much to consume and how much out of her income to declare, X_{1t}. It is assumed that $\frac{\partial E[U]}{\partial X_{1t}} |_{X_{1t}=0} > 0$ and $\frac{\partial E[U]}{\partial X_{1t}} |_{X_{1t}=w_{1t}A_t} < 0$. The (subjective) exogenous probability of detection is q. Furthermore, if caught cheating at time t, the individual faces the penalty rate, $F \in (1, +\infty)$, imposed on evaded tax liabilities (not on evaded income).[18]

When retired in period $t+1$, the old individual consumes out of her saving, the return *factor* of which from t to $t+1$, is R_{t+1} (the latter is simply 1+ *the interest rate*, or $1 + r_{t+1}$). It is assumed that the effective profits, Π_t, are distributed to the old households, who are also the owners of the capital stock. There is no tax on capital. It is straightforward to show that $\Pi_t \equiv F(K_t, A_t L_t) - \frac{\partial F(K_t, A_t L_t)}{\partial L_t} L_t$, collapses to $f'(k_t) K_t$. Since from $t-1$ to t old households receive $R_t K_t$, the return factor on saving made at $t-1$ becomes $f'(k_t)$.

We rule out any bequest motive. In each period people make their livings by supplying their labor or consuming out of saving. By the end of the next period all old

[17]Here w_{1t} carries the same meaning as w_t above, the only difference is in the subscript. This is to stress the earning of the young.

[18]It is widely accepted that $F < 2$ for most countries (see, e.g., Al-Nowaihi and Pyle 2000, p. 265). Further, we also assume in our model that the penalties will be paid during the working period.

individuals exit the model. We assume that individuals are identical within cohorts. At a given date, producing firms would hire labor for production and investing firms accumulate saving from the young and build up the capital stock for the next period.

Note that we assume the government sector is unproductive. In reality, the government provides public services which might be utility-enhancing for many. An individual taxpayer (a 'price-taker'), thus will have an incentive to free ride by evading more, and still receiving nearly the same flow of benefits from the public goods. That kind of free-riding can be linked to what is called 'tragedy of the commons'.

Under our assumptions, the individual's utility-maximization problem can be presented as follows:

$$\underset{\{C_{1t}, X_{1t}\}}{Max} E[U] \equiv U(C_{1t}) + \frac{1}{1+\rho}\left[(1-q)U(C_{2t+1}^{nc}) + qU(C_{2t+1}^{c})\right], \tag{4.1}$$

where $\frac{1}{1+\rho}$ is the psychological discount factor and ρ is the rate of time-preference ($\rho > 0$). $C_{2t+1}^{nc} \equiv R_{t+1}(w_{1t}A_t - \theta X_{1t} - C_{1t})$ and $C_{2t+1}^{c} \equiv R_{t+1}(w_{1t}A_t - \theta X_{1t} - F\theta(w_{1t}A_t - X_{1t}) - C_{1t})$, where "nc" and "c" stand for "not caught" and "caught", respectively.[19] Observe that, from the way we have defined the maximization problem, it follows that C_{1t} and $U(C_{1t})$ do not vary with state of the world. This might be difficult to acknowledge. However, stating the problem that way makes the model tractable. A similar approach has been undertaken also by Sengupta (1998, p 421) and Atolia (2003, p. 4).[20]

Problem (4.1) can be solved by substituting the expressions for C_{2t+1}^{nc} and C_{2t+1}^{c} into the objective function and then implementing the first-order necessary conditions with

[19]We assume that government purchases do not affect the taxpayer's decision. One possible explanation is that the utility is additively-separable, and the taxpayer takes government purchases as given in her decision-making process.

[20]Further, note that ex-post, if the agent gets caught, she will potentially borrow in the first period to pay the fine. And that might lead to the negative aggregate capital stock. However, as long as the total capital stock remains positive (which is likely to be since in reality only a small fraction of people end-up paying fines to the tax authorities), we do not have to impose a borrowing constraint. However, with logarithmic preferences one has to be careful. Our parameterization, though, does not lead to the negative consumption in period two.

respect to the choice variables, C_{1t} and X_{1t}. The resulting marginal relations are to be solved simultaneously to determine the optimal values, X_{1t}^* and C_{1t}^*.

Clearly, then, the maximization of (4.1) is equivalent to (4.1$'$) below:

$$\begin{aligned} \underset{\{C_{1t}, X_{1t}\}}{Max}\ E[U] &\equiv U(C_{1t}) + \frac{1}{1+\rho}[(1-q)U(R_{t+1}(w_{1t}A_t - \theta X_{1t} - C_{1t})) \\ &\quad + qU(R_{t+1}(w_{1t}A_t - \theta X_{1t} - F\theta(w_{1t}A_t - X_{1t}) - C_{1t}))]. \quad (4.1') \end{aligned}$$

The corresponding derivative with respect to C_{1t} gives us the following:

$$\begin{aligned} \frac{\partial E[U]}{\partial C_{1t}} &= U'(C_{1t}) + \frac{1}{1+\rho}[-(1-q)R_{t+1}U'(R_{t+1}(w_{1t}A_t - \theta X_{1t} - C_{1t})) \\ &\quad -qR_{t+1}U'(R_{t+1}(w_{1t}A_t - \theta X_{1t} - F\theta(w_{1t}A_t - X_{1t}) - C_{1t}))]. \end{aligned}$$

After setting the expression above equal to zero we get

$$\begin{aligned} U'(C_{1t}) &= \frac{R_{t+1}}{1+\rho}[(1-q)U'(R_{t+1}(w_{1t}A_t - \theta X_{1t} - C_{1t})) \\ &\quad +qU'(R_{t+1}(w_{1t}A_t - \theta X_{1t} - F\theta(w_{1t}A_t - X_{1t}) - C_{1t}))]. \quad (4.2) \end{aligned}$$

Analogously,

$$\begin{aligned} \frac{\partial E[U]}{\partial X_{1t}} &= \frac{1}{1+\rho}[-(1-q)\theta R_{t+1}U'(R_{t+1}(w_{1t}A_t - \theta X_{1t} - C_{1t})) \\ &\quad -q(\theta - F\theta)R_{t+1}U'(R_{t+1}(w_{1t}A_t - \theta X_{1t} - F\theta(w_{1t}A_t - X_{1t}) \\ &\quad -C_{1t}))] \\ &= 0. \end{aligned}$$

Here we can divide both sides of the preceding expression by $-\frac{\theta R_{t+1}}{(1+\rho)}$ and then bring the second term to the right-hand side. Thus, we obtain

$$\begin{aligned} & (1-q)U'(R_{t+1}(w_{1t}A_t - \theta X_{1t} - C_{1t})) \\ = \; & q(F-1)U'(R_{t+1}(w_{1t}A_t - \theta X_{1t} - F\theta(w_{1t}A_t - X_{1t}) - C_{1t})). \end{aligned} \tag{4.3}$$

Suppose that $U(C) \equiv \ln(C)$, i.e., the coefficient of relative risk aversion is 1. In fact, taking the utility function of the natural log variety is consistent with empirical evidence (Attanasio 1999). Then, (4.2) becomes

$$\begin{aligned} \frac{1}{C_{1t}} = & \frac{1}{1+\rho}[(1-q)\frac{1}{w_{1t}A_t - \theta X_{1t} - C_{1t}} \\ & + q\frac{1}{w_{1t}A_t - \theta X_{1t} - F\theta(w_{1t}A_t - X_{1t}) - C_{1t}}]. \end{aligned} \tag{4.2$'$}$$

From (4.3) it is clear that

$$\frac{1-q}{w_{1t}A_t - \theta X_{1t} - C_{1t}} = \frac{q(F-1)}{w_{1t}A_t - \theta X_{1t} - F\theta(w_{1t}A_t - X_{1t}) - C_{1t}}. \tag{4.3$'$}$$

After a series of algebraic manipulations, we find the optimal income declaration and consumption profiles:

$$X_{1t}^* = \frac{F\theta(2-q) + \theta\rho(F-1) + qF - \theta - 1}{\theta(F-1)(2+\rho)} w_{1t}A_t \tag{4.4}$$

and

$$C_{1t}^* = \frac{(1+\rho)(1-\theta)}{2+\rho} w_{1t}A_t. \tag{4.5}$$

That is, the behavior of an economic agent within the context of our model is fully described by equations (4.4) and (4.5). Observe that when $\theta = 0$, $C_{1t}^* = w_{1t}A_t(1+\rho)/(2+\rho)$, as in the basic form of the Diamond model. The taxpayer in time t decides how much to consume and how much of her income to declare to the tax authorities, given the parameters of the model. Note that (4.4) and (4.5) are independent of the interest rate as it is the case with logarithmic preferences. The usage of the logarithmic preferences is conventionally accepted in the literature. However, one may argue that in reality, since evaded taxes go into investment, the interest rate may be an important factor to consider. Before we proceed to the next section, we will state the following proposition.

Proposition 1 *To ensure absolute compliance, government should set* $qF = 1$*. As long as* $qF < 1$*, at least some evasion will take place.*

Proof. Indeed, if we require an interior solution (to satisfy two early assumptions $\frac{\partial E[U]}{\partial X_{1t}} |_{X_{1t}=0} > 0$ and $\frac{\partial E[U]}{\partial X_{1t}} |_{X_{1t}=w_{1t}A_t} < 0$), $\frac{F\theta(2-q)+\theta\rho(F-1)+qF-\theta-1}{\theta(F-1)(2+\rho)}$ has to be a positive fraction (refer to (4.4)). As for the fraction, the numerator is to be smaller than the denominator. After rearranging and simplifying, we obtain $qF < 1$. If we take $qF = 1$ and substitute it into the previous ratio, we will have $\frac{2F\theta-2\theta+\theta\rho(F-1)}{\theta(F-1)(2+\rho)}$. The latter collapses to 1 after some algebraic rearrangements, implying $X_{1t} = w_{1t}A_t$.[21] ■

Not surprisingly, we obtain the same condition as Yitzhaki (1974 p. 201) did for the interior solution. An important clarification is worth making here, that is, $X_{1t}^* < w_{1t}A_t$ may hold even if the taxpayer chooses to *evade* for certain *year*(s) and *over-report* her income for some other *year*(s). In essence, our condition for an interior solution is a "*per period*" one.

Further, to ensure that X_{1t} is positive, we state the following assumption.

Assumption 1. $F\theta(2-q) + \theta\rho(F-1) > 1 + \theta - qF$.[22]

Obviously, one can conclude with a closer inspection of (4.5) that $0 < \frac{(1+\rho)(1-\theta)}{2+\rho} < 1$. Hence, the agent's marginal propensity to consume belongs to $(0, 1)$ interval.

[21] Observe that if the tax rate is 100%, there will be absolute compliance, too. This can be explained by the fact that higher tax rate encourages compliance by increasing expected penalty payments. In essence, our model predicts that if the tax rate is 100%, the intolerably high penalty payments (when caught) would discourage the taxpayer from cheating. But for all practical purposes we can assume that the tax rate is never set that high.

[22] Kolm (1973) asserts that one can report more than the true income, even if this possibility is not the best choice. Such a possibility indeed exists when one, say, miscalculates her tax liability but it is hard to imagine that the taxpayer, trying to cheat, would deliberately declare more than what she earns. She could make a mistake, of course, but for now we assume it away.

Note that the key difference between the traditional static models and ours is that we do not treat wage income as a constant. An increase in one of the model parameters may affect the state of the economy (represented by the capital per effective worker), which, in turn, will affect $w_{1t}A_t$. So, before any comparative dynamics are analyzed, we need to find the explicit expression for k_t, which determines the real wage per worker. This will be addressed in subsequent sections.

Implications for Saving and Capital Accumulation.

In each period some taxpayers' returns are audited, and as a result some are convicted with tax fraud. We assume that the "conviction rate" is β $(\beta \leq q)$, and those who are caught cheating end-up paying the penalties to the tax-collecting agency. It would be reasonable to assume that β is an increasing function of q. We will simplify the analysis by supposing that q and β are independent. But still, one may guess that q must be equal to β. After all, if people know what the conviction rate is, why would they not set their subjective evaluation of getting caught to β? That line of reasoning is sensible. However, some taxpayers may perceive "too" high audit rate especially in the countries with the strong third-party reporting system. Most studies assume that the perceived likelihood of getting caught is around $1-3\%$, which is in line with our parameterization below. Taking a realistic value for the conviction rate can also help us to estimate the capital stock of the economy more accurately.

Let us denote the total saving of the group of people who are not caught cheating as S_{nc}, and that of the group of people who are caught as S_c, with $S_{nc} = (1-\beta)L_t\left[w_{1t}A_t - \theta X_{1t} - C_{1t}\right]$, and $S_c = \beta L_t[w_{1t}A_t - \theta X_{1t} - C_{1t} - F\theta(w_{1t}A_t - X_{1t})]$. By making corresponding substitutions from (4.4) and (4.5), and simplifying, we obtain

$$S_{nc} = \left[\frac{L_t F(1-\theta)(1-q)(1-\beta)}{(F-1)(2+\rho)}\right] w_{1t}A_t. \tag{4.6}$$

By recalling the similarities in the expressions for C^{nc}_{2t+1} and C^{c}_{2t+1}, utilizing (4.4) and

(4.5) again, we obtain

$$S_c = \left[\frac{L_t F(1-\theta)q\beta}{2+\rho}\right] w_{1t} A_t. \tag{4.7}$$

Equations (4.6) and (4.7) give us the overall saving for two groups of people: those who are not caught cheating and those who are caught cheating. Total saving in the economy is $S \equiv S_{nc} + S_c$ and this generates the capital stock in period $t+1$. This implies that

$$S \equiv K_{t+1} = \left[\frac{F(1-\theta)(1-q-\beta(1-qF))}{(F-1)(2+\rho)}\right] L_t w_{1t} A_t.^{23} \tag{4.8}$$

Now we can divide both sides of (4.8) by $L_{t+1}A_{t+1}$. As a result,

$$k_{t+1} = \left[\frac{F(1-\theta)(1-q-\beta(1-qF))}{(F-1)(2+\rho)(1+n)(1+g)}\right] w_{1t},$$

where $k_{t+1} \equiv \frac{K_{t+1}}{L_{t+1}A_{t+1}}$ is capital per effective worker. Recognizing that w_{1t} is the total output per effective worker less the share of the capital, we obtain

$$k_{t+1} = \left[\frac{F(1-\theta)(1-q-\beta(1-qF))}{(F-1)(2+\rho)(1+n)(1+g)}\right] (f(k_t) - f'(k_t)k_t). \tag{4.8$'$}$$

With the Cobb-Douglas technology where $f(k_t) = k_t^{\alpha}$, we can re-write (4.8$'$) as

$$k_{t+1} = \left[\frac{F(1-\theta)(1-q-\beta(1-qF))}{(F-1)(2+\rho)(1+n)(1+g)}\right] (1-\alpha)k_t^{\alpha}. \tag{4.9}$$

Condition (4.9) defines explicitly the capital per effective worker in period $t+1$ as a function of the previous period's capital per effective worker. Before we proceed we will state here a proposition.

[23] Recall that elderly receive the marginal product of the capital as the return factor on their savings. This can be even more easily seen by recognizing that the interest rate is $f'(k) - \varkappa = f'(k) - 1$. That is, $1 + \textit{interest rate}$ is just $f'(k)$, as was previously shown. Seniors at time t receive an income from their savings of the previous period and consume it all out. Certainly, it would make more intuitive sense if under the model parameterization $f'(k)$ ends up exceeding 1.

Proposition 2 *When $\beta < \frac{1-q}{1-qF}$, there exists a globally stable steady state on $\mathbb{R}_{++}$, that is unique and positive.*

Proof. From (4.9) we see that the right-hand side is indeed positive when $\beta < \frac{1-q}{1-qF}$. By noting that at the equilibrium, $k_{t+1} = k_t = k^*$, we can solve (4.9) for the positive steady state

$$k^* = \left[\left[\frac{F(1-\theta)(1-q-\beta(1-qF))}{(F-1)(2+\rho)(1+n)(1+g)}\right](1-\alpha)\right]^{\frac{1}{1-\alpha}}, \tag{4.10}$$

which is unique for exogenously given model parameters.[24] ■

In the absence of any exogenous shock, the economy will smoothly converge to a long-run equilibrium level of capital per effective worker, although the convergence process might be quite slow.

The Speed of Convergence.

Suppose that the economy starts with a level of capital per effective worker which is different from the equilibrium level. How fast will the economy converge to the equilibrium?

