TAX EVASION DYNAMICS IN THE CZECH REPUBLIC: FIRST EVIDENCE OF AN EVASIONAL KUZNETS CURVE

Jan Hanousek
Filip Palda

CERGE-EI

Charles University
Center for Economic Research and Graduate Education
Academy of Sciences of the Czech Republic
Economics Institute

WORKING PAPER SERIES (ISSN 1211-3298)
Electronic Version

360

Working Paper Series 360
(ISSN 1211-3298)

Tax Evasion Dynamics in the Czech Republic: First Evidence of an Evasional Kuznets Curve

Jan Hanousek
Filip Palda

CERGE-EI
Prague, September 2008

ISBN 978-80-7343-160-0 (Univerzita Karlova. Centrum pro ekonomický výzkum a doktorské studium)
ISBN 978-80-7344-149-4 (Národohospodářský ústav AV ČR, v.v.i.)

Tax Evasion Dynamics in the Czech Republic: First Evidence of an Evasional Kuznets Curve

Jan Hanousek* and Filip Palda**

CERGE-EI   ENAP

Abstract

Using surveys of the Czech Republic taken in 2000, 2002, 2004, and 2006 we measure how the percentage of tax evaders evolved from 1995 until 2006. We find that at first evasion rose, leveled off, and then fell along a quadratic path, suggesting the existence of what we call an evasional Kuznets curve. Our paper is the first to document the existence of an evasional Kuznets curve and to show how it can help improve Markov-chain predictions of tax evasion. We conclude by suggesting that the evasional Kuznets curve may be a subset of a larger trend in evasion for both transitional and developed economies.

Abstrakt

V této studii jsme měřili pomocí výběrových šetření provedených v letech 2000, 2002, 2004 a 2006, jak se vyvíjela šedá ekonomika v České republice v období 1995-2006. Zjistili jsme, že nejprve šedá ekonomika rostla, dosáhla svého maxima a potom začala klesat podle kvadratické křivky, kterou nazýváme Kuznetsovou křivkou šedé ekonomiky. Naše studie je první, která dokumentuje existenci Kuznetsovy křivky šedé ekonomiky a ukazuje, jak tuto skutečnost využít k predikcím šedé ekonomiky využívajících Markovských řetězců. Domníváme se, že Kuznetsova křivky šedé ekonomiky může být podmnožina trendů v chování šedé ekonomiky, které pozorujeme jak v rozvinutých, tak i v tranzitorních ekonomikách.

Keywords: Underground economy, tax evasion, Markov chains, transition, evasional Kuznets curve.

JEL Codes: H26, H43, K42, O17

We thank Libor Dusek, Randy Filer, Stepan Jurajda, and Jan Kmenta for their detailed comments. This research was supported by the Grant Agency of the Czech Republic (402/04/0166), the Phare ACE Programme of the European Union (P97-8119-R) and the Volkswagen Foundation (II/75831). The authors gratefully acknowledge comments from participants at the Czech Economic Association meetings of 2006 (September), and the Belgian Federal Administration Social Security Conference (May 2007).

* Jan Hanousek is Professor at CERGE-EI, a joint workplace of Charles University and the Academy of Sciences of the Czech Republic.

Address: CERGE-EI, P.O. Box 882, Politických vězňů 7, Prague 1, 111 21, Czech Republic

** Filip Palda is Professor at the École Nationale d'Administration Publique, Montreal, Canada and visiting professor at CERGE-EI.

Emails: Jan.Hanousek@cerge-ei.cz, Filip_Palda@enap.ca.

1. Introduction

All tax evasion research can be divided into three parts: measuring the value of evaded taxes, theorizing about and estimating the structural equations that predict the partial equilibrium response of an individual to a change in preferences or incentives, and measuring the social costs of evasion. Of all these, measurement of evasion has, in recent years, become a growth industry. Schneider and Enste (2000) provide a recent overview of efforts to measure the monetary value of the underground economy.

Knowing the *value* of tax evasion is important for those who worry about the excess burden of taxation. A high value of evasion relative to declared income forces a large tax burden on a small economic base. High burdens on small bases produce what Browning (1976) calls the social cost of public funds. Knowing the *number* of evaders and how this number is changing may be of equal importance to makers of tax policy, but studies of the number of evaders are few.

As evaders grow in number their example may tempt non-evaders to join their ranks. Gordon (1989) observed that "individuals are more likely to evade once they are aware of others evading, a phenomenon which does not square with the Allingham-Sandmo model where evasion reflects an independently made portfolio decision." If tax evasion is in part a herd phenomenon, as the research cited above suggests, policy makers should be aware that past a critical level, as Davis et al. (2003) have shown through simulation, the number of evaders may grow explosively. Research interest in the dynamics of the number of evaders has taken on new importance in light of the recent revisionist movement in the tax evasion literature, exemplified by Frey and Jegen (2001) and surveyed by McCaffrey and Slemrod (2006), which emphasizes sociological reasons for the individual's decision to evade. Knowing the number of evaders may also be of interest simply because the number of evaders is a rough proxy for the value of underground activity. We are interested in measuring the number of evaders because we want to sketch the dynamics of evasion in a transition country. Understanding the dynamics of evasion is crucial for predicting future evasion.

There is nothing new about the study of individual level tax evasion data. Clotfelter (1983) and Slemrod (1985) were among the first to analyze individual level tax data, but they confined their efforts to the search for factors that influence evasion. They had no interest in estimating either the number of evaders or the dynamics of evasion. An authoritative empirical study of the individual's choice to participate in the underground economy is that of Lemieux et al. (1994), but as with Clotfelter and Slemrod, they ignored the dynamics of evasion. The dynamics of individual tax evasion has been largely a theoretical pursuit, as exemplified by the works of Niepelt (2003), Davis et al. Busato et al. (2005). Engel and Hines (1999) draw conclusions about individual level evasion dynamics, but do so from aggregate data they feed into their model of individual evasion. Ours is the first study to use individual data to track aggregate dynamics.

With surveys we conducted in 2000, 2002, 2004, 2006 we track the transitions individuals make between evading and paying taxes. These surveys asked individuals whether they were presently evading and whether they evaded taxes two and five years in the past (section 4.2 explains how we asked these questions in a non-threatening manner). The data can help answer three questions:

1) *Between 1995 and 2006 how did the percentage of evaders change?* We find that within a 95% confidence interval the *number* and percentage of evaders rose until the early millennium and then started to fall. This rise and fall can also be gleaned from macroeconomic estimates of the *value* of the shadow economy in the Czech Republic, provided by Schneider et al. (2004), and Schneider (2005, 2006, 2007). We call this inverse-U path an evasional Kuznets curve and explain how knowledge of this curve is vital to predicting aggregate tax evasion. Simon Kuznets did not research tax evasion, but we invoke his name because through long use it has become associated with the rise and fall of undesirable by-products of economic development such as income inequality, and pollution. We avoid referencing formally the term "Kuznets Curve" because Kuznets never used it explicitly and it is not clear who introduced the term into economics.

2) *Is there a ratchet effect for evaders?* In labor economics a ratchet effect is called hysteresis. We lack the extended time-series necessary for a formal test, but we present some evidence that hysteresis may not be a feature of evasion in a transition economy.

3) *How will evasion evolve?* Using estimates of Markov transition probabilities we suggest that the number of evaders will flatten or fall in the decade to come. No Markov method will be able to predict the entire path of evasion in the presence of an evasional Kuznets, because Markov methods rely on linear matrix operators in their prediction formulas whereas the evasional Kuznets curve follows a quadratic path. Markov methods can be trusted only if we know on which side of the evasional Kuznets curves the economy lies.

The above three questions grew out of our desire to make practical use of tax evasion dynamics. The government that can predict whether evasion will rise or fall is a government alerted to the need to fight tax evasion. Our investigation of evasion dynamics suggests a method for predicting evasion. Two elements are needed for prediction: a mechanistic approach, such as that provided by Markov chain methods, and a more global approach which is the recognition of the side of the evasional Kuznets curve on which an economy lies. Prediction relies on a macro view (evasional Kuznets curve) allied to a micro method (Markov prediction).

Our paper is the first to note the existence of an evasional Kuznets curve. While we study a transition economy, our methods apply to tax evasion in developed economies. We are the first to analyze the dynamics of tax evasion using survey data, whether these data for transition economies or developed economies. In aid of this analysis we have developed a statistical methodology for merging seemingly unrelated surveys. The present paper does not simply contribute an interesting "stylized fact" about evasion, but identifies the questions that set apart the analysis of evasion dynamics from the analysis of dynamics in other fields such as labor studies.

Part 2 of the present paper describes the dataset on which we base our analysis and shows how tax evasion evolved in the Czech Republic over the last ten years. Part 3 presents evasion dynamics in the Czech Republic. Part 4 justifies using non-panel data to analyze evasion dynamics. Part 5 explains how to predict the number of people who will evade, based on our survey data of transitions between the state of evading and not evading. Part 6 explains the theoretical basis of the evasional Kuznets curve.

2. Data

Tax evasion is a broad term. Mirus and Smith (1997) provide a matrix covering conceivable notions of evasion but it is too complex to explain to respondents. We limited ourselves to asking respondents whether they worked without declaring the income from their labors.

Our data come from four face-to-face surveys made up of random stratified samples of residents of the Czech Republic. The surveys were done at our request in 2000, 2002, 2004, and 2006 by the firm MEDIA on samples ranging in size from 1041 to 1066 respondents.

Almost all respondents were Czechs or naturalized Slovaks, all with an excellent command of Czech. Our surveys are similar to those of Lemieux at al. (1994) and Fortin et al. (2000). Their interviews (in their case as well as in ours, face-to-face interviews) gathered information about how much tax people evade and why they evade. Their surveys differed from ours in that they did not ask questions pertaining to the dynamics of tax evasion. Fortin et al. (2000) were interested in the link between buying goods and services on which taxes were not declared, and evasion. Lemieux et al. (1994) were interested in the structural determinants of the decision to participate in the underground economy.

Table 1 cross-tabulates the decision to evade or not evade with some characteristics of Czech respondents but is much less detailed than the tables provided in Hanousek and Palda (2003). We present a limited portrait of evaders to show that our results are in line with the works of prominent researchers in the field of the survey analysis of evasion.

[Table 1]

Table 1 corresponds to the summary statistics of Lemieux et al.'s (1994) survey of tax evasion in Quebec City. Labor on which taxes are avoided tends either to rise and then fall with age or simply to fall. Our response rates for whether or not a respondent had worked in the underground economy in the previous year were close to 100% for all four surveys. In concordance with Lemieux et al. (1994) we also find (but do not show) that reported purchases of untaxed goods exceed reported work in the underground sector.

Our surveys differ from that of Lemieux et al. (1994) in that our Czech respondents admitted to rates of evasion between two and five times larger than those admitted in Quebec. The high number admitting they had evaded is not surprising in light of Schneider and Enste's (2000) survey article on evasion, showing much higher values of tax evasion as a percentage of GDP in transition economies than in developed economies. Our result may also be due to the manner in which we framed our questions about evasion. The obvious problem when asking people about their participation in the underground economy is that they will be reluctant to confess their participation. To avoid this problem we called our surveys "Satisfaction with government services". We started the surveys asking our respondents general demographic questions and questions related to government efforts to battle corruption and what they thought of the quality of services provided by government. When answering these questions respondents had no idea that questions about tax evasion would follow and they were probably put at ease by thinking that our interest in evasion motivated a minor part of the surveys. Flexman (1997) is the pioneering survey pointing the way to asking such questions.

3. Dynamics

Labor economic analysis of dynamics usually begins with time trends of variables of interest, be they labor force participation, or welfare dependency. In the present context of tax evasion the variable of interest is how many people evade taxes in different years. Table 2 uses contemporary as well as retrospective answers from our surveys on evasion to show the rates of evasion and their 95% confidence intervals (formulas for these are in part A of the Appendix) for the 2000, 2002, 2004, and 2006 surveys of the Czech Republic. The column labeled 2000 survey shows rates of evasion based on respondents' retrospective answers concerning 1995 and 1999 and their present answer concerning 2000. Other columns can be similarly read.

[Table 2]

To better grasp Table 2 we have mapped the confidence intervals it contains into Figure 1.

[Figure 1]

Some of the bars in Figure 1 are for different surveys but for overlapping years. The overlap accounts for the repeated values of certain years on the time axis. For example, the figure shows evasion for 2000 drawn from our 2000 survey, and evasion for 2000 drawn from our 2002 survey. There are also overlaps for 2002 and 2004, as can be gleaned from Table 3, from which Figure 1 derives. Overlaps appear in Figure 1 as bars encircled by ovals.

The pattern to emerge from Figure 1 is that evasion rose throughout the 1990's and leveled off since the new millennium. Chi-square tests indicate that between 1995 and 1997 tax evasion, as measured by the percentage of evaders, was rising but that after this evasion continued to fall into 2006.

The dynamics of the *rates* of evasion revealed by the surveys are similar to the dynamics of the *value* of evasion to be found in the macroeconomic estimates of evasion provided by Schneider (2005, 2006, 2007) and Schneider and Klinglmair (2004). These estimates appear in Figure 2. Figure 2 indicates that macroeconomic estimates of the value of evasion rise steeply for the 1990's and decline after the new millennium.