Note that $k_{t+1} - k^* \simeq \frac{dk_{t+1}}{dk_t}|_{k_t=k^*}(k_t - k^*)$.[25] If , say, $\frac{dk_{t+1}}{dk_t}|_{k_t=k^*} = 1/4$, then the distance between k_{t+1} and k^* will be roughly equal to the quarter of the original distance from k_t and k^*. Hence, each *period* k moves three-fourth of the way toward k^*. So, we can find the value for $\frac{dk_{t+1}}{dk_t}|_{k_t=k^*}$.

From (4.9)

$$\frac{dk_{t+1}}{dk_t}|_{k_t=k^*} = (1-\alpha)\alpha k^{*\alpha-1}\left[\frac{F(1-\theta)(1-q-\beta(1-qF))}{(F-1)(2+\rho)(1+n)(1+g)}\right] \tag{4.11}$$

[24] Observe that $\frac{\partial k^*}{\partial \theta} < 0$. However, $\frac{\partial k^*}{\partial F}$ and $\frac{\partial k^*}{\partial q}$ are both ambiguous in sign. On the one hand, higher enforcement parameters mean lower evasion (and thus, lesser capital stock), but also lower penalty payments when caught (and thus, higher than before capital stock).

[25] This is the standard result in the literature, stemming from a first-order approximation of (4.9) around k^* (see, e.g., Romer 2001, pp. 81-82). Here k_0 is any level of capital, different from the equilibrium one.

By substituting (4.10) into (4.11) and simplifying, we obtain $\frac{dk_{t+1}}{dk_t}\mid_{k_t=k^*}=\alpha$. Thus, the share of total income that goes to capital, α, determines the convergence to the balanced growth path.

Not surprisingly, this finding is similar to the one in the basic OLG models. Assume α is roughly 0.36 (see, e.g., Galasso 1999, p. 715). Then, the speed of convergence is $1-0.36^{\frac{1}{30}}=3.3\%$ per year.[26] This result is standard for the benchmark OLG models, but that speed might be quite fast (De La Croix and Michel 2002, p. 339).

Implications for Dynamic Inefficiency.

It is possible for the decentralized economy to over-accumulate capital up to the point that is not consistent with the Pareto-efficient level, and become *dynamically inefficient.*[27] Thus, k^* may be greater than the golden-rule level of the capital stock, k^*_{gr}. If the economy indeed over-accumulates capital, then $f'(k^*)<n+g+\varkappa(=1)$. Certainly, at k^*_{gr} the following condition is met: $f'(k^*_{gr})=n+g+1$. We will summarize the implications that follow in the proposition below.

Proposition 3 *If the capital share in total value added is sufficiently small, the economy will be dynamically inefficient.*

Proof. Clearly, $f'(k^*)\equiv\alpha k^{*\alpha-1}$. From (4.10), $f'(k^*)=\frac{(F-1)(2+\rho)(1+n)(1+g)}{F(1-\theta)(1-q-\beta(1-qF))}\frac{\alpha}{1-\alpha}$. In the limit, when α approaches zero, $f'(k^*)$ is zero. This completes the proof. ■

Interestingly, the dependence of the balanced-growth path level of capital on all

[26] We will try to visualize it in a heuristic way. Assume that the distance, k^*-k_0, is normalized to 1. Then, as before, $\frac{k_t-k_0}{k^*-k_0}=1-\alpha$, which is the per-period convergence speed around the steady state. Certainly, $\frac{k^*-k_t}{k^*-k_0}=\alpha$. Now, assume that the annual convergence speed is x, i.e., we "start" xth unit to the right of k_0. Since one period lasts for 30 years, the distance $1-x$ will keep shrinking, and will become $(1-x)^{30}$ eventually. But that distance is also equal to α. Thus, $x=1-\alpha^{\frac{1}{30}}$.

[27] A simple way to understand the possibility of oversaving, consider a case when $g=0$. Note, then, for every old person there are $1+n$ young persons in the economy. Hence, one way to transfer the resources between the generations is to persistently force the young to save a dollar, so that the old always receive $1+n$ dollars. An omnipotent social planner can force such a saving behavior. This, clearly, implies that the rate of return on saving is n. Now, consider the case when the market rate of return is below n. The young have no choice but to save more at such a *lower* rate, which is less efficient than the government-induced mandatory saving.

model variables does *not* produce additional sufficient conditions for over-accumulation of capital.

Comparative Dynamics.

Equipped with (4.4), (4.5) and (4.10), and by invoking that at the steady-state $w_{1t}A_t = A_t(1-\alpha)k^{*\alpha}$, we can discuss the impact of the change in the model's key parameters and variables, namely ρ, r_{t+1}, q, F, $w_{1t}A_t$ and θ, on the optimal level of the declared income and consumption. Again, a change in one of the model parameters will change $w_{1t}A_t$ by changing the steady state level of capital. But the most important question here to ask, is: "*when,* as a result of the change in one of the model parameters, the state of the economy will change?". And here we have to face the most unfortunate feature of a two-period OLG model. Such a change, of course, will change $w_{1t}A_t$ *after* the initial period is over. So, any additional effect, partially caused by $\frac{\partial w_{1t}A_t}{\partial k_t}$, will not be felt by the working members of the present generation because the mere structure of the two-period OLG model shuts off such a possibility! Sure, a change in any model parameter may have a strong behavioral effect, but at the time of the shock the economy's capital stock is still fixed by the previous level's capital stock, leaving no room for an extra effect. Thus, for a moment, we consider a behavioral impact only.

$$\frac{\partial X_{1t}^*}{\partial \rho} = \frac{w_{1t}A_t(qF-1)(\theta-1)}{\theta(F-1)(2+\rho)^2} \tag{4.12}$$

$$\frac{\partial X_{1t}^*}{\partial r_{t+1}} = 0 \tag{4.13}$$

$$\frac{\partial X_{1t}^*}{\partial q} = \frac{w_{1t}A_tF(1-\theta)}{\theta(F-1)(2+\rho)} \tag{4.14}$$

$$\frac{\partial X_{1t}^*}{\partial w_{1t}A_t} = \frac{F\theta(2-q)+\theta\rho(F-1)+qF-\theta-1}{\theta(F-1)(2+\rho)} \tag{4.15}$$

$$\frac{\partial X_{1t}^*}{\partial \theta} = \frac{w_{1t}A_t(1-qF)}{\theta^2(F-1)(2+\rho)} \tag{4.16}$$

$$\frac{\partial X_{1t}^*}{\partial F} = \frac{w_{1t}A_t(1-\theta)(1-q)}{\theta(2+\rho)(F-1)^2} \tag{4.17}$$

As expected, the amount of declared income is completely insensitive to the change in the interest rate. This is clear from (4.13). Intuitively, on the one hand, an increase in the interest rate increases the opportunity cost of being honest: the more you evade, the more you may have in your saving account, earning higher return than before (i.e., X_{1t} decreases). On the other hand, an increase in the interest rate tends to increase X_{1t} since lower evasion can achieve a given wealth accumulation. Apparently, these two effects exactly counteract each other.[28]

It is also clear that (4.14) and (4.17) are both positive, meaning that higher penalty rate or probability of detection lowers incentives to cheat. These results are also expected intuitively. An interesting result is presented by (4.16). As we have already required before, $1-qF$ has to be positive for the existence of an interior solution. But if so, (4.16) unequivocally shows that an increase in the tax rate increases the amount of declared income. So, we obtain the same results as Yitzhaki did.

Since we want $qF < 1$, from (4.12) we conclude that an increase in the rate of time-preference encourages cheating incentives. Intuitively, greater ρ means greater weight placed to the utility (from consumption) in the first period. Therefore, evasion should increase in order to have higher than before after-tax funds to finance purchases of goods and services in the first period.[29] Further, note that (4.15) is

[28] It is tempting to think that in general, the rise in the market rate of return has a positive wealth effect: as the interest rate increases, one (who has decreasing absolute risk aversion) tends to demand more risky assets (i.e., evasion increases). However, the wealth in our model is not just the immediate earning, but the overall annuity value of lifetime resources. If so, higher interest rate may actually imply a negative wealth effect, as it is associated with lower capital stock, and thus, lower earnings when young.

[29] One might wonder that since the penalty payment is paid in the first period, could not we alternatively say that higher ρ means more incentive to be honest in order to forgo the penalty payment? Indeed, in a two-period model, the individual can force her future self to bear all the risk (i.e., she can consume as much as she wants in period one and then let the penalty payment come from funds that otherwise were targeted for saving and hence, for period two consumption). In other words, even though the penalty payment is technically paid in period one, the taxpayer can force her period two self to shoulder the consequences, and this becomes increasingly the case as ρ increases.

to be positive in sign (since a positive fraction of income is to be declared, i.e., $0 < \frac{F\theta(2-q)+\theta\rho(F-1)+qF-\theta-1}{\theta(F-1)(2+\rho)} < 1$).

As for the other choice variable, we get

$$\frac{\partial C^*_{1t}}{\partial \rho} = \frac{(1-\theta)}{(2+\rho)^2} w_{1t} A_t \tag{4.18}$$

$$\frac{\partial C^*_{1t}}{\partial r_{t+1}} = 0 \tag{4.19}$$

$$\frac{\partial C^*_{1t}}{\partial q} = 0 \tag{4.20}$$

$$\frac{\partial C^*_{1t}}{\partial w_{1t} A_t} = \frac{(1+\rho)(1-\theta)}{(2+\rho)} \tag{4.21}$$

$$\frac{\partial C^*_{1t}}{\partial \theta} = -\frac{1+\rho}{2+\rho} w_{1t} A_t \tag{4.22}$$

$$\frac{\partial C^*_{1t}}{\partial F} = 0 \tag{4.23}$$

From (4.18) it is evident that an increase in the rate of time preference increases the first-period consumption. Again, it is logically obvious since higher ρ means higher relative weight assigned to the utility derived from the first-period. From (4.21) it is clear that higher wage income tends to encourage more consumption in the first period. However, as (4.22) shows, the higher the tax rate, the lower the consumption in the first-period. This is not contrary to macroeconomic intuition.[30] In addition, we note that the first-period consumption is completely insensitive to the key parameters of the model, namely, q and F (refer to (4.20) and (4.23)). The taxpayer adjusts solely her other choice variable, the amount of declared income, to the change in the probability of detection and the penalty rate. Her consumption decision remains unchanged. As was explained earlier, the latter results from the way we have stated problem (4.1). Finally, as expected, (4.19) is zero.

[30] Note that everywhere we assume that a statutory change in income tax rate does not affect the real wage, $w_{1t}A_t$.

But one should be careful here. True, when the shock (say, higher tax rate) hits the economy at the very start of a period, the impact on the steady state will be felt after that period when a young tax cheater becomes a law-obeying individual by assumption. But again, this is because of the two-period structure of the model. *Assume* that instantly after the shock, the economy jumps to the new steady-state. Therefore, we cannot ignore the term $\frac{\partial w_{1t}A_t}{\partial k_t}\frac{\partial k_t}{\partial \theta}$ anymore. However, some of our comparative dynamics results will take different forms. For example, we saw previously that an increase in ρ raises both consumption (decreasing saving) and evasion (increasing saving). Eventually, the capital stock will fall (see 4.10). Lower capital stock means lower income in the working period, depressing consumption. Therefore, in the very long-run, the impact of the change in the model parameters on evasion and/or consumption is less clear-cut. To illustrate our point more formally, consider the long-run impact of higher tax rate on income declaration incentives. In fact, it can be easily shown that in the very long-run, when an increase in the tax rate changes the steady-state level of the economy, the amount of declared income can move either way. To see this, just partially differentiate (4.4) with respect to θ, but this time note that in the long-run $\frac{\partial w_{1t}A_t}{\partial k^*}\frac{\partial k^*}{\partial \theta}$ is negative, adding ambiguity, which is absent from traditional static models.

$$\underbrace{\frac{\partial X_{1t}^*}{\partial \theta}}_{\text{total effect (?)}} = \underbrace{\frac{1-qF}{\theta^2(F-1)(2+\rho)}(1-\alpha)A_t{k^*}^{\alpha}}_{\text{immediate effect (+)}} + \underbrace{\frac{F\theta(2-q)+\theta\rho(F-1)+qF-\theta-1}{\theta(F-1)(2+\rho)(\theta-1)}\alpha A_t k_{new}^{*\alpha}}_{\text{long-run effect (-)}} \qquad (4.16')$$

Here k_{new}^* is the new steady state, due to the different tax rate. There is no change in the steady-state level of the economy in a static environment, i.e., we have only a short-run effect, encouraging compliance when the tax rate increases. In contrast, our model is capable of generating two distinct effects: a *behavioral* (the short-run, immediate) *effect* and the consequent *productivity* (the long-run) *effect*. The first effect reflects the individual's immediate alteration of her choice variables, and the resulting last effect reflects the fall in the capital stock, driving real wages down. A taxpayer (a representative of a new generation) hence, has some inclination to be *less* honest relative to her predecessors, in order to offset the fall in real income.[31] To sum up, *instantly*, a rise in the tax rates will encourage honesty. But since the capital stock will keep falling until it reaches the new steady-state level, what will happen to the magnitude of declared income in the economy over-time remains unknown. However, we will not feel the long-run effect any time before the end of the first period. But this is just a well-known dismal nature of a two-period OLG model. The message, nevertheless, should be clear: an impact of higher tax rate on compliance is less clear-cut as we move from a static, to a dynamic framework.

The bottom line behind the above analysis, is that any static model ignores the additional productivity effect by definition. Of course, we assumed an immediate convergence to the new steady-state to prove our point, but we did not have to make such an assumption, should the model be, say, in 75 periods to start with! If so, the convergence to the new steady-state will start much "faster", since every period lasts for a year, not for a half a lifetime.

A Remark on the Revenue-Maximizing Tax Rate.

The OLG model we have developed in this chapter has a distinctive characteristic that is absent from traditional intertemporal maximization problems. In our formulation, not only does the tax affect the taxpayer's disposable income, and thus saving decision, but also her incentives to cheat. In that regard it would be interesting to see how high the government should set the tax rate in order to maximize the revenue per worker in the long-run, when all the dynamic adjustments in the economy

[31] Recall partial derivative of X_{1t}^* with respect to $w_{1t}A_t$ (which, by no means is an entire, *lifetime* wealth) is positve.

are completed. We do not claim here to find the "optimal" tax rate since the latter is traditionally meant to minimize welfare losses, while still meeting certain revenue requirements and a desired income distribution. Further, individual utilities do not necessarily depend on consumption only, but may also be affected by the amount of leisure consumed, which, in turn, may have significant implications on the choice of the appropriate tax policy (Stern 1987). Hence, we are not pursuing here a goal of finding the best tax schedule, maximizing a social welfare function subject to fixed governmental revenue requirements. Rather, we focus on the revenue-maximizing government. After all, in the realm of large public indebtedness, a revenue-maximizing government, rather than a social-welfare maximizing planner, might be more tempting to model.

One should *not* expect the revenue-maximizing tax rate to grow without a bound since our model, implicitly assuming that the government marginal propensity to consume is unitary, predicts a negative relationship between the tax rate and the aggregate output through the adverse impact of taxation on private capital accumulation. On the other hand, we may not be able to come-up with a sufficiently low tax rate, predicted by some recent studies, which recognize the excess burden of taxation when the supply of labor is elastic, the phenomenon, which we ignore in our model. However, this is not a significant drawback of our model, since there are evidences that labor supply is inelastic for most individuals. Though, the important message here is that while traditional static models of the perfectly-competitive economic environment may theoretically predict that the revenue-maximizing tax rate can be as high as 100% (especially when higher taxes encourage compliance!), it will not necessarily be so in our dynamic framework. Note that this reasoning holds regardless of whether the tax rate encourages, or discourages compliance. Indeed, when higher taxes improve compliance, they also adversely affect the state of the economy, and thus, the government's ability to collect more revenue.[32] On the contrary, if higher tax rates were to *decrease* compliance and increase saving (an engine of the growth), they would, obviously, imply *lower* taxes collected to start with. An importance of such a trade-off becomes clear now, i.e., the trade-off between higher (lower) immediate revenue and low (high) long-run income level. That kind of trade-off cannot be captured in a static environment.

Although we will avoid making policy prescriptions based on the results of this section, we, nevertheless, will lay a foundation for the future works, which perhaps would enrich our framework by recognizing labor-leisure distortions, introducing the effect of the publicly funded goods on people's utilities, and so on.