[Figure 2]

Both microeconomic data (Figure 1) and macroeconomic data (Figure 2) suggest that the Czech Republic may have turned the peak of what might be called an "evasional Kuznets curve."

The evasion literature does not discuss an evasional Kuznets curve. How such a curve might arise and what forces compose it are points that need to be clarified but we leave a detailed discussion for the last. For the moment we simply note that we have invoked Kuznets' name because it has become standard to associate with his name a rising and then falling undesirable feature of economic development.

3.1 Evasion by age

With most kinds of survey data it is possible to get an idea of how time influences a variable by looking at the value of this variable for different age groups. Actuaries concerned with pension plans formulate their future payouts based on current demographic profiles. Future demographics are based on current demographic snapshots. An example comes from Constantatos and West (1991) who use a snapshot of age-earnings profiles to calculate the future benefits of investment in education.

Figure 3 shows evasion rates for each age group between 18 and 65 for each of our four surveys. These are not individual data as we have been discussing to this point, but averages taken for each age group. For example, the first row of Figure 3 is drawn from the 2000 survey. The leftmost cell of the first row of Figure 3 maps the age of respondents against the average calculated from each age group of their answer to the question of whether they evaded in 1995. Each point on this graph is the evasion rate for an *age group*, not for an *individual*. We calculated this average by isolating from the survey all persons of a certain age and calculating the rate of evasion for that age. The graphs in Figure 3 are not a time-series of evasion, but rather snapshots of evasion for a particular year, across age groups.

[Figure 3]

For all our surveys, answers to evasion in the present and two years past show a clear downward tendency (last two columns of figures) as established by the Wilcoxon (1985) rank-sum test across ordered groups (using the standard Stata 9 ‘nptrend’ command). For answers to evasion five years past, non-parametric tests confirm no downward tendency. Clotfelter (1983), Pommerehne and Weck-Hannemann (1996), and Orviska and Hudson (2003), have found that the money value of evasion diminishes with age. The tendency of evasion displayed in the last two columns of Figure 3 accords with these studies. The rise and fall of rates of evasion with age in the first column of Figure 3 may be random. The further back one is asked to remember one’s evasion the less well one may remember how one behaved.

Earlier we suggested the Czech Republic might have passed the turning point of its evasional Kuznets curve. The age-evasion profiles in Figure 3 lend further support to this suggestion. As the Czech population ages, *ceteris paribus*, the Czech Republic may continue down the eastern slope of its evasional Kuznets curve.

3.2 Multivariate analysis

So far this paper has focused on averages and cross-tabulations. Researchers have used surveys to analyze why people evade. We do not make the reasons for evasion the subject of this paper but we can use multivariate analysis to help us understand some aspects of evasion dynamics. Labor economists interested in dynamics, such as Blanchard and

Summers (1986), use multivariate analysis to determine where there is inertia in some variable of interest, such as unemployment. We follow their lead to ask whether there is inertia in tax evasion. We conducted regressions and logits (not shown here) of evasion in the current year on retrospective evasion, and a list of variables such as age, sex, income, education and other variables it is common to include in a multivariate examination of the causes of tax evasion. Our goal was to keep to a minimum the number of potentially endogenous variables on the right hand side of the equation so as to examine the coefficients that are most relevant to our discussion of evasion dynamics. The coefficients of greatest interest were those attached to retrospective answers.

Age had a negative but insignificant coefficient. This is due in part to the high correlation between age and retrospective answers on evasion. The two coefficients of greatest interest were the coefficients attached to the retrospective answers on evasion. These coefficients showed that evasion in the near past has a much stronger positive influence (ten times as much using point estimates) on current evasion than evasion five years ago. Results were robust to inclusion of different variables and runs conducted using different methods such as regressions and logits.

We have not formally tested for the absence of hysteresis. A formal test requires a time-series of panel data much longer than ours as Im et al. (1997) have shown, but the above results are interesting because they hint that hysteresis may not be a problem in tax evasion. Hysteresis is best known as a hypothesis from labor economics that holds that past increases in unemployment can permanently shift upward the natural rate of unemployment. In the present context, hysteresis would hold that past evasion does not have a decaying effect on current evasion. The very small value of the coefficient we found on evasion five years ago compared to evasion two years ago suggests that hysteresis of tax evasion may not be a problem Czechs need fear.

4. Using non-panel data to analyze dynamics: consistency of surveys

4.1. Survey consistency test

Juxtaposing surveys that are not panel data gives rise to the question of whether such juxtaposition has meaning. One of the main variables of interest in the surveys is the individual's answer to whether or not he or she evaded taxes. Each survey asked people about their current and past evasion. If answers about evasion in 2002 given to questions in the 2004 survey are statistically indistinguishable from answers about evasion in 2002 given by respondents in the 2002 survey one might conclude that memory is good and that surveys in 2002 and 2004 are consistent with each other. Consistency means that answers in surveys conducted at different times do not contradict each other.

To test the consistency of present answers about evasion with retrospective answers for the same year from a later survey, consider two independent sample surveys of n and m observations respectively $\mathbf{x_1}=(x_{11}, x_{12}, \ldots, x_{1n})$ and $\mathbf{x_2}=(x_{21}, x_{22}, \ldots, x_{2m})$, where x_{ij} denotes the j[th] observation of the i[th] survey. Survey 1 is taken in the year 2000 and survey 2 is taken in the year 2002. The x's in the 2000 survey are the answers of each respondent to whether he or she evaded in 2000 and the x's in the 2002 survey are the answers to whether a respondent remembered evading in 2000. " Yes" answers are coded as ones, no answers as zeroes. The data are non-panel. Our variables of interest are the proportions of evaders in each sample $p_1 = \frac{1}{n}\sum_{i=1}^{n} x_{1i}$, and $p_2 = \frac{1}{m}\sum_{i=1}^{m} x_{2i}$ and we wish to test the hypothesis

H$_0$: $p_1=p_2$,

i.e., that the proportion tax evaders in both samples is the same. Consider the following u test statistic

$$u = \frac{p_1 - p_2}{\sqrt{\overline{p}(1-\overline{p})\left(\frac{1}{n}+\frac{1}{m}\right)}} \qquad (1)$$

where $\overline{p} = \frac{1}{(n+m)}\left(np_1 + mp_2\right)$, Under the null hypothesis, the test statistic u has a standard normal distribution in large samples. The above is a test statistic that allows us to distinguish whether certain variables have been drawn from different distributions. Our results are summarized in Table 3.