With that said, assume now that the government's sole goal is to maximize the revenue per worker, by choosing the appropriate tax rate, and given the fact that the taxpayer optimizes her compliance and consumption decisions. Clearly, the government's problem is to

$$\underset{\{\theta\}}{Max} R \equiv (1-\beta)\theta X_{1t}^{*} + \beta[\theta X_{1t}^{*} + F\theta(w_{1t}A_t - X_{1t}^{*})].^{33} \tag{4.24}$$

where R is the tax revenue per worker.[34] After implementing the first-order necessary condition, utilizing expressions (4.4) and (4.10), and undertaking a series of lengthy algebraic manipulations, we finally find that the revenue-maximizing tax rate, θ^*, is as follows:[35]

$$\theta^* = 1 - \alpha + \frac{\alpha(1-\beta F)(1-qF)}{F(2-\beta-q+\beta qF+\rho)-\rho-1}. \tag{4.25}$$

[32] Remember, as in many studies we assume the unproductive government sector.

[33] Clearly, here we have an unconstrained maximization problem.

[34] Note that the total revenue collected, TR, can be expressed as $L_t(1-\beta)\theta X_{1t}^{*} + L_t\beta[\theta X_{1t}^{*} + F\theta(w_{1t}A_t - X_{1t}^{*})]$. Divide both sides by L_t to obtain $\frac{TR}{L_t}(\equiv R)$.

[35] It can be shown that the sufficient condition, $\frac{d^2R}{d\theta^2} < 0$, collapses to $(1-F)(2+\rho) < 0$, which is always true.

Now we are ready to establish the following proposition.

Proposition 4 *A revenue-maximizing government cannot set the tax rate equal to 100%.*

Proof. *It is straightforward to show that* $\theta^* = 1$ *if, and only if* $(1 - F)(2 + \rho) = 0$. *Neither* F *is unitary, nor the rate of time-preference is negative in our model.* ■

It is tempting to conclude here that when the penalty rate is lower and exactly equal to 1, then the government can "compensate", and set the tax rate equal to 100% in order to maximize its revenue. Not in our model. Recall that when $F = 1$, the consumption of those who are caught when old becomes $C^c_{2t+1} \equiv R_{t+1}((1-\theta)w_{1t}A_t - C_{1t})$, which implies that the taxpayer did not cheat to start with. Obviously, this is contradictory.

Before ending this chapter, we would like to adopt a slight, but important modification to our OLG model, which ignored the cost of compliance faced by the taxpayer. We render that to the next section, though.

4.2. Tax Evasion and Tax Compliance Costs: a Simple Scenario.

It has long been proposed that an efficient tax system should be simple and least costly in terms of compliance. Fundamental tax reforms in this direction have long been proposed in many developed countries. For instance, in Australia it has been suggested that complexity and compliance costs are one of the major disadvantages of the country's taxation system (Australian Chamber of Commerce and Industry 2004). In general, tax compliance costs put a considerable burden on the government and individuals.

Rarely one can avoid some expenditures in preparing the individual tax return. In recent years, the literature on costs of tax administration considerably broadened, encompassing not only enforcement costs, but also "...the costs incurred by taxpayers in meeting the requirements laid on them by the tax law and revenue authorities. They are the costs over and above the actual payment of tax and over and above any distortion costs inherent in the nature of the tax; costs which would disappear if the tax was abolished. Thus, for individuals, they include the costs of acquiring sufficient knowledge to meet their legal obligations; the time taken in completing their personal tax returns, and obtaining, filing and storing the data to enable them to complete their returns; the fees paid to any advisers or tax agents; incidental expenses, such as travel costs to visit a tax advisor or the revenue authorities; and, more difficult to measure, psychic or psychological costs—the stress and anxiety experienced by some in seeking to deal with their tax affairs" (Sandford 1995 p. 1).

In this section, hence, we will try to capture those costs and see how they affect the behavior of taxpayers. Realizing the trade-off between the realism and tractability, we will introduce a simple cost structure, given exogenously. One final argument is worth remembering before we proceed. A part of the individual's compliance cost (such as a payment for using Intuit Inc.'s Turbo Tax® tax preparation software) becomes someone else's income in the economy (Intuit Inc.'s, in our example). So, even though the payment drains the savings account of the taxpayer, it flows into the recipient's account. Therefore, the overall societal impact maybe a "zero-sum game". The implication is that in interpreting the impact of tax compliance costs on the overall saving and capital accumulation in the economy, we should ideally include only those costs, which affect the economy as a whole, for instance, deadweight losses, psychic costs, time lost and so on.

However, there is yet another important reason why we are trying to introduce the costs of compliance to this study. Introduction of such a cost structure can help to shed some light on *cross-country* comparisons, when different economies face different tax compliance costs due to their institutional structure. When, for example, in one country the tax compliance is relatively less costly, while in the other one it is substantial, one may expect different compliance level in those countries. Hence, studying tax compliance costs can serve as a convenient and straightforward tool in analyzing the compliance decision of various economies with various earnings, holding all other things the same. This is important both at the country, and individual level.[36]

The Model Specification.

Assume that we have to study a different economy, where a representative taxpayer faces the exogenous compliance cost, $\gamma \equiv \xi \overline{w_{1t}A_t}$, where $\xi \in (0,1]$ and the "bar" is used to distinguish the country with compliance costs from the one without. Note first that the statement that the tax compliance costs are proportional to income may not picture the full story since some people, in avoiding the costs of compliance, may want to reduce their income declaration, say, by filing a simpler tax return. But if so, the cost of compliance is no longer a constant fraction of the wage. On the other hand, richer people may face higher costs of compliance because of the larger opportunity costs, but whether these are strictly proportional to one's income is an open question, which we try to avoid in this study.

With these specification, the problem is to

$$\underset{\{C_{1t},X_{1t}\}}{Max}\; E[U] \equiv U(\overline{C}_{1t}) + \frac{1}{1+\rho}\left[(1-q)U(\overline{C}^{nc}_{2t+1}) + qU(\overline{C}^{c}_{2t+1})\right], \tag{4.26}$$

where $\overline{C}^{nc}_{2t+1} \equiv R_{t+1}(\overline{w_{1t}A_t} - \theta\overline{X}_{1t} - \overline{C}_{1t} - \gamma)$ and $\overline{C}^{c}_{2t+1} \equiv R_{t+1}(\overline{w_{1t}A_t} - \theta\overline{X}_{1t} - F\theta(\overline{w_{1t}A_t} - \overline{X}_{1t}) - \overline{C}_{1t} - \gamma)$.

[36] At this point we are not aware of any empirical work, studying tax evasion in different countries with different compliance costs. However, that might be an extremely interesting study to undertake. After all, we often hear that complicated taxation rules deter honesty within an economy. But the similar argument also holds for across countries, facing different institutional structures, and thus, various complications stemming from endless tax reforms. Sandford (1995 p. 6) reports that an increasing complication of the tax code is especially relevant for Australia, New Zealand, Canada, and to a lesser extent, the United States. Even at individual level the problem should not be ignored. For instance, Alm et al. (1993) find that larger income is associated with higher probability of underreporting and the level of underreporting in Jamaican economy. Their results accord with those of Clotfelter (1983). An income, of course, can be largely affected by compliance costs.

After a series of tedious, but straightforward manipulations we obtain the new income declaration and consumption profiles:

$$\overline{X}_{1t}^{*} = \frac{F\theta(2-q)+\theta\rho(F-1)+qF-\theta-1}{\theta(F-1)(2+\rho)}\overline{w_{1t}A_t} - \frac{qF-1}{\theta(F-1)(2+\rho)}\xi\overline{w_{1t}A_t}, \quad (4.27)$$

and

$$\overline{C}_{1t}^{*} = \frac{(1+\rho)(1-\theta)}{2+\rho}\overline{w_{1t}A_t} - \frac{1+\rho}{qF+1+\rho}\xi\overline{w_{1t}A_t}. \quad (4.28)$$

Since we are concerned with the interior solution, we will establish here the following proposition.

Proposition 5 *To ensure absolute compliance, government should set $qF = 1$.*

Proof. Clearly, for an interior solution we have to require $\frac{F\theta(2-q)+\theta\rho(F-1)+qF-\theta-1-\xi(qF-1)}{\theta(F-1)(2+\rho)}$ to be a positive fraction (refer to (4.27)). As for the fraction, the latter collapses to $(1-\theta-\xi)(qF-1) < 0$. This implies that either qF is less than 1, or $1-\theta$ is less than ξ. If we take $qF = 1$ and substitute into the previous ratio, we will have it collapse to 1 after some algebraic rearrangements, implying $X_{1t} = w_{1t}A_t$. ■

Further, to ensure that X_{1t} is positive, we state the following assumption.

Assumption 2. $F\theta(2-q)+\theta\rho(F-1) > \theta + (1-\xi)(1-qF)$.

From (4.27) and (4.28) we do not know how $\overline{X}_{1t}^{*}$ and $\overline{C}_{1t}^{*}$ are related to X_{1t}^{*} and C_{1t}^{*}, since we do not know what happens to the new steady-state level of capital (call it $\overline{k}^{*}$), and therefore, to the taxpayer's true income. Certainly, if the steady-state of the economy in two different countries were identical, we would have concluded that $\overline{X}_{1t}^{*} > X_{1t}^{*}$ (since $qF < 1$), and $\overline{C}_{1t}^{*} < C_{1t}^{*}$. Nevertheless, we are ready to establish here the following proposition.

Proposition 6 *In an economy with exogenous compliance costs, the amount of declared income as a fraction of the true income increases, while consumption as a fraction of the true income decreases.*

Proof. By closely examining (4.4), (4.5), (4.27) and (4.28), we see that $\frac{X^*_{1t}}{w_{1t}A_t} < \frac{\overline{X}^*_{1t}}{\overline{w_{1t}A_t}}$, while $\frac{C^*_{1t}}{w_{1t}A_t} > \frac{\overline{C}^*_{1t}}{\overline{w_{1t}A_t}}$. ■

Hence, *relative* to their true incomes, people of the country, finding compliance more costly, become more honest in declaration of income, as opposed to the residents of another country, who do not face as much compliance costs.[37] But note that we cannot say for sure whether one country has larger *absolute* amount of declared income. We will elaborate more on that below, after determining the steady-state level of the economy.

In analogy to our previous discussions, we can see that the total saving for the group of people who are not caught cheating is

$$\begin{aligned} \overline{S}_{nc} &= \frac{L_t F(1-\theta)(1-q)(1-\beta)}{(F-1)(2+\rho)}\overline{w_{1t}A_t} + \frac{L_t(qF-1)(1-\beta)}{(F-1)(2+\rho)}\xi\overline{w_{1t}A_t} \\ &+ \frac{L_t(1+\rho)(1-\beta)}{qF+1+\rho}\xi\overline{w_{1t}A_t}. \end{aligned} \tag{4.29}$$

Similarly, the total saving for the group of people who are caught cheating is

$$\begin{aligned} \overline{S}_{c} &= \frac{L_t F(1-\theta)q\beta}{2+\rho}\overline{w_{1t}A_t} + \frac{L_t(qF-1)\beta}{(F-1)(2+\rho)}\xi\overline{w_{1t}A_t} \\ &+ \frac{L_t(1+\rho)\beta}{qF+1+\rho}\xi\overline{w_{1t}A_t} - \frac{L_t F(qF-1)\beta}{(F-1)(2+\rho)}\xi\overline{w_{1t}A_t}. \end{aligned} \tag{4.30}$$

Consequently,

$$\begin{aligned} \overline{K}_{t+1} &\equiv \overline{S}_{nc} + \overline{S}_c = \frac{F(1-\theta)(1-q-\beta(1-qF))}{(F-1)(2+\rho)}L_t\overline{w_{1t}A_t} \\ &+ \frac{(1-F\beta)(qF-1)}{(F-1)(2+\rho)}L_t\xi\overline{w_{1t}A_t} + \frac{(1+\rho)}{qF+1+\rho}L_t\xi\overline{w_{1t}A_t} \\ &= \left[\begin{array}{c} \frac{F(1-\theta)(1-q-\beta(1-qF))+\xi(1-F\beta)(qF-1)}{(F-1)(2+\rho)} \\ +\frac{\xi(1+\rho)}{qF+1+\rho} \end{array} \right] L_t\overline{w_{1t}A_t}. \end{aligned} \tag{4.31}$$

[37]This result has to be interpreted with some caution. The proponents of a less complex and costly tax system stress that making the tax code simpler might actually help to free up government recourses to more efficiently combat tax evasion.

Comparing expression (4.31) and (4.8), we cannot say whether the total capital stock increases, or decreases. Once again, with the Cobb-Douglas technology and the equilibrium condition, $\overline{k}_t = \overline{k}_{t+1} = \overline{k}^*$, we obtain

$$\overline{k}^* = \left[\left[\begin{array}{c} \frac{F(1-\theta)(1-q-\beta(1-qF))+\xi(1-F\beta)(qF-1)}{(F-1)(2+\rho)(1+n)(1+g)} \\ +\frac{\xi(1+\rho)}{(qF+1+\rho)(1+n)(1+g)} \end{array} \right] (1-\alpha) \right]^{\frac{1}{1-\alpha}} . \tag{4.32}$$

Since *a priori* we do not know how $\overline{k}^*$ is related to k^*, we cannot say whether $\overline{X}^*_{1t} \gtrless (=) X^*_{1t}$, or $\overline{C}^*_{1t} \lessgtr (=) C^*_{1t}$. This is an interesting finding since one might think that the economy, where everyone loses "few dollars" because of the complex taxation rules, is poorer than the economy, facing no such losses. That is not necessarily the case. Indeed, if "poorer" country saves more, it might actually end-up generating larger output per person in the long-run.

Conclusions.

In this chapter we developed a two-period model of income tax evasion and solved it analytically. We discussed the dynamics of the economy and obtained the closed-form solution for the steady-state value of capital per effective worker. Introducing tax cheaters did not produce any additional sufficient conditions for oversaving, and mostly the results associated with dynamic efficiency are consistent with the literature. Further, we were able to distinguish between the behavioral effects caused by changing policy variables from the long-run productivity effects, which is what makes our study essentially different from static models. The case of revenue-maximizing government was also discussed and the main finding was that the revenue-maximizing tax rate set by the government has to be strictly less than 100%. Finally, we were able to introduce simple compliance cost structure and found that countries with higher costs of tax compliance will tend to evade smaller portion of their domestic income.

Chapter Five: Review of Dynamic Programming, Stochastic Control and Application to Tax Evasion.

5.1. Basic Dynamic Programming.

We start this chapter by reviewing the basic principles of dynamic programming. Predominantly throughout this chapter we will rely on the presentation by Dixit and Pindyck (1994, pp. 59-120), Dixit (1993, pp. 1-8), and Kamien and Schwartz (1981, pp. 238-243).

Let us suppose, for clarity, that a decision-maker at time t is trying to maximize the net present discounted value of all the payoffs by choosing the optimal time pace of the variable she controls, a *control variable*, y_t. In addition, there is a *state variable*, x_t, reflecting the information about the current state of the decision-maker. At time t the state variable is known and in general, the evolution of the state variable is a random *Markov process* by assumption.[38] Both the state and control variables at time t determine the flow of immediate payoffs, Π_t, which is a function of x_t and y_t. The agent has to choose y_t optimally by utilizing the known information about x_t.

The decision-making process is as follows: to maximize the sum of the value of the *immediate payoff* at time t, and the *continuation value* thereafter. In essence, the optimization process is split into two parts. To better understand the optimization process, suppose the agent makes optimal decisions from now onwards. Given these *optimal* decisions, the resulting present value of *all* the payoffs at time t will depend on the value of the state variable at time t. Denote this present value as $J_t(x_t)$ (also called *value function*). But since x_t is a random process, the agent does not know x_{t+1} and so on. From time $t+1$ perspective, given optimal decisions then and

[38] A Markov process x_t possesses the following property: the probability distribution of x_t depends only on that of x_{t-1}. That is, two periods or earlier the influence of the state variable on its standing at t vanishes, meaning that only present information is relevant to predict the immediate future state of the process.

onwards, the present value of all the payoffs can be denoted as $J_{t+1}(x_{t+1})$. From time t perspective, again, x_{t+1} is not known and thus, the agent takes the expected value of it, $E[J_{t+1}(x_{t+1})]$, where $E[\bullet]$ is the expectation operator. $E[J_{t+1}(x_{t+1})]$ is what the continuation value is. At time t the agent has to discount it back to the present, which leads to $\frac{1}{1+\rho}E[J_{t+1}(x_{t+1})]$. Therefore, the problem can be stated as

$$J_t(x_t) = \underset{\{y_t\}}{Max}\left\{\Pi_t(x_t, y_t) + \frac{1}{1+\rho}E[J_{t+1}(x_{t+1})]\right\}. \tag{5.1}$$

The first term in the brackets, $\Pi_t(x_t, y_t)$ is the immediate payoff. Hence, we come to the basic principle of dynamic programming, or (Richard) *Bellman's Principle of Optimality*, stating that "[a]n optimal path has the property that whatever the initial conditions and control values over some initial period, the control (or decision variables) over the remaining period must be optimal for the remaining problem, with the state resulting from the early decisions considered as the initial condition" (Kamien and Schwartz 1981, p. 238). The continuation value subsumes that y_{t+1} and so on are chosen optimally, i.e., at time t the only problem is to choose y_t optimally. Equation (5.1) is called the Bellman equation or the *fundamental equation of optimality.*[39]

One way to solve the dynamic programming problem is to work backward. If there is a fixed time horizon, T, there is a termination payoff at that time. So, one can try to choose y_{T-1}, aiming at time $T-1$ to maximize the sum of the immediate payoff and the discounted continuation value (the latter being as the *expectation* of the termination payoff). But once y_{T-1} is optimally chosen that way, we will know

[39]Interestingly, L. S. Pontryagin, one of the most brilliant Russian mathematicians of the 20th century and the foremost contributor to the theory of optimal control, stated in his biographical notes that what we call the Bellman equation nowadays was in fact obtained first by another American mathematician, Rufus Isaacs (1914-1981). According to Pontryagin, Isaacs was working on the secret military projects for a while, and thus, his findings were not widely publicized. Whether Pontryagin was jealous to his colleague Bellman and was trying to diminish his findings that way, or was just trying to be fair, is a mystery to me. An interested reader may refer to the Pontryagin's *Biography* (*Zhizneopisanie* 1998, pp. 138-232), published in Moscow after his death.

$J_{T-1}(x_{T-1})$. Similarly, one can proceed by finding y_{T-2} and so on.

However, the analysis can be greatly simplified if one assumes an infinite planning horizon, so that the payoff flow, the discount rate and the conditional (upon the current information) cumulative probability distribution of the state variable are independent of the actual label of the date. In an infinite setup, there is no known value function from which one can work backward. Therefore, the backward solution is not feasible. But that implies that the Bellman equation is exactly the same no matter what time we are at. Loosely speaking, if $T = 10$ and the agent is at time $t = 2$, the value function at that time is not the same as the one at, say, $t = 3$ (because the continuation value, depending on x and y at time 4 onwards, would be different). But when time is infinite, that difference will be irrelevant. The value function will still depend upon what value of the state, x_t, we have. Thus, we can rewrite the Bellman equation for any t as

$$J(x_t) = \underset{\{y_t\}}{Max} \left\{ \Pi(x_t, y_t) + \frac{1}{1+\rho} E[J(x_{t+1})] \right\}.$$

In more general notations, the value of the state in one period, call it $x^{'}$, is conditioned on its value in the previous period, call it x, and also the control in that period. Consequently, for all possible values of x (not just for t and $t+1$), we obtain

$$J(x) = \underset{\{y\}}{Max} \left\{ \Pi(x, y) + \frac{1}{1+\rho} E[J(x^{'} \mid x, y)] \right\}. \tag{5.2}$$

Equation (5.2) is the version of the Bellman equation for infinitely repeating (recursive) problem of dynamic programming (Dixit and Pindyck 1994, p. 102).

Dynamic Programming: Continuous Time.

We now consider the case when time is continuous. Further, we will complicate the notation a little bit by allowing the payoff flow or terminal payoff to depend on

t as well. If $\Pi(x,y,t)$ and ρ are the rates of the profit flow and the discount rate, respectively, over the time period of length Δt they will become $\Pi(x,y,t)\Delta t$ and $\rho\Delta t$, correspondingly. Thus, equation (5.2) can be re-written as

$$J(x,t) = \underset{\{y\}}{Max}\left\{\Pi(x,y,t)\Delta t + \frac{1}{1+\rho\Delta t}E[J(x',t+\Delta t \mid x,y)]\right\}.$$

After multiplying both sides by $(1+\rho\Delta t)$ and bringing one of the resulting terms, $J(x,t)$, on the left-hand side to the right-hand side, we obtain

$$\begin{aligned} J(x,t)\rho\Delta t &= \underset{\{y\}}{Max}\left\{\Pi(x,y,t)\Delta t(1+\rho\Delta t) + E[J(x',t+\Delta t) - J(x,t)]\right\} \\ &= \underset{\{y\}}{Max}\left\{\Pi(x,y,t)\Delta t(1+\rho\Delta t) + E[\Delta J]\right\}. \end{aligned}$$

The final step is to divide both sides by Δt and take the limit of the expression when Δt is approaching to zero. This will result in

$$\rho J(x,t) = \underset{\{y\}}{Max}\left\{\Pi(x,y,t) + \frac{1}{dt}E[dJ]\right\}, \tag{5.3}$$

which is just another form of the Bellman equation.

Application to Stochastic Control.

Recall that an optimal control problem incorporates a differential equation which describes the motion of the state variable. In our analysis the latter is denoted as x. However, instead of a trivial differential equation, we have its stochastic counterpart, describing the increment of x. Now we are ready to introduce the so-called *Brownian motion* or a *Wiener process,* or *white noise.* "Brownian motion is a continuous-time scalar stochastic process such that, given the initial value x_0 at time $t=0$, the random variable x_t for any $t>0$ is normally distributed with mean $(x_0+\mu t)$ and variance $(\sigma^2 t)$...We can think of the Brownian motion as the cumulation of independent identically normally distributed increments, the infinitesimal random increment dx over the infinitesimal random time dt having mean μdt and variance $\sigma^2 dt$" (Dixit 1993, p. 1). It should be noted that the Brownian motion obeys Markov properties as well. If we had any general normal (μ, σ) variable, we could have presented it as $\mu + \sigma w$, where w has zero mean and unit variance. Similarly, we can state

$$dx = \mu dt + \sigma dw, \tag{5.4}$$

where w is a *standardized* Brownian motion whose increment, dw, has zero mean and variance dt. Moreover, the increments of dw is equal to $\epsilon_t \sqrt{dt}$, where ϵ_t is a random variable having zero mean and unit standard deviation. Equation (5.4) is also called the *Brownian motion with drift,* where μ is a drift parameter (measuring the trend), and σ is the variance parameter (measuring the volatility of the process) (Dixit and Pindyck 1994, p. 65). We will generalize the above differential equation by allowing the drift and the variance parameter to depend on the state, the control and the time variables. Thus, equation (5.4) can be re-written as

$$dx = \mu(x, y, t)dt + \sigma(x, y, t)dw.[40] \tag{5.5}$$

With all these in mind, let us refer back to equation (5.3). We can see that J depends on x. But if x is stochastic, so will be $J \equiv J(x, t)$. The question here is whether we, by knowing the increment of the stochastic process, x, can somehow deduce dJ. Why would we be interested in finding dJ in the first place? Just have a look at the last term of equation (5.3). Now, this is where we will have to rely on the elements of the *Stochastic Calculus of Itô,* which is introduced in Kamien and Schwartz (1981, p. 243). The Itô's Calculus provides us with the rules for integration of a stochastic differential equation, which are different from the ones in ordinary calculus. For example, if $dy = ydx$, we can solve for y by means of ordinary

[40] This is a generalized form of a Brownian motion.

calculus: $\frac{dy}{y} = dx \Rightarrow \int \frac{dy}{y} = \int dx \Rightarrow \ln y = x \Rightarrow y = e^x$. But if y is stochastic, then dy is a *stochastic differential* and we cannot apply the ordinary integration rule. The rule of finding the stochastic differential dy when in general $y = J(x,t)$ is known as the *Itô theorem*:

$$dy = J_t dt + J_x dx + \frac{1}{2} J_{xx} (dx)^2. \tag{5.6}$$

If y were deterministic, we could just use the ordinary rule for total differentials, which is just the sum of the first two terms on the right-hand side of (5.6)[41].

Let us now substitute (5.5) into (5.6):

$$\begin{aligned} dy &= J_t dt + J_x(\mu(x,y,t)dt + \sigma(x,y,t)dw) \\ &\quad + \frac{1}{2} J_{xx}(\mu(x,y,t)dt + \sigma(x,y,t)dw)^2 \\ &= J_t dt + J_x \mu(x,y,t)dt + J_x \sigma(x,y,t)dw \\ &\quad + \frac{1}{2} J_{xx} \mu^2(x,y,t)(dt)^2 + \frac{1}{2} J_{xx} \sigma^2(x,y,t)(dw)^2 \\ &\quad + \mu(x,y,t)\sigma(x,y,t) J_{xx} dt dw. \end{aligned}$$

Given that the Brownian motion has the properties we described above, it will follow that $(dw)^2 = dt$, $dwdt = 0$, and $(dt)^2 = 0$ (Ibid. p. 244). Then,

$$\begin{aligned} dy &= J_t dt + J_x \mu(x,y,t)dt + J_x \sigma(x,y,t)dw + \frac{1}{2} J_{xx} \sigma^2(x,y,t)dt \\ &= (J_t + J_x \mu + \frac{1}{2} J_{xx} \sigma^2)dt + J_x \sigma dw. \end{aligned} \tag{5.7}$$

Now we can use the last expression to find $E[dJ]$ for (5.3):

$$\begin{aligned} E[dJ] &= E[(J_t + J_x \mu + \frac{1}{2} J_{xx} \sigma^2)dt + J_x \sigma dw] \\ &= (J_t + J_x \mu + \frac{1}{2} J_{xx} \sigma^2)dt. \end{aligned} \tag{5.8}$$

[41]Kamien and Schwartz (1981, p. 245-246) provide some examples of the solutions to simple stochastic differental equations.

Finally, we substitute (5.8) into (5.3) and obtain

$$\rho J(x,t) = \underset{\{y\}}{Max} \left\{ \Pi(x,y,t) + J_t + J_x \mu + \frac{1}{2} J_{xx} \sigma^2 \right\}. \tag{5.9}$$

Equation (5.9) is the fundamental equation for stochastic optimal control, or *stochastic Bellman equation.*

Now, with infinite planning horizon we drop the time dependence to obtain an ordinary differential equation with x as its only independent variable:

$$\rho J(x) = \underset{\{y\}}{Max} \left\{ \Pi(x,y) + J_x \mu + \frac{1}{2} J_{xx} \sigma^2 \right\}, \tag{5.9$'$}$$

which will be particularly useful in our application.[42]

5.2. A Dynamic Continuous-time Income Tax Evasion Model with Tax Compliance Costs.

As we have previously mentioned, tax compliance costs play an important role in reality. We have already analytically analyzed the introduction of a simple tax compliance cost structure in the OLG model. Here we will do the same thing but for another tax evasion model, first developed by Lin and Yang (2001). The model is appealing for an important reason: without introducing many additional assumptions and departing from the traditional neoclassical framework, it obtains analytically rigorous and intuitively appealing results, which we discussed in Section 2.2. Further, the authors themselves acknowledged the importance of tax compliance costs on tax evasion behavior: "Some results obtained in this paper might need to be qualified after taking into account the social context. Another reason for interpreting our results cautiously is that 'transaction costs' arising from taxpayers' compliance and/or evasion activities, as in most models, are ignored in this paper. Increasing the extent

[42]Note that we arrived at equation (5.9$'$) rather in a heuristic way, for that we have not derived the stochastic counterpart of the Bellman equation "from the scratch". Instead, we heavily relied on the Bellman equation for the deterministic case. Nevertheless, the end result would be identical.

of tax evasion raises evasion costs on the one hand, but at the same it lowers compliance costs on the other hand. How the incorporation of these costs might modify our results is unclear at the present date" (Ibid. p. 1837).

We have seen in the context of an OLG model, that a very simple tax compliance costs ought to be assumed in order to keep the model tractable. Although that may not answer all the interesting questions implied by the quote above, it, nevertheless, is an attempt to consider costs of compliance in general on evasion decisions, and will serve as a basis for more comprehensive cross-model and cross-country empirical comparisons. Future works may arise from our study, in an attempt to make the modeling more realistic and interesting. The questions of simplicity and tractability are very vital for dynamic models, especially those developed in the framework of optimal control. "...[E]ven in fairly simple [dynamic] problems, the solution and analysis procedure may be quite lengthy and tedious. It is for this reason that simple specific functions are often invoked in economic models to render the solution more tractable, even though such specific functions may not be totally satisfactory from an economic point of view. It is for the same reason that writers often assume that the parameters in the problem, including the discount rate, remain constant throughout the planning period. The constancy assumption becomes especially problematic in infinite-horizon problems, where the parameters are supposed to remain at the same levels from here to eternity. Yet the cost—in terms of analytical complexity—of relaxing this and other simplifying assumptions can be extremely high" (Chiang 1992, p. 314). Thus, we modify the Lin and Yang (2001) model in a simple way, and see what implications follow.

The Model.

The model has been described briefly in Section 2.2. Thus, we start this section by deriving the key elements of the model. First, we will need the parameters for the Brownian motion. The key here, of course, is to start with a simple discrete time problem, and then extend it to the continuous time, in order to obtain proper drift and variance parameters. We suspect our results will be slightly different, since we modify the model itself a little bit. Although our derivation steps will closely follow Lin and Yang (2001), we decide to undertake them here in more details for the sake of analytical clarity (since the authors skip most of derivations).

Note that the return on a dollar of concealed income is a random variable, η, taking the value θ with likelihood $1-q$, and the value $-F\theta$ with likelihood q. That is, we consider a simple Bernoulli process, where η moves either "up" or "down" by some magnitude. Time is discrete and divided into periods Δt, meaning that the number of steps, n, is $t/\Delta t$, and the cumulated random return on concealed income with time interval of length 1 has mean

$$E(\eta) = \frac{[(1-q-Fq)\theta]}{\Delta t}. \tag{5.10}$$

The variance, call it $V(\eta)$, is by definition equal to $E(\eta^2) - [E(\eta)]^2$. Clearly, η^2 becomes θ^2 with probability $1-q$ and $F^2\theta^2$ with probability q. Hence, per time interval Δt, $E(\eta^2) = (1-q+F^2q)\theta^2$, and per time interval of length 1

$$\begin{aligned}
V(\eta) &= \frac{\left[(1-q+F^2q)\theta^2 - (1-q-Fq)^2\theta^2\right]}{\Delta t} \\
&= \frac{\left[(1-q)\theta^2 + F^2q\theta^2 - (1-q)^2\theta^2 + 2(1-q)Fq\theta^2 - F^2q^2\theta^2\right]}{\Delta t} \\
&= \frac{\left[(1-q)q\theta^2 + F^2q(1-q)\theta^2 + 2(1-q)Fq\theta^2\right]}{\Delta t} \\
&= \frac{(1-q)q[\theta^2 + F^2\theta^2 + 2F\theta^2]}{\Delta t} \\
&= \frac{(1-q)q[\theta(1+F)]^2}{\Delta t}.
\end{aligned} \tag{5.11}$$

As we see, both $E(\eta)$ and $V(\eta)$ depend on the model parameters: θ, q, F and also on Δt. Thus, the objective is to make the first and second moments of the process independent of those parameters, and reach equation (5.5) in the limit, when Δt

approaches zero. When Δt indeed approaches zero, the number of steps, n, approaches infinity, and the binomial distribution converges to a normal one. Thus, following Lin and Yang (2001, p. 1838), we set

$$\theta = \sigma(\Delta t)^{1/2}, \tag{5.12}$$

$$F = 1 - \frac{(\overline{r}\Delta t)^2}{1 - \overline{r}\Delta t}, \tag{5.13}$$

$$q = \frac{(1 - \overline{r}\Delta t)^2}{1 + (1 - \overline{r}\Delta t)^2}. \tag{5.14}$$

Note that $1 - q - Fq = 1 - q(1 + F)$. Using (5.13), we see that $1 + F = 2 - \frac{(\overline{r}\Delta t)^2}{1-\overline{r}\Delta t} = \frac{2-2\overline{r}\Delta t-(\overline{r}\Delta t)^2}{1-\overline{r}\Delta t}$. The latter, coupled with (5.14), allows us to write

$$\begin{aligned} q(1+F) &= \frac{(1-\overline{r}\Delta t)}{1+(1-\overline{r}\Delta t)^2}(2 - 2\overline{r}\Delta t - (\overline{r}\Delta t)^2) \\ &= \frac{2 - 2\overline{r}\Delta t - (\overline{r}\Delta t)^2}{2 - 2\overline{r}\Delta t + (\overline{r}\Delta t)^2}(1 - \overline{r}\Delta t). \end{aligned} \tag{5.15}$$