[Table 3]

Table 3 indicates that:

1) No difference can be found for the 2000 survey estimate of evasion in 2000 and the 2002 survey retrospective estimate of evasion in 2000 because the U-statistic of U = -0.359 is not significant.
2) The same can be said of the 2002 survey estimate of evasion in 2002 and the 2004 survey estimate of evasion in 2002 (U = 0.382, not significant).
3) The same can be said of the 2002 survey estimate of evasion in 1999 and the 2000 survey and its estimate of evasion in 1999 (U = -0.863, not significant).
4) Similarly, the 2006 survey estimate of evasion in 2004 and the 2004 survey and its estimate of evasion in 2004 show a consistent pattern (U = -0.955, not significant).

The apparently strong consistency between surveys gives some justification for building time-series from the surveys, indicates that answers to questions about past evasion in a survey taken in one year are statistically indistinguishable from answers to questions about contemporary evasion given in a survey two years earlier. Even though the surveys are independently drawn, we are tempted to say that *people remember*.

4.2. Credibility and usefulness of the answers

Individuals lie in surveys. The percentage of respondents to political surveys who claim they vote in US federal elections is consistently above 80% (Matsusaka and Palda 1999) but the real figure is closer to 55%. Lying about tax evasion is more difficult to measure than lying about voter participation because no satisfactory objective data exists on the rate of evasion. There are no exit polls for tax evasion.

Audits and penalties by tax authorities give biased estimates of evasion because tax auditors do not perform random samples from the population in deciding whom to audit. Without much academic precedent we venture that more survey respondents seek to hide their evasive activities than there are respondents who falsely boast of having evaded. Average rates of evasion calculated from all respondents will underestimate evasion. To test this idea we asked respondents to estimate what percentage of other people in their country evaded taxes. We believe this an unthreatening question to which people will respond honestly. In 2000 respondents believed that on average 38.3% of the population evaded taxes. The average for their answers to the more threatening question of “did you evade in 2000?” is 25.2% (see Table 1). Similar discrepancies arise for the other surveys. If respondents are on average correct in their estimation of how much other people evade then the above comparison suggest that answers to personal questions underestimate the rate of evasion. Our estimates may also underestimate evasion because almost all of our respondents were native Czech speakers. The Czech Republic is home to an unknown number of illegal immigrants working in the underground economy. Omitting these workers from our surveys may bias downwards our estimates of evasion rates.

Hanousek and Palda (2006) explain in their critique of tax evasion measures that lying about evasion, or forces that bias estimates downwards (such as a large illegal immigrant population) may vex those who wish to know the *absolute* rate of evasion. Hanousek and Palda argued that to those interested in evasion *dynamics*, lying might be less of a problem provided the lies are systematic. If 20% of respondents lie about evasion and the same 20% lie equally the following year then taking the difference between evasions in the two years will cancel the lies. We take a similar position in the present paper. Our interest is in dynamics and not in absolute levels. Our estimates of the dynamics of evasion will be biased to the degree that liars and honest respondents differ in their transition from evasion to non-evasion. To this potential problem with our estimates we have no satisfactory answer.

5. Predicting evasion

Dynamic analysis is not just about rummaging through the past, but about looking into the future. Our data can give some idea of how tax evasion will evolve. We can use Markov chains to predict evasion, but Markov chains do not suffice because their perspective is linear. To predict the future we might be helped by knowing on which side of the non-linear evasional Kuznets an economy lies. Prediction must proceed in two steps. First establish a method for predicting the future, and then judge this method by determining on which side you sit of the evasional Kuznets curve.

To see how tax evasion will evolve in the Czech Republic we focus on the probability of changing between states of evasion. A proportion of the new labor force arriving on the market will not evade and others will jump to evading. Those who are not new to the labor force and do not evade can make similar jumps. Those who evade may jump to not evading. The 2x2 "stage/transition matrix" can summarize these flows in and out of tax evasion for each individual for the $t+1$ survey given in Table 4:

[Table 4]

Each cell gives for an individual the probability he or she will go from one state in t to another state in $t+1$. For example, P_{en} gives the probability an individual who evaded in t will never evade in $t+1$.

We assume that the transitions from evader to non-evader (and vice-versa) satisfy the Markov property; that is, the best forecast of future transitions depends on the current behavior of the taxpayer. Stated differently, the Markov assumption says that the path by which one has arrived at one's current state has no influence on the probability of acceding to a future state. Evading in previous states then has no bearing on evading in a subsequent state. To those who believe in learning-by-doing our Markov assumption will seem objectionable. We do not model current transition probabilities as depending on past behavior for practical reasons. To forecast with path dependency means that one must know the contributions of the path and the contributions of individual characteristics that determine jumps from one state to another. We do not have a long-enough time series to isolate both effects nor do we know how this omission in our modeling might bias or not bias our forecasts. Given the novelty of the present research we wish to lay bare what will need to be done in future research and also to state that if the Markov assumption seems a strong one, we have not made this assumption gratuitously.

Each individual's transition probability will differ from the transition probability of other individuals. To precisely estimate how total evasion will evolve we would need to calculate a stage-transition matrix for each individual and then see what "percentage" of that individual moves from cell to cell. We would then add all these percentages in each year to arrive at the total number of evaders in each of the two categories. A simpler, though somewhat less precise way of arriving at the same calculation is to calculate aggregate transition probabilities. Calculating the percentage of people who moved from cell to cell between 1999 and 2000 does this. Aggregate probabilities are slightly less accurate than if we used a transition matrix for each individual, but given the large numbers we surveyed, the central limit theorem suggests that the variance of our calculations around the true mean (provided that individual transition probabilities are uncorrelated with each other) will not be far off their true values.

Our technique for predicting the evolution of tax evasion is new and needs to be considered in the context of past research on evasion prediction. Engel and Hines (1999) have built a model to simulate long-term evasion dynamics in the US using aggregate data. Aggregate evasion shows cycles if a sufficiently large number of individual taxpayers cycle together, as happens under the influence of aggregate shocks, which tend to influence all in the same direction. In the absence of such shocks Engle and Hines find the interesting result that the cross-section of evasion rates converges to a steady state and aggregate tax evasion approaches a limit even though individual rates cycle. The distinction between aggregate and individual cycles arises because an individual's steady state is conditional on not being audited, while the economy's steady state is conditional on a distribution of individual audits across taxpayers with differing evasion histories. The distinction between aggregate and individual cycles in tax evasion is similar to the distinction between family and societal sex ratios.

We can use Engle and Hines' (1999) insight that tax evasion converges to a steady state to draw conclusions about the evolution of tax evasion in the Czech Republic. Engle and Hines used their model to examine continuous aggregate data on tax evasion. Our data is on individuals, is discrete, and spans five years over which we can see how the respondent jumped between the categories of evading and not evading. As far as we know, such a dataset is unique.

Using the formulas described in Appendix B we calculated the short-term transition matrices and present these in Table 5.