Then,

$$\begin{aligned} 1 - q(1+F) &= 1 - \frac{2 - 2\overline{r}\Delta t - (\overline{r}\Delta t)^2}{2 - 2\overline{r}\Delta t + (\overline{r}\Delta t)^2}(1 - \overline{r}\Delta t) \\ &= \frac{2 - 2\overline{r}\Delta t + (\overline{r}\Delta t)^2 - (2 - 2\overline{r}\Delta t - (\overline{r}\Delta t)^2)(1 - \overline{r}\Delta t)}{2 - 2\overline{r}\Delta t + (\overline{r}\Delta t)^2} \\ &= \frac{2 - 2\overline{r}\Delta t + (\overline{r}\Delta t)^2 - 2 + 2\overline{r}\Delta t + (\overline{r}\Delta t)^2 + 2\overline{r}\Delta t - 2(\overline{r}\Delta t)^2 - (\overline{r}\Delta t)^3}{2 - 2\overline{r}\Delta t + (\overline{r}\Delta t)^2} \\ &= \frac{2\overline{r}\Delta t - (\overline{r}\Delta t)^3}{2 - 2\overline{r}\Delta t + (\overline{r}\Delta t)^2} = \frac{\overline{r}\Delta t[2 - (\overline{r}\Delta t)^2]}{2 - 2\overline{r}\Delta t + (\overline{r}\Delta t)^2}. \end{aligned} \tag{5.16}$$

Now it becomes evident that

$$\begin{aligned} \lim_{\Delta t \to 0}\left[\frac{1 - q(1+F)}{\Delta t}\theta\right] &= \lim_{\Delta t \to 0}\left[\frac{\overline{r}[2 - (\overline{r}\Delta t)^2]}{2 - 2\overline{r}\Delta t + (\overline{r}\Delta t)^2}\theta\right] \\ &= \overline{r}\theta. \end{aligned} \tag{5.17}$$

Now let us consider (5.11) step-by-step. We already know that $1+F = \frac{2-2\overline{r}\Delta t-(\overline{r}\Delta t)^2}{1-\overline{r}\Delta t}$. Using (5.14) we get

$$\begin{aligned} 1-q &= 1-\frac{(1-\overline{r}\Delta t)^2}{1+(1-\overline{r}\Delta t)^2} = \frac{1}{1+(1-\overline{r}\Delta t)^2} \\ &= \frac{1}{2-2\overline{r}\Delta t+(\overline{r}\Delta t)^2}. \end{aligned} \tag{5.18}$$

Hence,

$$\begin{aligned} (1-q)q &= \frac{1}{(2-2\overline{r}\Delta t+(\overline{r}\Delta t)^2)}\frac{(1-\overline{r}\Delta t)^2}{(1+(1-\overline{r}\Delta t)^2)} \\ &= \frac{(1-\overline{r}\Delta t)^2}{(2-2\overline{r}\Delta t+(\overline{r}\Delta t)^2)^2}. \end{aligned} \tag{5.19}$$

Consequently,

$$\begin{aligned} q(1-q)(1+F)^2 &= \frac{(1-\overline{r}\Delta t)^2}{(2-2\overline{r}\Delta t+(\overline{r}\Delta t)^2)^2}\frac{(2-2\overline{r}\Delta t-(\overline{r}\Delta t)^2)^2}{(1-\overline{r}\Delta t)^2} \\ &= \frac{(2-2\overline{r}\Delta t-(\overline{r}\Delta t)^2)^2}{(2-2\overline{r}\Delta t+(\overline{r}\Delta t)^2)^2}, \end{aligned} \tag{5.20}$$

with

$$\lim_{\Delta t\to 0}\left[q(1-q)(1+F)^2\right] = \lim_{\Delta t\to 0}\left[\frac{(2-2\overline{r}\Delta t-(\overline{r}\Delta t)^2)^2}{(2-2\overline{r}\Delta t+(\overline{r}\Delta t)^2)^2}\right] = 1. \tag{5.21}$$

Finally, recall from (5.12) that $\theta^2 = \sigma^2\Delta t$. This implies that $\frac{\theta^2}{\Delta t} = \sigma^2$. Combining the latter result with (5.21) and referring to (5.11), we find that

$$\lim_{\Delta t\to 0}\left[\frac{(1-q)q[\theta(1+F)]^2}{\Delta t}\right] = \sigma^2. \tag{5.22}$$

Since the return on a dollar of concealed income is random, the corresponding change in capital stock, $\Delta\overline{k}$, is random, too.[43] Indeed, per discrete time interval, Δt,

[43]Again, here we use "bar" to show that the capital stock in our specification is not necessarily

$\Delta\overline{k}$ is distributed as a simple Bernoulli process with probability $1-q$ (with $\Delta\overline{k} = [(1-\theta+\theta\overline{e})\overline{y}-\overline{c}-\xi\overline{y}]\Delta t$), and probability q (with $\Delta\overline{k} = [(1-\theta-F\theta\overline{e})\overline{y}-\overline{c}-\xi\overline{y}]\Delta t$), where $\xi\overline{y}$ is the exogenous total cost of compliance. Note when $\xi = 0$, our model collapses to Lin and Yang's specification. Further, notice that the *change* in capital stock *cannot* be due to the factor $(1-\theta)\overline{y}-\overline{c}-\xi\overline{y}$, since the latter occurs in both states of the world. Hence, the cumulated change $(\overline{k}-\overline{k}_0)$ per a time interval of length 1 is a binomial random variable with mean

$$(1-\theta)\overline{y}-\overline{c}-\xi\overline{y}+\frac{[(1-q-Fq)\theta\overline{ey}]}{\Delta t}. \tag{5.23}$$

Using (5.17), observe that in the limit, when $\Delta t \to 0$, (5.23) collapses to

$$(1-\theta)\overline{y}-\overline{c}-\xi\overline{y}+\overline{r}\theta\overline{ey} = [1-\theta+\overline{r}\theta\overline{e}-\xi]\overline{y}-\overline{c}. \tag{5.24}$$

Again, since $(1-\theta)\overline{y}-\overline{c}-\xi\overline{y}$ is state independent, the variance of $(\overline{k}-\overline{k}_0)$ per a time interval of length 1 can be calculated as

$$\begin{aligned}
& \left\{[\theta^2\overline{e}^2\overline{y}^2(1-q)+qF^2\theta^2\overline{e}^2\overline{y}^2]-[\theta\overline{ey}(1-q-Fq)]^2\right\}/\Delta t \\
= & \left\{\theta^2\overline{e}^2\overline{y}^2[1-q+qF^2-1+2q-q^2+2qF-2q^2F-q^2F^2]\right\}/\Delta t \\
= & \left\{\theta^2\overline{e}^2\overline{y}^2[q+qF^2-q^2+2qF-2q^2F-q^2F^2]\right\}/\Delta t \\
= & \left\{\theta^2\overline{e}^2\overline{y}^2[q(1+F^2+2F)-q^2(1+2F+F^2)]\right\}/\Delta t \\
= & \left\{\theta^2\overline{e}^2\overline{y}^2[q(1-q)(1+F)^2]\right\}/\Delta t. \qquad (5.25)
\end{aligned}$$

Using (5.22), observe that in the limit, when $\Delta t \to 0$, (5.25) collapses to

$$(\overline{ey}\sigma)^2. \tag{5.26}$$

the same as in Lin and Yang (2001) model. This is because the tax compliance costs reduce the disposable income, and may also affect the behavior choice of the agent.

Expressions (5.24) and (5.26) give us the values we need for the Brownian motion. Consequently, each individual's financial asset account evolves according to the following stochastic differential equation

$$d\overline{k} = ([1 - \theta + \overline{r}\theta\overline{e} - \xi]\overline{y} - \overline{c})dt + \overline{ey}\sigma dz, \tag{5.27}$$

and

$$\overline{k}(t_0) = \text{given.}^{44} \tag{5.28}$$

Finally, the agent's problem is to

$$\underset{\{\overline{c}(t),\overline{e}(t)\}}{Max} \int_{t=0}^{+\infty} \exp(-\rho t) \ln \overline{c}(t) dt \tag{5.29}$$

subject to (5.27) and (5.28). The term associated with public goods is left out in the objective function due to the "small agent" assumption.

Using (5.9′) we write the stochastic Bellman equation:

$$\rho J(\overline{k}) = \underset{\{\overline{c}t),(\overline{e}(t)\}}{Max} \left\{ \ln \overline{c}(t) + J_{\overline{k}}([1 - \theta + \overline{r}\theta\overline{e} - \xi]\overline{y} - \overline{c}) + \frac{1}{2} J_{\overline{kk}}(\overline{ey}\sigma)^2 \right\}. \tag{5.30}$$

Setting the partial derivatives of the right-hand side of (5.30) with respect to the control variables to zero, we get

$$\overline{c}(t) = \frac{1}{J_{\overline{k}}} \tag{5.31}$$

and

$$\overline{e}(t) = -\frac{J_{\overline{k}}\overline{r}\theta}{J_{\overline{kk}}\overline{y}\sigma^2}. \tag{5.32}$$

[44] It is implicitly assumed that the government is not aware of the initial level of capital, i.e., it is not capable of determining the true flow of income (Eichhorn 2004, p. 6).

Let us substitute (5.31) and (5.32) back to (5.30):

$$\rho J(\overline{k}) = \ln\frac{1}{J_{\overline{k}}} + J_{\overline{k}}([1-\theta-\frac{J_{\overline{k}}(\overline{r}\theta)^2}{J_{\overline{kk}}\overline{y}\sigma^2}-\xi]\overline{y}-\frac{1}{J_{\overline{k}}}) + \frac{1}{2}\frac{J_{\overline{k}}^2(\overline{r}\theta)^2}{J_{\overline{kk}}\sigma^2}, \tag{5.33}$$

which, using Ak production technology, can be re-stated as

$$\rho J(\overline{k}) = \ln\frac{1}{J_{\overline{k}}} - 1 + J_{\overline{k}}(1-\theta-\xi)A\overline{k} - \frac{1}{2}\frac{J_{\overline{k}}^2(\overline{r}\theta)^2}{J_{\overline{kk}}\sigma^2}. \tag{5.33$'$}$$

Equation (5.33$'$) is a nonlinear second-order ordinary differential equation. The particular solution is

$$J(\overline{k}) = \frac{\ln A\overline{k} + \Gamma}{\rho}, \tag{5.34}$$

where $\Gamma \equiv \ln\frac{\rho}{A} - 1 + \frac{(1-\theta-\xi)A}{\rho} + \frac{(\overline{r}\theta)^2}{2\rho\sigma^2}$. Note from (5.34) that $J_{\overline{k}} = \frac{1}{\rho\overline{k}}$ and $J_{\overline{kk}} = -\frac{1}{\rho\overline{k}^2}$. Therefore, the optimal path of our control variables is of the forms

$$\overline{c}(t) = \rho\overline{k}(t) \tag{5.35}$$

and

$$\overline{e}(t) = \frac{\overline{r}\theta}{A\sigma^2}. \tag{5.36}$$

Referring to Lin and Yang (2001 p. 1832), we see that our solutions for the optimal path of consumption and the fraction of evaded income are exactly identical.[45] There are some comments worth making here, too. The general solution to (5.33$'$) has to include the natural logarithm of the capital stock, $\overline{k}$, raised to the power one. Otherwise, it would not be the solution. This is because we are dealing with $A\overline{k}$ production technology, where the capital stock enters as it is. More importantly, we see that introduction of the simple costs of compliance did not change the exact

[45]Lin and Yang (2001, p. 1832) state that $c(t) = \rho Ak$. This is an obvious (what is likely to be) typo. To see this, just differentiate their equation (9) with respect to the capital stock and plug into their equation (7).

interior solutions for the control variables. In general, even if the agent faces any proportional "sunk costs", she will still stubbornly evade exactly the same fraction of income. The same is true for the consumption. Note that $\overline{c}(t)$ stands for the total consumption per person. Nevertheless, we can re-write it as $\overline{c}(t) = \frac{\rho}{A}\overline{y}$, meaning that consumption-output ratio is constant with and without compliance costs. Observe, though, that *total* economy-wide evasion and consumption should be different since $\overline{y}$ and y differ. So, what can we conclude? The fundamental difference between the OLG model of ours and the Lin and Yang's (2001) model is that interior solutions for the latter are very rigid. This is, perhaps, due to the nature of their model, as the differential equation (5.33$'$) does not allow any solution where the capital stock does not enter the natural logarithm linearly. Now, would that be reasonable? After all, the fact that consumption and evasion ratio stay the same in our modification of the stochastic dynamic tax evasion program means that no matter how high the exogenous costs of compliance are (corresponding to various ξ), the *behavior* does not change neither in terms of relative consumption, nor in terms of evasion. With all that said, one could conclude that according to the infinite-planning horizon model considered above, "poor" economies would choose to be relatively as "dishonest" as the rich ones. Unfortunately, since there is no cross-country econometric comparisons of tax compliance, we may not be able to directly run a "horse race" between the real-world evidences and the predictions of the model variations discussed in this study. Further, in the calibration section of this study we will try to choose a reasonable value for ξ in the U.S. and see how evasion would respond to that. Finally, we are planning to investigate how reliable the infinite-horizon model is in terms of fitting the reality. But we will render a simple empirical test of Lin and Yang (2001) model to the following chapter.

Conclusions.

We started this chapter by reviewing the basics of dynamic programming and control theory. We, then, analyzed a pioneering study of income tax evasion in the context of an infinite-planning horizon dynamic framework. To be able to compare a two-period OLG model developed above and a continuous-time model, we introduced the exogenous costs of compliance and obtained the analytical solutions for the optimal fraction of evasion and consumption. The most striking observation is that evasion and consumption as a fraction of income does not change in the stochastic control problem, while it does in the OLG model. That sounds counter-intuitive, since ξ is indeed capable of differentiating between the economies of different wealth levels. And we might expect that the economies with different wealth levels to differ in their compliance attitudes. Therefore, the analysis of this section alone should provoke future research on the tax evasion behavior when time is continuous, and households are infinitely-lived.

Chapter Six: Data, Simulations and Discussion.

6.1. Introduction.

In this chapter we will first describe the data to be used for our simulation purposes. Then we will carefully simulate our theoretical findings. Our main focus will be on the U.S. economy, for which we will also show the welfare implications of lower audit rates. In addition, we will also calibrate the value of the revenue-maximizing tax rate and see how different it is from the existing average tax rate.

We will concisely test the explanatory power of the OLG model for a selected economies, for which we were able to gather the relevant microdata. It is worth noting that the parameters governing the behavior of taxpayers might vary across countries due to the differences in institutional settings. Calibrating and simulating analytical solutions for evasion, consumption and saving, and cross-checking the results with some real-world evidences will help to critically evaluate our theoretical findings.

Finally, we will conduct a brief empirical test of the Lin and Yang (2001) model in the context of the U.S. economy. The above study is indeed purely theoretical. As Eichorn (2004, p. 17) has suggested: "[a]n empirical estimation of the model might be interesting", and we will attempt to do just that.

Data.

Data for the parameter values will be mostly gathered from previous studies, including Manasan (1988), Maddison (1992), Ríos-Rull (1996), Andreoni et al. (1998), Galasso (1999), Ambler (2000), Altig et al. (2001), De La Croix and Michel (2002), Chen (2003), Gupta (2004), Niepelt (2005), Sandford (2005), which are conventional sources of reference in the literature of consumption, saving and tax evasion. We will rely on the saving rates reported by the World Bank World Development Indicators (The World Bank Group 2007), which are consistent with those in Maddison (1992) who estimates the long-run gross saving rates for 11 countries. WDI data are also useful in tracking down the average population growth rate for various countries. For that purpose we use the average value for 1965-1995, since most of the available model parameters fall within that range. Ríos-Rull (1996), Galasso (1999), Ambler (2000) and Altig et al. (2001) carefully simulate different OLG models, while De La Croix and Michel (2002) is the classic modern reference in OLG modelling. The latter comprehensively covers the theory of OLG models, and extensively discusses its various policy implications. The book also concisely provides some guides for simulation within two-period OLG frameworks, which will prove useful for the purposes of this study.[46] Most of our simulation analysis below will be based on the authors presentations. Chen (2003) models and calibrates tax evasion in the context of endogenous growth, and incorporates the Ak model, thus, being useful to compare some of the microdata with the predictions of the Lin and Yang (2001) framework. We will also rely on Manasan (1988), Andreoni et al. (1998) and Gupta (2000), for they provide various tax compliance and microdata for a number of industrialized and developing economies. In addition, Sandford (1995) will also serve as one of the most recent sources for tax compliance costs around the world.