[Table 5]

The value of 0.21 in the upper left cell of the top matrix indicates the probability an individual will go from a "state" of evading in 1999 back to the same state in 2002. The sum of values in the cells of each matrix must come to one because the individual must move from one state to another state or the same state with complete certainty.

Table 6 shows what evasion rates were in fact, and compares them to our projections using short-term transition probabilities (formulas for projections are in Appendix C).

[Table 6]

Values on the diagonal (bold faced numbers) have been estimated from the survey (already presented in the Table 1). Outside of the diagonal are predictions that used short-term transition matrices with the Markov property. Their confidence intervals were constructed using a bootstrap procedure (for formulas and technical details see the appendix). To understand how to read this table consider the first column. The top cell of this column indicates that the true evasion rate in 2000 was 25%. The cell below is the prediction that based on the 2000 survey, evasion in 2002 would be 29%. Now jump one cell right of the cell containing the 29%. The bold-faced number **0.24** indicates that actual evasion rates, based on the 2002 survey were 24%. As we work our way down the first column we see that the furthest prediction is also the least accurate. The 2000 survey predicts evasion in 2006 will be 36%. Actual evasion, as read from the last cell of the last column was 22%. More formally, for the shaded cells we reject the hypothesis that predicted values based on the 2000 survey are statistically indistinguishable from the actual values as given on the diagonal.

The inability of the 2000 survey to give credible forecasts of evasion is due to the Kuznets property of evasion, which we documented in Table 2 and Figure 1. Around 2002 it appears that the Czech Republic had gone over the peak of the evasional Kuznets curve. Any survey taken before 2002 will be inappropriate for prediction because Markov predictions are based on linear formulas whereas the evasional Kuznets curve is quadratic. Surveys on or after 2002 predict evasion quite well, as Table 6 shows, because the linear equations on which the predictions are based past 2002 are fair approximations of the downward sloping part of the evasional Kuznets curve.

In calculations not shown here we found that whatever survey we used, predictions about the percent of evaders and non-evaders converge to some constant percentage. Such convergence is similar to Engle and Hines' (1999) discovery of a steady state in evasion. However, our work goes a step further than theirs in that we use transition probabilities estimated from individual data to make our projections. We must of course be sensitive to the criticism that our convergence to a steady state is an artifact of the computational method we use. The essence of this computation lies in the assumption of stable transition probabilities. Our finding of an evasional Kuznets curve suggests that transition probabilities are not constant and that as a result evasion may not be predicted by a mechanical application of linear formulas such as those derived from a Markov model. To predict with some success using linear methods we need to know on which side of the evasional Kuznets curve an economy sits.

To test our view that Markov transition probabilities are not stable we can resort to a Chi-square test, which asks whether the probabilities calculated from each of our four surveys are drawn from the same distribution of variables that underlie the calculation of our probabilities.

For the test of stability (homogeneity) of the Markov transition matrices, both short and long-term, let us consider the four following categories: [E->E] (someone who evaded in an earlier period continues to evade in the current period), [E->N] (someone who evaded in an earlier period no longer evades in the current period), [N->E] (someone who did not evade in an earlier period begins to evade in the current period) and [N->N] (someone who did not evade in an earlier period continues to not evade in the current period). The numbers *n* transiting from one state to another are the basis of the calculation of our transition probabilities. A non-parametric test that the underlying transition matrices are the same across all surveys can be carried out by using a standard test of homogeneity of distributions. Let us summarize all outcomes in the Table 7.

[Table 7]

Under the null hypothesis, the test statistic is

$$\chi^2 = \sum_{i=1}^{4} \sum_{j}^{4} \frac{\left(n_{ij} - \frac{1}{n} n_{i.} n_{.j}\right)^2}{\frac{1}{n} n_{i.} n_{.j}} = n.\sum_{i=1}^{4} \sum_{j}^{4} \frac{{n_{ij}}^2}{n_{i.} n_{.j}} - n \qquad (2)$$

a chi-square distribution with 9 degrees of freedom. The test statistic of 60.45 indicates that jointly the surveys cannot be said to be drawn from the same distribution as far as transition probabilities are concerned.

Is there any contradiction between the non-stability of transition probabilities and our earlier finding that all four surveys were consistent with each other? When we tested for the consistency of surveys we looked at tax evasion for the *same year* measured using current and retrospective answers from *different surveys*. As we expected, we found that we could not statistically distinguish between answers about evasion for a particular year drawn from different surveys. The consistency of those surveys does not mean that there is no change in evasion probabilities or that the Markov probability matrix cannot change from year to year. We expect the probability of evasion and the Markov matrix to change as demographic variables change. Similar tests of stability for variables, which could be linked to change in evasion status, such as household income (compared to one year and five years ago, respectively), and satisfaction with government services, also reject the stability of these variables.

6. Components of the evasional Kuznets curve

The present paper has so far only documented the possible existence of a curve that maps a rising and then falling level of tax evasion against time in the Czech Republic. We have called this relation an evasional Kuznets curve without explaining why it should exist or why it should not provoke skepticism. A critique of our findings is implicit in the work of Enste (2005) who found that evasion was consistently higher in high-tax OECD countries than in low-tax OECD countries. Does Enste's cross-sectional analysis of countries not suggest that as the Czech Republic passed from low taxation in the early 1990's to ever-higher taxation, evasion there should have done nothing but rise?

The answer is "yes" if we believe the only change gripping the Czech economy in this period was the rise of taxes. The answer is closer to "no" if we allow that different forces affected tax evasion and that over time some of these forces worked against each other in a way that produced at first a rising, and then a falling level of evasion. An illustrative example appears in Figure 4.

[Figure 4]

From period one to period ten an index of tax rates *T* rises from one to ten. In this period an index of the perceived quality of government services *Q* also rises from one to ten. Evasion for this example is an additive function of the tax and quality indices taking the general form *E=f(T)+g(Q).* The particular form we give this function for illustrative purposes is

$$E = \underbrace{50T^{\frac{1}{2}}}_{f(T)} - \underbrace{(Q^2 + 49)}_{g(Q)} \qquad (3)$$

Evasion is a rising function of taxes and a falling function of the quality of government services. Figure 4 maps both the *f* and *g* functions and their sum, which gives the rate of evasion over time and resembles an evasional Kuznets curve. The above exercise proves nothing, but suggests that the evolution of evasion over time depends on:

1) The functional dependence of evasion on possibly countervailing forces such as quality of government services, and tax rates.

2) The evolution of quality and tax rates. Had we joined unchanging quality to rising tax rates evasion would have risen steadily over time, as Enste's (2005) work suggests it might. Instead, we joined rising quality with rising tax rates in a way that produced an evasional Kuznets curve.

The above example is simplistic because it deems tax rates and the quality of government services are the only influence on evasion, or that changes in these factors are the only influence on evasion over time. We formulated the example in such a manner as to join the two forces that students of evasion believe to be among the most potent determinants of evasion. We also chose this example because it has some empirical backing. Taxes rose in the Czech Republic after1989. Perceived quality of government services are harder to measure than taxes, but following the methodology of Hanousek and Palda (2004) we found from our four surveys that since 2000 Czechs are increasingly satisfied with government services, and see corruption as declining. Czechs also increasingly believe it is immoral to evade and that family reactions to evasion are becoming increasingly negative. If both taxes and quality were rising and working against each other in their effect on evasion, the sum of their opposite influences might have given rise to an evasional Kuznets curve over the period we studied.

The challenge the above interpretation of a Kuznets curve raises to the theme of the present paper is that trends in countervailing forces such as taxes and quality may change. The right tail of the evasional Kuznets curve may flatten out, or may start to rise again if taxes continue to rise and the quality of government services stagnates. We have proposed a method for predicting evasion using Markov chains, provided one knows on which side of the Kuznets curve one lies. The above example suggests that the evasional Kuznets curve may be a feature particular to transition economies. After a country passes through its transition, the useful horizons of predictions may shrink. The challenge for prediction is knowing whether one is still in transition. To this question we provide no answer as yet. Future efforts might bend themselves to answering this question.

7. Conclusion

The present paper has laid out how the percentage of tax evaders in the Czech Republic has evolved since 1995. We have argued that surveys may be merged to provide a time-series of evasion. Surveys hold two keys to predicting evasion. Surveys may show on which side of an evasional Kuznets curve and economy lies. Once one knows where one lies on such a curve Markov methods may be used to predict how evasion will evolve.

Our work may be of interest to scholars of transition economies but we believe it is also the first attempt to analyze the dynamics of tax evasion using survey data, whether for transition or for developed economies. We have explained how non-panel survey data may be used with confidence to track the evolution of evasion. Our discovery of an evasional Kuznets curve is a first in the tax evasion literature and has not been anticipated by theoretical developments or by previous empirical studies. We have gone beyond simply noting the existence of an evasional Kuznets curve to show how it may be wedded to Markov prediction techniques to foretell trends in tax evasion.

In sum, the present paper has contributed not only an interesting "stylized fact" about tax evasion but has show how this stylized fact can be of practical use to the forecaster.

References

Blanchard, Olivier J. and Lawrence H. Summers (1986). "Hysteresis and the European Unemployment Problem." *NBER Macroeconomic Annual*, 1:15-78.

Browning, Edgar K. (1976). "The Marginal Cost of Public Funds." *Journal of Political Economy*, 84:2, 283-98.

Busato, Francesco, Bruno Chiarini, and Guido M. Reys (2005). "Equilibrium Implications of Fiscal Policy with Tax Evasion." University of Aarhus working paper.

Clotfelter, Charles T. "Tax Evasion and Tax Rate: An Analysis of Individual Returns." *The Review of Economics and Statistics*, 65:363-373.

Constantatos, Christos and Edward G. West (1991). "Measuring Returns From Education: Some Neglected Factors." *Canadian Public Policy*, 17:127-138.

Cuzick, J. (1985). "A Wilcoxon-type test for trend." *Statistics in Medicine*, 4:87-90.

Davis, Jon S., Gary Hecht, and Job D. Perkins (2003). "Social Behaviors, Enforcement, and Tax Compliance Dynamics." *The Accounting Review*, 78:39-69.

Efron, B. *The Jackknife, the Bootstrap and Other Resampling Plans*. Philadelphia: Society for Industrial and Applied Mathematics, 1982.

Efron, B. (1979). "Bootstrap methods: another look at the jackknife." *Annals of Statistics,* 7:1–26.

Efron, B. and R. J. Tibshirani. *An Introduction to the Bootstrap*. New York: Chapman & Hall, 1993.

Engel, Eduardo M. R. A. and Hines, James R., Jr. 1999. "Understanding Tax Evasion Dynamics." Working Paper: 6903, National Bureau of Economic Research, January 1999.

Enste, Dominik H. (2005), "The Shadow Economy in OECD and EU Accession Countries: Empirical Evidence for the Influence of Institutions, Liberalization, Taxation and Regulation." Pages 123-138 in Size, *Causes and Consequences of the Underground Economy: An International Perspective.* Edited by Friedrich Schneider and Christopher Bajada. England: Aldershot, 2005.

Fortin, Bernard; Lacroix, Guy and Montmarquette, Claude (2000). “Are Underground Workers More Likely to be Underground Consumers?” *Economic Journal,* 466:838-60.

Flexman, Bruce. “Canadian Attitudes Towards Taxation.” Pages 53-74 in *The Underground*

Economy: Global Evidence of its Size and Impact. Edited by Owen Lippert and Michael Walker. Vancouver: The Fraser Institute, 1997.

Frey, Bruno S. and Reto Jegen (2001). “Motivation Crowding Theory.” *Journal of Economic Surveys*, 15:589–611.

Gordon, James F. (1989). “Individual Morality and Reputation Deterrents to Tax Evasion.” *European Economic Review*, 33:797-805.

Hanousek, Jan and Filip Palda (2006). “Problems measuring the underground economy in transition.” *The Economics of Transition,* 14:707–718.

-------- (2004). “Quality of Government Services and the Civic Duty to Pay Taxes in the Czech and Slovak Republics, and other Transition Countries.” *Kyklos*, 57: 237–252.

-------- “Why People Evade Taxes in the Czech and Slovak Republic: A Tale of Twins.” Pages 139-174 in *The Informal Economy in the EU Accession Countries: Size, Scope, Trends and Challenges in the Process of EU Enlargement*, edited by Boyan Belev. Sofia: Center for the Study of Democracy, 2003.

Im, K.S., Pesaran, M.H. and Shin, Y. (1997). "Testing for unit roots in heterogeneous panels." Mimeo, DAE University of Cambridge.

Lemieux, Thomas, Bernard Fortin, and Pierre Fréchette (1994). “The Effects of Taxes on Labor Supply in the Underground Economy.” *The American Economic Review,* 84 :231-254.

Matususaka, John G. and Filip Palda (1999). “Voter Turnout: How Much Can We Explain?″ *Public Choice*, 98:431-446.

McCaffrey, Edward J. and Joel Slemrod, *Behavioral Public Finance*. New York: Russel Sage Foundation, 2006.

Mirus, Rolf and Roger S. Smith. “Canada’s Underground Economy: Measurement and Implications.” Pages 3-10 in *The Underground Economy: Global Evidence of its Size and Impact.* Edited by Owen Lippert and Michael Walker. Vancouver : The Fraser Institute, 1997.