Simulation and Discussion: the Case of the Falling Audit Rate.

We first calibrate our theoretical results for the U.S., and then focus on the welfare impact of low audit rates on the economy. Recent dramatic decline in the U.S. income tax audit rate raised many concerns. Existing static models of tax evasion predict that such a decline in the audit rate, if accompanied by a decrease in taxpayers' subjective evaluation of getting caught, unequivocally lowers the amount of declared income. However, we have theoretically shown in our general equilibrium OLG model of tax evasion that the magnitude of the decrease in income declaration critically depends

[46] At this point we would like to thank Professor De La Croix for his cooperation, which helped us in the simulation section. The usual disclaimer applies.

on the time span and may intuitively increase in the long run.

In calibrating an OLG model one should keep in mind that each period corresponds roughly to half a lifetime. We assume each period lasts for 30 years. We set annualized population growth rate and the rate of technological change equal to 1.24 and 1 percents, respectively. Then $1+n = (1+0.0124)^{30} \approx 1.45$ and $1+g = (1+0.01)^{30} \approx 1.35$. To match the long-run capital-output ratio and the interest rate, we take the annual rate of time preference roughly equal to 2.71%, implying that $1+\rho = (1+0.0271)^{30} \approx 2.23$. If there are no market imperfections, no externalities and no unemployment, then we can take the interest rate in the loanable funds market as the appropriate proxy for the individual discount rate. In the U.S., for instance, the returns on long-term, high-grade bonds in the postwar period averaged at about 3.72% (McGrattan and Prescott 2003, p. 395)[47]

We set $F = 1.75$, $\theta = 0.3$ and the annual tax audit rate equal to 1.7%, corresponding to the mid 90s statistics. The latter parameter usually is taken as a proxy for the probability of being caught, as perceived by the taxpayer, meaning that in our set-up $q = 0.51$.

In the U.S. approximately 4% of all audited taxpayers end-up paying penalties for "...fraud, negligence, false withholding, failure to report tips, or other miscellaneous infractions" (Andreoni 1998, p. 821). This corresponds to β equal to 0.0204.

The parameter values we use for the simulation purposes are summarized below in Table 1. Consequently, the *annual* steady-state capital-output ratio, $30k^{*1-\alpha}$, is about 2.43, the long-run saving rate ($\equiv \frac{K_{t+1}}{Y_t} = \frac{K_{t+1}}{A_t L_t k^{*\alpha}} = \frac{K_{t+1}(1+n)(1+g)}{A_{t+1}L_{t+1}k^{*\alpha}} = (1+n)(1+g)k^{*1-\alpha}$) and the annual interest rate are 15.79 and 5.102 ($\Leftarrow \sqrt[30]{\alpha k^{*\alpha-1}} - 1$) percents, respectively. We also provide in Table 2 the values for capital per effective worker and intensive output on the balanced growth path.[48] Note that the rates of return

[47] We are rounding the numbers in the text, while the subsequent macro variables will be calculated more precisely.

[48] The parameter values and results we have are broadly consistent with those in standard macro models and studies, which are summarized in previous section (see exclusively, yet not exhaustively:

are converted to their annual equivalents.

Table 1. Parameterization of the baseline economy.

Symbol	**Definition**	**Values**
n	Per period population growth rate	0.45
g	Per period rate of technological change	0.35
ρ	Per period rate of time preference	1.23
α	Relative share of capital	0.36
θ	Income tax rate	0.3
q	Per period subjective probability of detection	0.51
F	Penalty rate	1.75
β	Per period conviction rate	0.0204

Table 2. Preliminary simulation results.

Symbol	**Definition**	**Value**
k^*	Steady-state capital per effective worker	0.01967
$f(k^*)$	Steady-state intensive output	0.243
$n+g$	Per period output growth rate	0.80
K/Y	Annualized capital-output ratio	2.428
s	Long-run saving rate (%)	15.785
r	Annualized interest rate (%)	5.101

It is, perhaps, of somewhat greater interest to analyze the impact of a falling audit rate on the well-being of the generation, born in period $t = 1$ (at the time of the shock), and thereafter. If not caught, the life-cycle utility of generation t will

Maddison 1992, p. 185, Ríos-Rull 1996, pp. 474-475, Andreoni et. al. 1998, p. 821, Altig et. al. 2001, pp. 580-584, De La Croix and Michel 2002, p. 339, Niepelt 2005, p. 1624 and so on). Also recall that A_t is just a scale parameter and we will set it for the convenience equal to 75 to generate the capital-labor ratio around 1.5.

increase through its favorable impact on saving (*less* income is reported to the tax authority, i.e., more funds flow into the financial asset account), but not so much for the generations to come, since rising k (and thus, real wages) in subsequent periods will encourage *more* income declaration. More importantly, since the capital stock increases, the rate of return on saving falls (recall that the return factor on saving from the perspective of t, R_{t+1}, is equal to $f'(k_{t+1}) = \alpha k_{t+1}^{\alpha-1}$), thus, adversely affecting per-period utility. We can unequivocally conclude that the shock (declined q) permanently decreases interest payments. But which effect dominates? The one which tends to boost lifetime resources due to more tax evasion, or, the second one, decreasing the return on saving? Thus, we are establishing here the following proposition.

Proposition 7 *While the caught taxpayer at the time of the shock will be relatively worse-off due to the decline in audit rate, the impact of the lower audit rate on the not caught taxpayer's life-cycle utility, relative to the previous generation's utility, is ambiguous.*

Proof. First, observe that the maximized expected life-cycle utility of generation t is given as

$$\begin{aligned} E[U] &= \underbrace{\ln\left(\frac{(1+\rho)(1-\theta)}{2+\rho}A_t(1-\alpha)k_t^{\alpha}\right)}_{\equiv C_{1t}^*} \\ &+\frac{1-q}{1+\rho}\ln\underbrace{\left(\alpha k_{t+1}^{\alpha-1}\left[\frac{F(1-\theta)(1-q)}{(F-1)(2+\rho)}A_t(1-\alpha)k_t^{\alpha}\right]\right)}_{\equiv C_{2t+1}^{nc}=R_{t+1}s_{nc}} \\ &+\frac{q}{1+\rho}\ln\underbrace{\left(\alpha k_{t+1}^{\alpha-1}\left[\frac{F(1-\theta)q}{2+\rho}A_t(1-\alpha)k_t^{\alpha}\right]\right)}_{\equiv C_{2t+1}^{c}=R_{t+1}s_c}, \end{aligned} \tag{6.1}$$

where s_{nc} and s_c are per capita saving in two states of the world, respectively. Certainly, *ex-post* utilities for two different persons (caught and not caught), born at time t, are given by

$$U^c = \ln\left(\frac{(1+\rho)(1-\theta)}{2+\rho}A_t(1-\alpha)k_t^\alpha\right) + \frac{1}{1+\rho}\ln\left(\alpha k_{t+1}^{\alpha-1}\left[\frac{F(1-\theta)q}{2+\rho}A_t(1-\alpha)k_t^\alpha\right]\right), \tag{6.2}$$

and

$$U^{nc} = \ln\left(\frac{(1+\rho)(1-\theta)}{2+\rho}A_t(1-\alpha)k_t^\alpha\right) + \frac{1}{1+\rho}\ln\left(\alpha k_{t+1}^{\alpha-1}\left[\frac{F(1-\theta)(1-q)}{(F-1)(2+\rho)}A_t(1-\alpha)k_t^\alpha\right]\right). \tag{6.3}$$

Now, let us take the derivative of (6.2) with respect to q. Remember, the first term will vanish since the capital stock at time t is *fixed* by the decisions about saving in $t-1$. We get

$$\frac{\partial U^c}{\partial q} = \frac{1}{(1+\rho)C_{2t+1}^{nc}}[\alpha k_{t+1}^{\alpha-1}\left[\frac{F(1-\theta)}{2+\rho}A_t(1-\alpha)k_t^\alpha\right] + \alpha\frac{\partial(k_{t+1}^{\alpha-1})}{\partial q}\frac{F(1-\theta)q}{2+\rho}A_t(1-\alpha)k_t^\alpha] \tag{6.2$'$}$$

Since $\frac{\partial(k_{t+1}^{\alpha-1})}{\partial q} = (\alpha-1)k_{t+1}^{\alpha-2}\frac{\partial k_{t+1}}{\partial q}$, and we know that $\frac{\partial k_{t+1}}{\partial q}$ is negative, it follows that $\frac{\partial U^c}{\partial q}$ is positive.[49] Analogously,

$$\frac{\partial U^{nc}}{\partial q} = \frac{1}{(1+\rho)C_{2t+1}^{c}}[\alpha k_{t+1}^{\alpha-1}\left[\frac{F(1-\theta)}{(F-1)(2+\rho)}A_t(\alpha-1)k_t^\alpha\right] + \alpha\frac{\partial(k_{t+1}^{\alpha-1})}{\partial q}\frac{F(1-\theta)(1-q)}{(F-1)(2+\rho)}A_t(1-\alpha)k_t^\alpha] \tag{6.3$'$}$$

Clearly, (6.3$'$) is ambiguous in sign. ∎

That is, when the probability of getting caught declines, the agent born and caught

[49]Remember, k_{t+1} is the capital stock per effective worker in the next period; it is *not* the capital stock at the new long-run equilibrium.

in time t will be worse off. Interestingly, we *cannot* say whether the taxpayer who were able to get away with cheating is better off when q falls! Yes, when old she will face a lower return on her savings, but she also has more funds in her financial asset account to start with: she evaded more, and none noticed that! While we know that the taxpayer who is caught for sure will be *worse off*, we cannot say that the one never caught for sure will be *better off*.

Speaking about the welfare of the caught taxpayers born *after* the shock, we cannot tell whether it will be lower relative to the time $t = 0$. After the shock, the "post-shock" generations' paychecks are higher, since the capital stock in the economy is higher.

Let us see what simulation has to say about our theoretical findings. Assume that at time $t = 0$, the economy is at the steady-steady equilibrium. For illustrative purposes, assume that at time $t = 1$, q permanently declines by 35% (= 1.105). We purposely take here a conservative estimate for the decline in the likelihood of getting caught. First, after 70% initial decline in the audit rate, there were some increases in the rate. Further, note that a lower audit rate does not necessarily mean that fewer people will be caught cheating, and the taxpayers should be aware of that. In addition, it might take some time for the public to learn that the audit rate fell well below its historic average. Finally, our qualitative conclusions will be the same even when q declines by 70%. The new value of q then becomes 0.3315. Assume no change in β. As a result, the capital stock increases and reaches the value of 0.03154 at the new balanced growth path. However, the speed of adjustment is pretty slow: it takes about 270 years to "almost" hit the new steady-state! Although, the majority of empirical studies would assume even slower convergence of the economy to the new balanced growth path (De La Croix and Michel 2002, p. 339).

Now we are ready to demonstrate the results of the simulation exercise. They can be seen below in Table 3.

Table 3. A permanent fall in the audit rate.[50]

t	k	X^*	*Ann. Int. R.*, %	$E[U]$	U^c	U^{nc}
0	0.01967	10.46035	5.102	2.8183	2.7639	2.8750
1	0.01967	6.95004	5.102	2.7735	2.4771	2.9205
2	0.02727	7.81727	4.372	2.9164	2.6199	3.0634
3	0.03000	8.09098	4.159	2.9570	2.6605	3.1040
4	0.03099	8.18559	4.087	2.9705	2.6741	3.1176
5	0.03134	8.21918	4.062	2.9753	2.6789	3.1223
6	0.03147	8.23121	4.053	2.9770	2.6806	3.1240
7	0.03152	8.23554	4.050	2.9776	2.6812	3.1246
8	0.03153	8.23709	4.049	2.9778	2.6814	3.1248
9	0.03154	8.23765	4.048	2.9779	2.6815	3.1249
$+\infty$	0.03154	8.23785	4.048	2.9779	2.6815	3.1250

The amount of declared income *falls immediately* due to the behavioral effect, then starts *rising* due to the productivity effect, but still remains low. But what about case-by-case life-cycle utilities of those people who are lucky to avoid the audit, versus those who are not so fortunate? By utilizing equations (6.3) and (6.2), we obtain the last two columns in Table 3. We see that the expected utility of a typical individual is strictly lower in the period of the shock by about 1.59%. Further, the life-cycle utility of caught at $t = 1$ would decrease quite significantly by 10.38%, while that of not caught would increase relatively moderately by roughly 1.58%.

We can conclude that calibrating the model for the U.S. economy produces quite reasonable values for the existing macrodata. Recall that the parameters of our model

[50]Recall that we assume everyone in the economy is a tax cheater. Although such an approach is very common in the tax evasion literature, it still might be interesting to consider a share of the population who (for some reason) never run the expected utility maximization problem, and just decide how much to consume out of their after honestly paying meeting tax liabilities. Mixing two types of individuals and considering how they fair in the wake of the falling audit rate, might be quite interesting.

were chosen based on existing data in the literature. The only parameter we had to adjust was ρ, but we took its value equal to 2.71% per year, which is within the range of estimates for the U.S. (see, e.g., Galasso, 1999 and De La Croix and Michel, 2002). We were also able to analyze within our framework the impact of the falling U.S. audit rate on the economy, as well as on the well-being of individuals.

In addition, our model predicts that the amount of declared income of an average person, $X_{1t}/w_{1t}A_t$, is about 89.65%, which is broadly consistent with existing empirical evidence. For instance, Andreoni et al. (1998, p. 822), based on the 1992 IRS study, report that "...91.7 percent of all [U.S.] income that should have been reported, was in fact reported". According to Engel and Hines (1998, p. 2), individuals in the U.S. underreport about 10.6% of their incomes annually. Hence, we can reasonably assume that an average American tax cheater understates her income roughly by 10%. Further, if we were to refer to the OLG model with simple compliance costs, we had to make an assumption about the value of ξ, which is about 1.5% in the U.S.[51] If so, the fraction of declared income increases and stays at 90.55%, which is even closer to the existing estimates. Thus, we can conclude with a fairly good sense of confidence that our model has an acceptable fitting power for the U.S. economy. But can we say the same for the rest of the world? Will our model have a good predicting power for only some major developed countries, or not necessarily so? This question is important because country-specific institutional factors might affect tax compliance. As Das-Gupta (2004, p. 6) reports: "[r]esearch, primarily in the United States, suggests that what maybe termed "cultural" factors may significantly influence taxpayer attitudes. Included in this are such things as fiscal knowledge, income and social class, risk aversion, risk, age, sex, occupation, peer attitudes to tax evasion and bribe payment, deterrence to authority, and acquaintance with tax offenders". Hence, we are

[51] This value is obtained from Blumenthal and Slemrod (1995, p. 152) (Ed.). They state the total resource cost in 1989, on average, at $354 in real terms per household, which is roughly the same as in 1982; That is, $\xi \approx 0.015$ (we assume 2000 hours worked per year and $11.5 in wages per hour; for the latter figure refer to the Bureau of Labor Statistics' Historical Listings 2002).

interested in seeing how our model explains the tax compliance behavior in various countries with different social and institutional structures.

Finally, note that unlike the Lin and Yang (2001) model, which shows that tax evasion unambiguously falls when tax rate falls, the impact is ambiguous in our model. Indeed, a complete removal of ambiguity is not necessarily a good thing: as we have mentioned above, there are studies predicting a positive relationship between compliance and the tax rate. It is also interesting to see how the Lin and Yang (2001) model fits the reality, and how we can draw some comparisons between their and our OLG model. We will try to address those in the following section.

6.2. Brief Cross-country Comparisons Regarding Noncompliance.

Unfortunately, there are not many studies that provide tax compliance estimates for a variety of countries, and the few that are available, are very difficult to obtain. Therefore, we will review only those economies, for which there exists some relatively easily accessible empirical studies of income tax evasion. We will also try to consider both developed and developing countries to see whether the model results we derived are sensitive to a country specifications. The countries under consideration are: Germany, UK, Italy, Belgium, Spain, Philippines and Jamaica. The data for the income tax rates in the European countries as well the share of capital in total value added are taken from Gupta (2004). We assume that the technological progress in Jamaica, Philippines, and in the U.S. was the same, while that in European countries is averaged at 1.6%. The latter is broadly consistent with existing evidence (e.g., Maddison 1991). Fortunately, the results are very robust to that parameter. As far as the enforcement parameters are concerned, we used similar parameters for both Jamaica and Philippines ($q = 0.25$ and $F = 1.5$).[52] That penalty rate is consistent

[52] Alm et al (1993, p. 14) state that in Jamaica the mean value of the income reported by the taxpayer was J$7,123, while the corrected income after the audit is averaged at J$12,585. That roughly corresponds to $X_{1t}/w_{1t}A_t = 56.6\%$. The effective income tax rate is taken at 43.3%.

with the evidence from Alm et al. (1993) for Jamaica. The per-period probability of getting caught equal to 0.25 implies that less than 1% of returns are audited yearly (see, e.g., Atolia 2003). We use the relatively high U.S. estimates of the enforcement parameters for the Germany and U.K. and use more conservative values for Spain, Belgium and Italy, where tax evasion is historically higher (namely, we set $q = 0.48$ and $F = 1.65$).

Finally, data for the population growth rates and saving rates are derived from the World Development Indicators (1965-1995). Table 4 shows the average actual size of tax evasion (as a fraction of the true income), as well as that predicted by the OLG model.

Table 4. Some cross-country evidences on non-compliance.

Countries	*Actual evasion* (%)	*Predicted evasion* (%)	$\lvert Error(\%) \rvert$
Germany	13	13.6	4.6
UK	11.5	11.4	0.09
Italy	21.4	26.3	22.9
Belgium	18.4	21.8	18.5
Spain	19	32	68.4
Philippines	55	56.5	2.7
Jamaica	56.6	43.3	23.5

The third column in Table 4 is simply a forecasting error. The key parameter value we choose is the rate of time preference, in order to match the data on the saving rate. For every country, the parameter ρ more or less falls within reasonable ranges (e.g., Philippines with the highest annual discount rate about 3.86%, and in

Further, for both Jamaica and Philippines we take α equal to 0.40, which is consistent with the empirical estimates for the majority of developing countries (Atolia 2003, p. 20). Finally, according to Easterly and Rebelo (1992, p. 18), the marginal tax rate in Philippines is 35%.

Belgium it was about 0.07%—extremely low discount rate). Remember that both the amount of declared income and the actual income are flow variables, i.e., "long-run evasion" corresponds to the annual one.

We clearly see that our model does a very good job in explaining the evasion attitudes in some countries, while a poor job for some other ones. The main reason is a sensitivity of the results to the enforcement parameters, especially to the likelihood of getting caught. On the other hand, the lack of precise parameter values does not let us very accurately evaluate the predictive power of the OLG tax evasion model. Thus, the brief analysis undertaken in this section raises some important questions for the future research elaborations. Namely, whether there are any additional factors (cultural, institutional, etc.) we are not capturing in our simple neoclassical framework.

On the "Optimal" Income Tax Rate.

Recall expression (4.25), describing the level of the tax rate the government should choose in order to maximize the long-run revenue per worker. Using the parameter values we have chosen to represent the U.S. economy, we find $\theta^* \approx 65.5\%$.[53] According to Barro and Sahasakul (1986, p. 563), the average marginal tax rate from the Social Security and the individual income tax for the year 1983 was 33.9%. Based on their and our findings, can we conclude that the present tax rate in the U.S. is "inefficiently" low? Not at all. There are at least two things that we should always keep in mind. First, we did not take into account the deadweight-loss due to the tax-induced distortion of the work choice. Second, it might be that indeed, higher tax rates discourage tax compliance even in the "short-run" (contrary to our behavioral effect), and thus, the corresponding estimate of θ^* can be much more conservative than what our model suggests. These observations lay important justifications for

[53] Though, if you assume inelastic labor supply, then combined Social Security payroll tax (10.6%) and the highest marginal tax rate for every filing status (35%) will come much closer to θ^*.

the future modifications and extensions of our theoretical model.

6.3. Some Notes on the Empirics of the Lin and Yang (2001) Model.

Ever since Lin and Yang proposed a dynamic model of income tax evasion with infinite-planning horizon that completely removes the Yitzhaki puzzle, the model has never been tested empirically. We do not know up to now whether the model can fit the reality well, based on the reasonable values of the model's key parameters. We have seen that our two-period OLG model has relatively attractive explanatory power, without assuming unrealistic values for the parameters. Can we say the same thing about the Lin and Yang (2001) model specification? Although we are not trying to undertake a comprehensive empirical testing of their model, we, nevertheless, are going to see whether it can explain some given macroeconomic trends, without placing too restrictive requirements on the parameter values.

In this section we will consider the original formulation of the Lin and Yang model, i.e., when in our reformulation ξ is given the value of zero. We do so for the sake of brevity, since if ξ is non-zero (say, equal to 0.015 for the U.S. economy), our empirical conclusions will be very similar. Thus, to save space, we start with original model specification, and thus, drop the "bar" notation from now on.

We start by asking ourselves a question: what basic real-world macro evidences can be supported by the model's main results? First, let us rewrite formula (5.35), so that consumption per person, $c(t)$, becomes $\frac{\rho}{A}y$ (recall that output per person, y, is $Ak(t)$). Since individuals are identical and assumed to be all household-producers, the consumption-GDP ratio becomes ρ/A, which is about 70% in the U.S. The parameter A is just a scale parameter, so its value of little interest. However, most dynamic models with infinitely lived households and traditional real-business-cycles and tax evasion models assume ρ to be around few percentage points. Second, according to (5.36), a fraction of income evaded, $e(t)$, is $\frac{\bar{r}\theta}{A\sigma^2}$, which, for an average U.S. tax cheater is estimated to be 10%. In addition, Lin and Yang (2001, p. 1833) show that in the presence of tax evasion, the growth rate of output per capita is $(1-\theta)A+(\frac{\overline{r}\theta}{\sigma})^2-\rho$,[54] which, according to the U.S. 1965-1995 World Bank data, is averaged to be about 2.1% yearly.

Therefore, we will try to track down the values of the model parameters (A, ρ, σ^2) by solving the following system of equations (6.4)-(6.6):

$$\frac{\rho}{A} = 0.7, \tag{6.4}$$

$$\frac{\overline{r}\theta}{A\sigma^2} = 0.1, \tag{6.5}$$

$$(1-\theta)A + (\frac{\overline{r}\theta}{\sigma})^2 - \rho = 0.021. \tag{6.6}$$

The above system generates unique solutions, presented by (6.7)-(6.9) below:

$$A = \frac{0.021}{1-\theta+0.1\overline{r}\theta-0.7}, \tag{6.7}$$

$$\rho = \frac{0.0147}{1-\theta+0.1\overline{r}\theta-0.7}, \tag{6.8}$$

$$\sigma^2 = \frac{10\overline{r}\theta(1-\theta+0.1\overline{r}\theta-0.7)}{0.021}. \tag{6.9}$$

Recall that $\overline{r} = 1-q-qF$, which is roughly 0.953 for our baseline annualized model parameters. The latter number is consistent with the literature (e.g., see Dhami and Al-Nowaihi 2006, p. 1). With $\theta = 0.3$, we find that $A \approx 0.735$, $\rho \approx 0.514$, and $\sigma^2 \approx 3.892$. We see that the rate of time preference of the average agent has to be unusually high, equal to 51.4%! That is way above the level considered in the

[54] To see this, let us substitute (5.35) and (5.36) into the stochastic differential equation (5.27), and divide both sides by k (remember to drop "bar" and let ξ be zero). We get $\frac{dk}{k} = [(1-\theta)A+(\frac{\overline{r}\theta}{\sigma})^2 - \rho]dt + \frac{\overline{r}\theta}{\sigma}dz$. Recall that for the case of Ak production technology, $\frac{\dot{y}}{y} = \frac{\dot{k}}{k}$. Further, although tax evasion is risky for an individual, that risk, when spread throughout a large economy, is negligible. Hence, we simply take the expected value of $\frac{dk}{k}$ and set it equal to the growth rate of the output per person.

literature. Nevertheless, we will proceed by trying to gain a better view of the key model variable, the capital stock. For that purpose we will simulate its path, and then compare it with some real world evidences.

Recall that the change of the capital stock is in the form of the following stochastic differential equation: $dk = ([1-\theta+\bar{r}\theta e]y - c)dt + ey\sigma dz = ([1-\theta+\bar{r}\theta e]A - \rho)kdt + (eA\sigma)kdz$. By using our simulation numbers, we can state that

$$dk = 0.02099kdt + 0.14491kdz. \tag{6.10}$$

Equation (6.10) is a special case of the *generalized Brownian Motion* (also called the *Itô process*), where the drift and variance coefficients are the functions of current state, and thus, time (exactly as in equation (5.5)). Expression (6.10) is the special case indeed, called the *geometric Brownian motion with drift*, and the drift and variance coefficients are just constants, as opposed to being the functions. As stated in Dixit and Pindyck (1994, p. 71), the expected value of k itself, $E[k]$, is $k_0 \exp(\alpha t)$, where α is the drift, equal to 0.02099 in our case. The latter tells us that the annual expected growth rate of the capital per person is approximately 2.1% (equal to growth rate of GDP per capita, given the parameters we have chosen), while the annual standard deviation is about 14.5%. For comparison, the annual expected growth rate and standard deviation of the New York Stock Exchange Index is roughly equal to 9 and 20 percents, respectively (Ibid. p. 72).

Finally, we are ready to simulate the actual path of the capital stock. Following Dixit and Pindyck (1994, p. 72), we calculate $k(t)$ using the equation

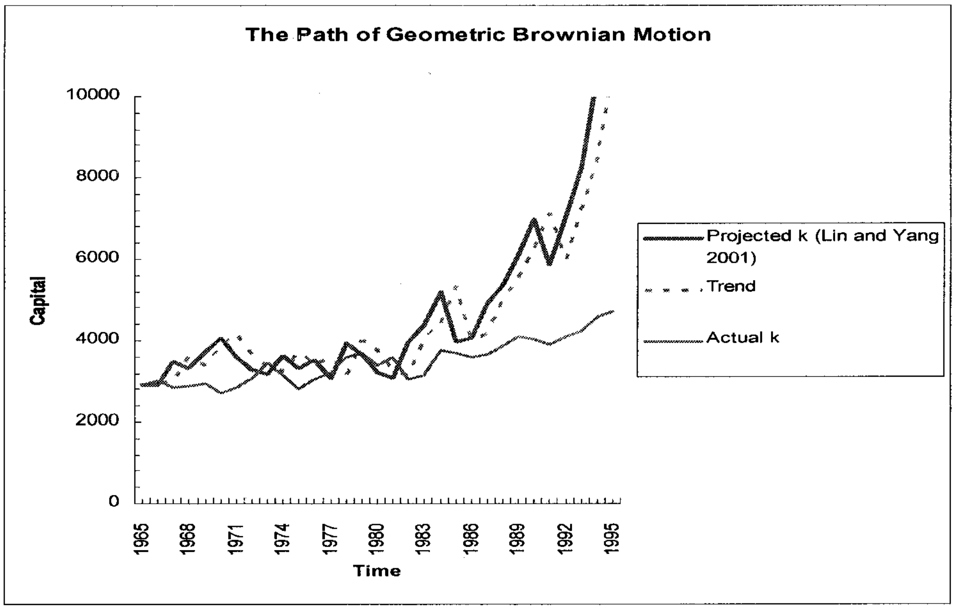
$$k_t = 1.02121k_{t-1} + 0.14491k_{t-1}\epsilon_t.[55] \tag{6.11}$$

At each time t, we will draw ϵ_t from a normal distribution with the first and

[55]The coefficient $1.02121 = \exp(0.02099)$.

second moments equal to zero and one, respectively. For better comparison with the real world data, we take the initial value k_{t-1} equal to 2932, since this is what per person gross real domestic savings (as a proxy for the capital stock) in the U.S. was in 1965. We aim to simulate the path of the capital stock over the time period 1965 to 1995. The results of the simulation is shown in Figure 1 below. The graph is shown in annual terms. We also report the actual domestic savings per person, as well as a trend line, corresponding to equation (6.11) with $\epsilon_t = 0$.

Figure 1.



We see that the actual trend of the capital stock in the U.S. is relatively stable (the growth rate is about 1.9% per year, and the relative standard deviation is about 15.5%, the latter being not considerably different from that of the projected path). However, we see that except the 1972-1982 period, the projected path was consistently overestimating the actual path of the capital. Further, the projections are way too "optimistic" to be realistic: the capital stock average annual growth rate is roughly 5.5%, almost three times as high as the actual one.

To sum up, we found that although Lin and Yang (2001) completely remove the ambiguity from the A-S model by stretching the planning horizon till infinity, the predictive power of their model is rather weak, if not completely inadequate. They key variable of the model, the capital stock, shows way too optimistic growth rate, and the economic agents have to be assumed unrealistically impatient. The model solves one important theoretical puzzle at the expense of being too unattractive from any practical considerations.

Conclusions.

In this chapter we simulated our two-period OLG model and showed that under reasonable assumptions it describes well the state of the U.S. economy. Then, we considered a decline in the tax audit rate and demonstrated that the taxpayers who are caught cheating at the time of the shock will be strictly worse off, while those who are never caught might be worse off, too, from theoretical perspectives. Our simulation results revealed that the expected life-cycle utility of the generation born at the time of the shock were lower. We also simulated the model for a variety of countries and found that the predictions of noncompliance are very sensitive to the change in the enforcement parameters. The revenue-maximizing tax rate is found to be very high, as one might have expected intuitively.

Finally, for the first time we have carefully simulated the main findings of the Lin and Yang (2001) model with infinite-horizon. We were able to show that the model fails to fit the reality under reasonable parameter values.

Chapter Seven: Summary and Some Remarks on Future Extensions.

In this study we carefully reviewed the main theoretical studies devoted to the income tax evasion phenomenon. We saw the result obtained by Yizthaki (1974), — about the inverse relationship between the tax rate and the amount of income concealed, — is central for the majority of theoretical models. Current solutions for the Yitzhaki puzzle share a common feature: predominantly they all deviate from the original framework substantially by, e.g., endogenizing the labor supply decision (and assuming a backward-bending labor supply curve), introducing stigma costs, or even suggesting to abandon the expected utility framework, and assuming that the taxpayers are all loss-averse.

Instead, we were able to resolve the Yitzhaki puzzle in the context of the original framework, simply by introducing a progressive income tax rate structure (which is realistic) without a need for additional assumptions. Since there are studies suggesting that people become more honest when the marginal tax rate goes up, we have never considered the resolution of the Yizthaki puzzle to be of the utmost importance to our theoretical analysis in this study.

Further, we developed a new two-period overlapping-generations model of income tax evasion and solved it analytically. We saw how the short-run implications on compliance may differ from the long-run ones. We also theoretically treated the government's revenue-maximizing objective on the choice of the appropriate tax rate. We, then, incorporated a simple tax compliance structure in the model, and resolved it. The main reason was to see how it affects the decision-making. We saw that both the absolute and the relative compliance decisions would change, which can be plausible.

Then, we switched our interest to the Lin and Yang (2001) pioneering study of tax evasion in the context of the stochastic optimal control, incorporating infinitely-lived households. We introduced the cost of compliance into their model as well, and found that no matter how high they are, the taxpayers will not change much their decisions on compliance.

In the calibration part of our study we simulated the theoretical findings of the OLG model both for the U.S. and a handful of developed and developing countries. We observed that the OLG model has a good explanatory power for the U.S. Perhaps, the most interesting result in that part of our study was the intergenerational welfare comparison for the taxpayers when facing different tax audit rates. We found that the taxpayers, who comply less at the initial time when the audit rate is low, will be strictly worse-off in terms of their expected life-cycle utilities. Finally, we undertook a careful simulation of the Lin and Yang (2001) model, and showed that their study, unlike ours, must tolerate very unreasonable parameterization in order to be able to pass the test for explaining the reality reasonably well.

However, the results of the present study suggest some important extensions for future work. It would be interesting to see whether an explicit tax rate schedule (with progressive marginal rate structure) will preserve the theoretical ambiguity found in our extension of the static Allingham-Sandmo (1972) model. Furthermore, how progressive the tax rate schedule must be to show a negative relationship between the tax rate and the amount of income reported?