Niepelt, Dirk (2003). “Tax Evasion Dynamics.” Institute for International Economic Studies, Stockholm, Seminar Paper No. 721.

Orviska, Marta and John Hudson (2003). "Tax evasion, civic duty and the law abiding citizen." *European Journal of Political Economy*, 19 :83-102.

Pommerehne, Werner W. and Hannelore Weck-Hannemann (1996). "Tax Rates, Tax Administration and Income Tax Evasion in Switzerland." *Public Choice*, 88:161-170.

Schneider, Friedrich (2007). “Shadow Economies and Corruption all over the World: New Estimates for 145 Countries.” University of Linz working paper.

-------- (2006). “Shadow Economies of 145 Countries all over the World: What do we really know?” University of Linz working paper.

-------- (2005). “Shadow Economies Around the World: What do We Really Know?” *European Journal of Political Economy*, 21:598-642.

Schneider, Friedrich and Robert Klinglmair (2004). “Shadow Economies Around the World: What do We Know?” Institute for the Study of Labor (IZA) working paper No. 1043.

Schneider, Friedrich and Dominik H. Enste (2000). “Shadow Economies: Size, Causes, and Consequences.” *Journal of Economic Literature*, 38:77-114.

Slemrod, Joel (1985). “An Empirical Test for Tax Evasion.” *The Review of Economics and Statistic*, 67:232-238.

Appendix:

A. Calculating confidence intervals for estimated probabilities of evasion

The sample relative frequencies *pe, p a*nd p_n allow us to construct confidence intervals for underlying probabilities P_e, and P_n and for a transition matrix Π_{ij} . Since we analyze a random sample from the Czech population, population size, *N*, will refer to several millions, and therefore we can use the well-known normal approximation to show that

$$\frac{p-P}{\sqrt{\frac{p(1-p)}{n-1}\frac{N-n}{N}}} \approx \frac{p-P}{\sqrt{\frac{p(1-p)}{n-1}}} \sim N(0,1)\,, \qquad \text{(A.1)}$$

where *p* and *P* can refer to p_e, and p_n and P_e, and P_n , respectively. Hence, a (1- α)% confidence interval is simply determined by (we use *1/2n* correction for non-continuous random variables)

$$p-\frac{1}{2n}-u_{1-\alpha/2}\sqrt{\frac{p(1-p)}{n-1}} < P < p+\frac{1}{2n}+u_{1-\alpha/2}\sqrt{\frac{p(1-p)}{n-1}}\,, \qquad \text{(A.2)}$$

where *u* denotes the quintile of the standard normal distribution.

B. Estimating the probability of evasion and transition probability matrices

Let $$T = T_t = \begin{bmatrix} T_{EE} & T_{EN} \\ T_{NE} & T_{NN} \end{bmatrix} \qquad \text{(A.3)}$$

denote the transition probability matrix between evading and non-evading states. Each cell holds for an individual the probability he or she will go from one state in period *(t-1)* to another state in period *t*. For example, T_{EN} gives the probability an individual who evaded in *(t-1)* will not evade in the period *t*, etc.

Similarly, $$E = E_t = \begin{bmatrix} E_E \\ E_N \end{bmatrix} \qquad \text{(A.4)}$$

is a vector containing probabilities of an individual evading (E_E) and non-evading (E_N) at time *t*.

Using individual responses to the set of the retrospective questions, we can construct the following set of dummy variables:

$$e_t = e_{t,i} = \left\langle \begin{array}{l} 1 \text{ if individual i evaded } in \text{ period t} \\ 0 \text{ otherwise} \end{array} \right. \tag{A.5}$$

$$t_{EE,t} = t^i_{EE,t} = \left\langle \begin{array}{l} 1 \; if \;\; e_{t,i} = 1 \;\&\; e_{t-1,i} = 1 \\ 0 \text{ otherwise} \end{array} \right. \tag{A.6}$$

$$t_{EN,t} = t^i_{EN,t} = \left\langle \begin{array}{l} 1 \; if \;\; e_{t,i} = 1 \;\&\; e_{t-1,i} = 0 \\ 0 \text{ otherwise} \end{array} \right. \tag{A.7}$$

$$t_{NE,t} = t^i_{NE,t} = \left\langle \begin{array}{l} 1 \; if \;\; e_{t,i} = 0 \;\&\; e_{t-1,i} = 1 \\ 0 \text{ otherwise} \end{array} \right. \tag{A.8}$$

$$t_{NN,t} = t^i_{NN,t} = \left\langle \begin{array}{l} 1 \; if \;\; e_{t,i} = 0 \;\&\; e_{t-1,i} = 0 \\ 0 \text{ otherwise} \end{array} \right. . \tag{A.9}$$

Realizations of the random variables defined above in (3)-(7) form sample counterparts to the probabilities of evasion and the transition probability matrix, respectively.

Therefore, $$\hat{E}_E = \hat{E}_{E,t} = \frac{1}{n}\sum_{i=1}^{n} e_{t,i} \text{ and } \hat{E}_N = \hat{E}_{N,t} = 1 - \hat{E}_{E,t} \tag{A.10}$$

and $$\hat{T}_{EE} = \hat{T}_{EE,t} = \frac{1}{n}\sum_{i=1}^{n} t^i_{EE,t} \,, \qquad \hat{T}_{EN} = \hat{T}_{EN,t} = \frac{1}{n}\sum_{i=1}^{n} t^i_{EN,t}$$

$$\hat{T}_{NE} = \hat{T}_{NE,t} = \frac{1}{n}\sum_{i=1}^{n} t^i_{NE,t} \,, \quad \text{and} \qquad \hat{T}_{NN} = \hat{T}_{NN,t} = \frac{1}{n}\sum_{i=1}^{n} t^i_{NN,t} \,. \tag{A.11}$$

Since all variables defined in (3)-(7) are sample realizations of Bernoulli (0-1) variables, their estimated sample variance is equal to

$$est.\operatorname{var}\left(\hat{\Theta}\right) = \frac{1}{n}\hat{\Theta}(1 - \hat{\Theta}) \,, \tag{A.12}$$

for all estimators defined in (A.10)-(A.11).

C. Prediction of future evasion using current evasion and transition probability matrices

We assume that we know (at time *t*) the probability of evasion and the past transition probability matrix. Using the Markov-type computation, we can construct the predicted probability of evasion as:

$$\hat{E}_E = \hat{E}_{E,t+1} = E_{E,t}\, P(evading(t+1)\,|\,evading(t)) + E_{N,t} P(evading(t+1)\,|\,non-evading(t)) =$$
$$= E_{E,t} \frac{T_{EE}}{T_{EE}+T_{EN}} + E_{N,t} \frac{T_{NE}}{T_{NE}+T_{NN}} \qquad \text{(A.13)}$$

The above says that the probability of evading at the time *t+1* is equal to probability of evading at time *t* times the probability that those evading at the time *t* will be still evading at the time *t+1* plus the probability of non-evading, times the probability that those not evading at the time *t* will start evading at *t+1*.