Our analysis considered a revenue-raising, not a revenue-using government. In the OLG model we assumed that the government sector is completely "unproductive". That sounds defective, since intuitively, one of the objectives of many tax evaders is to free ride on government provided public goods. Further, public expenditure may not only affect the utility of the individual, but also the steady-state level of the economy, via enhancing the infrastructure and/or partially crowding out private investments. In addition, if taxpayers believe that public goods are over-provided, then higher tax rate will encourage tax evasion because such an increase in the tax rate worsens the feeling of overproducing. To compensate, the taxpayers with decreasing absolute risk aversion will choose to comply less (e.g., Cowell and Gordon 1988, pp. 318-319; Cullis and Jones 1998, p. 200).

In addition, an incorporation of the labor-leisure choice would be interesting as well, for at least two reasons. First, it would allow us to design a more realistic revenue-maximizing tax rate, taking into account the excess burden of income taxation. Second, it may help to introduce an otherwise absent behavioral effect (in the manner of Cowell (1985)), stemming from higher tax rates.

In most models of income tax evasion it is assumed that the agents are identical in terms of their earning capabilities. That assumption, of course, is inconsistent with the reality. In reality one might expect poor people to evade taxes, while richer people to avoid them. On the contrary, there can be much truth to the popular view that richer people are the ones who violate the laws. Introducing agents with different earning capabilities might be quite a difficult task, but can help to reconcile many of the aforementioned issues.

Finally, to better assess the immediate economic repercussions of the change in various policy parameters, a more realistic multi-period overlapping-generations model can be designed. An immediate drawback is that an analytical solution may no longer be feasible to obtain (see, e.g., İmrohoroğlu et al. (2003)). Thus, as usual, we would have to face a trade-off between realism and tractability which is worth considering at least for the sake of analytical curiosity.

References

[1] Abel, A. B., N. G. Mankiw, L. H. Summers and R. J. Zeckhauser (1989), 'Assessing Dynamic Efficiency: Theory and Evidence', *The Review of Economic Studies*, Vol. 56, 1, pp. 1-19 (January).

[2] Allingham, M. G. and A. Sandmo (1972), 'Income Tax Evasion: a Theoretical Analysis', *Journal of Public Economics*, Vol. 1, 3-4, pp. 323-338 (November).

[3] Alm, J., R. Bahl and M. N. Murray (1993), 'Audit Selection and Income Tax Underreporting in the Tax Compliance Game', *Journal of Development Economics*, Vol. 42, 1, pp. 1-33 (October).

[4] Alm, J., I. Sanchez and A. De Juan (1995), 'Economic and Noneconomic Factors in tax Compliance', *KYKLOS*, Vol. 48, 1, pp. 3-18.

[5] Al-Nowaihi, A. and D. Pyle (2000), 'Income Tax Evasion: A Theoretical Analysis' in MacDonald and Pyle (Ed.), Ashgate Dartmouth.

[6] Altig, D., A. Auerbach, L. J. Kotlikoff, K. A. Smetters and J. Walisser (2001), 'Simulating Fundamental Tax Reform in the United States', *The American Economic Review*, Vol. 91, 3, pp. 574-595 (June).

[7] Ambler, S. (2000), 'Optimal Time Consistent Fiscal Policy with Overlapping Generations', *Working Paper*, pp. 1-23, CREFÉ, Université du Québec à Montréal (May).

[8] Andreoni, J., B. Erard and J. Feinstein (1998), 'Tax Compliance', *Journal of Economic Literature*, Vol. 36, 2, pp. 818-860 (June).

[9] Annicchiarico, B. and N. Giammarioli (2004), 'Fiscal Rules and Sustainability of Public Finances in an Endogenous Growth Model', Working Paper Series, 381, pp. 1-46, European Central Bank (August).

[10] Atolia, M. (2003), 'An OLG Model of Tax Evasion with Public Capital', *Manuscript*, pp. 1-33, Florida State University (April).

[11] Attanasio, O. (1999), 'Consumption', In John B. Taylor and Michael Woodford, eds., *Handbook of Macroeconomics*, Vol. 1B, pp. 741-812.

[12] Australian Chamber of Commerce and Industry (2004), 'Taxation Reform Blueprint: A Strategy for the Australian Taxation System 2004-2014', pp. 1-103 (November).

[13] Barro, R. and Sahasakul C. (1986), 'Average Marginal Tax Rates from Social Security and Individual Income Tax', *The Journal of Business*, Vol. 59, 4(1), pp. 555-566 (October).

[14] Barro, R. and X. Sala-i-Martin (2004), *Economic Growth*, The MIT Press, 2^{nd} Edition.

[15] Becker, G. S. (1968), 'Crime and Punishment: an Economic Approach', *The Journal of Political Economy*, Vol. 76, 2, pp. 169-217 (March-April).

[16] Blumenthal, M. and J. Slemrod (1995), 'Recent Tax Compliance Costs Research in the U.S.', in Sandford (Ed.), Bath.

[17] Bureau of Labor Statistics (2002), 'Employer Costs for Employee Compensation Historical Listings (Annual), 1986-2001', *ECI Standard Errors*, pp. 1-213 (June).

[18] Busato, F. and B. Chiarini (2004), 'Market and Underground Activities in a Two-Sector Dynamic Equilibrium Model', *Economic Theory*, Vol. 23, 4, pp. 831-861 (May).

[19] Chen, B-L. (2003), 'Tax Evasion in a Model of Endogenous Growth', *Review of Economic Dynamics*, Vol. 6, 2, pp. 381-403 (April).

[20] Chiang, A. C. (1992), *Elements of Dynamic Optimization*, McGraw-Hill, Inc., New York.

[21] Christiansen, V. (1980), 'Two Comments on Tax Evasion', *Journal of Public Economics*, Vol. 13, 3, pp. 389-393 (June).

[22] Clotfelter C. (1983), 'Tax Evasion and Tax Rates: an Analysis of Individual Returns', *The Review of Economics and Statistics*, Vol. 65, 3, pp. 363-373 (August).

[23] Cowell, F. A. (1985), 'Tax Evasion with Labor Income', *Journal of Public Economics*, Vol. 26, 1, pp. 19-34 (February).

[24] Cowell, F. A. and J. P. F. Gordon (1988), 'Unwillingness to Pay: Tax Evasion and Public Goods Provision', *Journal of Public Economics*, Vol. 36, 3, pp. 305-321 (August).

[25] Cremer, H. and F. Gahvari (1994), 'Tax Evasion, Concealment and Optimal Linear Income Tax', *Scandinavian Journal of Economics*, Vol. 96, 2, pp. 219-239.

[26] Cullis, J. and P. Jones (1998), *Public Finance and Public Choice,* Oxford University Press, 2^{nd} Edition.

[27] Das-Gupta, A. (2004), 'The Economic Theory of Tax Compliance with Special Reference to Tax Compliance Costs', *Working Paper*, pp. 3-27, Goa Institute of Management.

[28] De La Croix, D. and P. Michel (2002), *A Theory of Economic Growth: Dynamics and Policy in Overlapping Generations*, Cambridge University Press.

[29] Dhami, S. and A. Al-Nowaihi (2006), 'Why Do People Pay Taxes? Prospect Theory Versus Expected Utility Theory', Working Paper 05/23, pp. 1-29, University of Leicester (August).

[30] Diamond, P. A. (1965), 'National Debt in Neoclassical Growth Model', *The American Economic Review*, Vol. 55, 5, pp. 1126-1150 (December).

[31] Dixit, A. (1993), *The Art of Smooth Pasting,* Harwood Academic Publishers.

[32] Dixit, A. and R. S. Pindyck (1994), *Investment Under Uncertainty,* Princeton University Press, Princeton, NJ.

[33] Easterly W. and Rebelo S. (1992), 'Marginal Income Tax Rates and Economic Growth in Developing Countries', *Working Paper Series* 1050, pp. 2-23 (November).

[34] Engel, E. and J. R. Hines (1998), '*Understanding Tax Evasion Dynamics*', Serie Economia 47, Centro de Economia Aplicada, Universidad de Chile, pp. 1-68 (December).

[35] Eichhorn, C. (2004), 'Implications of Tax Evasion for Economic Growth', *Working Paper*, pp. 1-19, Ludwig-Maximilians University of Munich (July).

[36] Feinstein, J. S. (1991), 'An Econometric Analysis of Income Tax Evasion and its Detection', *RAND Journal of Economics*, Vol. 22, 1, pp. 14-35 (Spring).

[37] Feldstein, M. (1985), 'The Optimal Level of Social Security Benefits', *The Quarterly Journal of Economics*, Vol. 100, 2, pp. 303-320 (May).

[38] Friedland, N., S. Maital and A. Rutenberg (1978), 'A Simulations Study of Income Tax Evasion', *Journal of Public Economics*, Vol. 10, 1, pp. 107-116 (August).

[39] Galasso, V. (1999), 'The U.S. Social Security System: What Does Political Sustainability Imply?', *Review of Economic Dynamics*, 2, pp. 698-730 (July).

[40] Geeroms, H. and H. Wilmots (1985), 'An Empirical Model of Tax Evasion and Tax Avoidance', *Public Finance/Finances Publiques*, Vol. 40, 2, pp. 190-209.

[41] Gordon, J. (1989), 'Individual Morality and Reputation Costs as Deterrents to Tax Evasion', *European Economic Review*, Vol. 33, 4, pp. 797-805 (April).

[42] Gupta, R. (2004), 'Endogenous Tax Evasion and Reserve Requirements: a Comparative Study in the Context of European Economies', *Working Paper*, pp. 1-30, University of Connecticut (November).

[43] İmrohoroğlu, A., S. İmrohoroğlu and D. Jones (2003), 'Time-inconsistent Preferences and Social Security', *The Quarterly Journal of Economics*, Vol. 18, 2, pp. 745-784 (May).

[44] Internal Revenue Service: The United States Department of the Treasury (2006), *IRS Updates Tax Gap Estimates*, IR-2006-28 (February).

[45] Jorgenson D. W. and K. J. Stiroh (2000), 'Raising the Speed Limit: U.S. Economic Growth in the Information Age', *Brookings Papers on Economic Activity*, 1:2000, pp. 125-235.

[46] Kakwani N. (1976), 'Measurement of Tax Progressivity: an International Comparison', *The Economic Journal*, Vol. 87, pp.71-80 (March).

[47] Kamien, M. I. and N. L. Schwartz (1981), *Dynamic Optimization: The Calculus of Variations and Optimal Control in Economics and Management*, North Holland.

[48] Kolm, S.-C. (1973), 'A Note on Optimum Tax Evasion', *Journal of Public Economics*, Vol. 2, 3, pp. 265-270 (July).

[49] Léonard, D. and N.-V. Long (1992), *Optimal Control Theory and Static Optimization in Economics*, Cambridge University Press.

[50] Lin, W.-Z. and C. C. Yang (2001), 'A Dynamic Portfolio Choice Model of Tax Evasion: Comparative Statics of Tax Rates and Its Implication for Economic

Growth', *Journal of Economic Dynamics and Control*, Vol. 25, 11, pp. 1827-1840 (November).

[51] MacDonald Z. and D. Pyle (2000), *Illicit Activity: the Economics of Crime, Drugs and Tax Fraud*, Ashgate Dartmouth.

[52] Maddison A. (1991), *Dynamic Forces in Capitalistic Development*, New York: Oxford University Press.

[53] Maddison, A. (1992), 'A Long-run Perspective on Saving', *The Scandinavian Journal of Economics*, Vol. 94, 2, pp. 181-196 (June).

[54] McCaffery, E. J. and J. Slemrod (2004), 'Toward an Agenda of Behavioral Public Finance', *Working Paper* No. 36, pp. 1-28 (August).

[55] McGrattan, R. and Prescott E. (2003), 'Average Debt and Equity Returns: Puzzling?', *The American Economic Review*, Vol. 93, 2, pp. 392-397 (May).

[56] Manasan, R. G. (1988), 'Tax Evasion in the Philippines: 1981-1985', *Journal of Philippine Development*, No. 27, Vol. XV, 2, pp. 167-190.

[57] Niepelt, D. (2005), 'Timing Tax Evasion', *Journal of Public Economics*, Vol. 89, 9-10, pp. 1611-1637 (September).

[58] Ogaki, M., J. D. Ostry and C. M. Reinhart (1996), 'Saving Behavior in Low-and-Middle Income Developing Countries: A Comparison', *IMF Staff Papers*, Vol. 43, 1, pp. 38-71 (March).

[59] Organization for Economic Cooperation and Development (2004), *Tax Administration in OECD Countries: Comparative Information Series (2004)*, Centre for Tax Policy and Administration, pp. 1-70 (October).

[60] Pencavel, J. H. (1979), 'A Note on Income Tax Evasion, Labor Supply and Nonlinear Income Tax Schedule', *Journal of Pubic Economics*, Vol. 12, 1, pp.115-124 (August).

[61] Piketty, T. and E. Saez (2003), 'Income Inequality in the United States, 1913-1998', *The Quarterly Journal of Economics*, Vol. CXVIII, 1, pp. 1-39 (February).

[62] Pontryagin, L. (1998), *Zhizneopisanie*, Moscow.

[63] Pudney, S., D. Pyle and T. Saruc (2000), 'Income Tax Evasion: An Experimental Approach' in MacDonald and Pyle (Ed.), Ashgate Dartmouth.

[64] Ríos-Rull, J-V (1996), 'Life-Cycle Economies and Aggregate Fluctuations', *The Review of Economic Studies*, Vol. 63, 3, pp. 465-489 (July).

[65] Romer, D. (2001), *Advanced Macroeconomics*, McGraw-Hill Higher Education, 2^{nd} Edition.

[66] Russel, A. M., J. A. Rickard and T. D. Howroyd (1982), 'A Tax Evasion Model with Allowance for Retroactive Penalties', *Economic Record*, Vol. 58, 163, pp. 386-394 (December).

[67] Sandford, C. (Ed.) (1995), *Tax Compliance Costs Measurement and Policy*, Fiscal Publications, Bath.

[68] Sandmo, A. (2005), 'Tax Evasion: a Retrospective View', *National Tax Journal*, Vol. LVIII, 4, pp. 643-663 (December).

[69] Samuelson, P. A. (1958), 'An Exact Consumption-Loan Model of Interest with or without the Social Contrivance of Money', *The Journal of Political Economy*, Vol. 66, 6, pp. 467-482 (December).

[70] Sengupta, P. (1998), 'Tax Evasion and Intertemporal Choice', *Atlantic Economic Journal*, Vol. 26, 4, pp. 420-430 (December).

[71] Silberberg, E. and S. Wing (2001), *The Structure of Economics: a Mathematical Analysis*, McGraw Hill, 3^{rd} Edition.

[72] Solow, R. M. (1956), 'A Contribution to the Theory of Economic Growth', *Quarterly Journal of Economics*, Vol. 70, 1, pp. 65-94 (February).

[73] Srinivasan, T. M. (1973), 'Tax Evasion: a Model', *Journal of Public Economics*, Vol. 2, 4, pp. 377-389.

[74] Stern, D. (1987), 'The Theory of Optimal Commodity and Income Taxation: An Introduction', in The Theory of Taxation for Developing Countries, in D. Newberry and N. Stern, New York, Oxford University Press, pp. 22-59.

[75] Swan, T. W. (1956), 'Economic Growth and Capital Accumulation', *Economic Record*, Vol. 32, 2, pp. 334-361 (November).

[76] The World Bank Group (2007), *World Development Indicators Online*, coverage: 1960-present.

[77] Toossi, M. (2002), 'A Century of Change: the U.S. Labor Force, 1950-2000', *Monthly Labor Review*, pp. 15-28 (May).

[78] Tresch, R. W. (2002), *Public Finance: A Normative Theory*, Boston: Academic Press, 2^{nd} Edition.

[79] United States General Accounting Office (2001), 'IRS Audit Rates', *Report to the Chairman, Subcommittee on Oversight, Committee on Ways and Means, House of Representatives*, pp. 1-28 (April).

[80] Wagstaff A. and Doorslaer E. (2001), 'What Makes the Personal Income Tax Progressive? A Comparative Analysis for Fifteen OECD Countries', *International Tax and Public Finance*, Vol. 8, pp. 299-315.

[81] Yaniv, G. (1999), 'Tax Compliance and Advance Tax Payments: Prospect Theory Analysis', *National Tax Journal*, Vol. LII, 4, pp.753-764 (December).

[82] Yitzhaki, S. (1974), 'A Note on Income Tax Evasion: Theoretical Analysis', *Journal of Public Economics*, Vol. 3, 2, pp. 201-202 (May).