Similarly

$$\hat{E}_N = \hat{E}_{N,t+1} = E_{E,t} \frac{T_{EE}}{T_{EE}+T_{EN}} + E_{N,t} \frac{T_{NE}}{T_{NE}+T_{NN}} = 1 - \hat{E}_{E,t+1}\,. \qquad \text{(A.14)}$$

Point estimates of the predicted probabilities of evasion can be easily constructed from (A.12) and (A.13). Because of the non-linear relationship and possible interdependence between estimates of $\hat{T}$ and $\hat{E}$, estimation of the variances of (A.12) and (A.13) is not straightforward. One could try to employ a delta method to get an estimated asymptotic variance of the predicted probability of evasion at time *(t+1)*, however, one would still need to compute/estimate the covariance between T and E, which together with the first derivatives will lead to a complicated formula. In addition, computing the variance via delta methods relies on a certain set of assumption and more importantly, it gives the asymptotic behavior of the variance. Let us note that its finite sample properties could be rather different. Therefore, we opted for a simpler method that uses the bootstrap method (for the original setup see Efron , 1979 and 1982; Efron and Tibshirani, 1993).

Table 1: Descriptive statistics for tax evasion in the Czech Republic for the 2000, 2002, 2004, 2006 surveys

	% of evading individuals by category			
survey	*2000*	*2002*	*2004*	*2006*
Total	**25.2%**	**23.9%**	**21.4%**	**22.0%**
Sex: Male	34.6%	29.8%	29.9%	29.5%
Female	16.4%	17.8%	12.6%	14.5%
Age: 18 to 25 years	30.3%	29.6%	25.1%	29.0%
26 to 35 years	25.9%	26.1%	23.3%	21.4%
36 to 45 years	31.4%	26.6%	16.5%	27.5%
46 to 55 years	22.2%	18.6%	23.0%	18.9%
56 to 65 years	13.3%	16.3%	18.2%	12.5%
Education: Primary	29.0%	25.1%	28.3%	29.5%
Without GCE	32.6%	26.0%	24.5%	26.5%
With GCE	14.6%	23.3%	16.4%	16.1%
Higher	14.8%	13.6%	13.9%	12.4%
Income[CZK]: < 10,000	22.1%	19.5%	17.6%	19.2%
10,001 to 15,000	33.1%	24.5%	23.9%	25.7%
15,001 to 20,000	23.1%	30.7%	23.7%	20.5%
20,001 to 25,000	50.0%	42.9%	39.0%	29.6%
25,001 to 30,000	55.6%	50.0%	28.6%	20.8%
30,001 to 40,000	100.0%	25.0%	0.0%	27.3%
Rejected	15.9%	20.5%	15.8%	8.5%

Source: 2000-2006 survey's, authors' computations

Table 2: Tax evasion rates and confidence intervals for the 2000, 2002, 2004, and 2006 surveys

Year	2000 survey	2002 survey	2004 survey	2006 survey
1995	15.4% (13.3%, 17.6%)	NA	NA	NA
1997	NA	23.1% (20.5%, 25.7%)	NA	NA
1999	20.6% (18.2%, 23.1%)	NA	22.2% (19.7%, 24.7%)	NA
2000	25.2% (22.6, 27.9%)	25.9% (23.2%, 28.6%)	NA	NA
2001	NA	NA	NA	21.2% (18.7%, 23.8%)
2002	NA	23.9% (21.3%, 26.5%)	23.2% (20.6%, 25.7%)	NA
2004	NA	NA	21.4% (18.9%, 23.8%)	23.4% (20.8%, 26.1%)
2006	NA	NA	NA	22.0% (19.4%, 24.5%)

Source: Our 2000, 2002, 2004, 2006 surveys of tax evasion in the Czech Republic. NA indicates "not applicable". The first lines contain the mean of each category expressed in percents, the second lines give estimated 95% confidence interval.

Table 3. Tests of consistency of surveys: comparison of retrospective estimates of evasion.

A. Tax evasion in 2000 (test of consistency 2000&2002)

Survey	Evaders	Non-evaders	Total
2000	268	794	1062
2002	268	766	1034
Test statistics	**-0.359**	p-value:	**0.360**

B. Tax evasion in 2002 (test of consistency 2002&2004)

Survey	Evaders	Non-evaders	Total
2002	247	788	1035
2004	245	813	1058
Test statistics	**0.382**	p-value:	**0.649**

C. Tax evasion in 1999 (test of consistency 2000&2004)

Survey	Evaders	Non-evaders	Total
2000	219	843	1062
2004	234	822	1056
Test statistics	**-0.863**	p-value:	**0.194**

D. Tax evasion in 2004 (test of consistency 2004&2006)

Survey	Evaders	Non-evaders	Total
2000	227	836	1062
2004	229	762	991
Test statistics	**-0.955**	p-value:	**0.170**

Table 4. Markov transition matrix between years t and t+1.

Tax evasion		t+1	
		Evade	*Never evade*
t	*Evade*	P_{ee}	P_{en}
	Never evade	P_{ne}	P_{nn}

Table 5. Estimated short-term transition matrices, with 95% confidence intervals.

1999/2000		2000 Evaders	2000 Non-evaders
1999	Evaders	**0.21** (0.18 , 0.23)	**0** (0.0 , 0.0)
	Non-evaders	**0.05** (0.03 , 0.06)	**0.75** (0.72 , 0.77)

2000/2002		2002 Evaders	2002 Non-evaders
2000	Evaders	**0.21** (0.18 , 0.24)	**0.05** (0.04 , 0.06)
	Non-evaders	**0.03** (0.02 , 0.04)	**0.71** (0.69, 0.74)

2002/2004		2004 Evaders	2004 Non-evaders
2002	Evaders	**0.18** (0.16 , 0.21)	**0.05** (0.03 , 0.06)
	Non-evaders	**0.03** (0.02 , 0.04)	**0.74** (0.71 , 0.76)

2004/2006		2006 Evaders	2006 Non-evaders
2004	Evaders	**0.19** (0.17 , 0.22)	**0.04** (0.03 , 0.05)
	Non-evaders	**0.03** (0.02 , 0.04)	**0.74** (0.71 , 0.77)

Source: Authors' computation using 2000-2006 surveys. The first lines contain mean of each frequency expressed in percents, the second lines give estimated 95% confidence interval. Confidence interval formulas are given in the appendix.

Table 6. Predictions using fixed Markov (short-term) transition matrices.

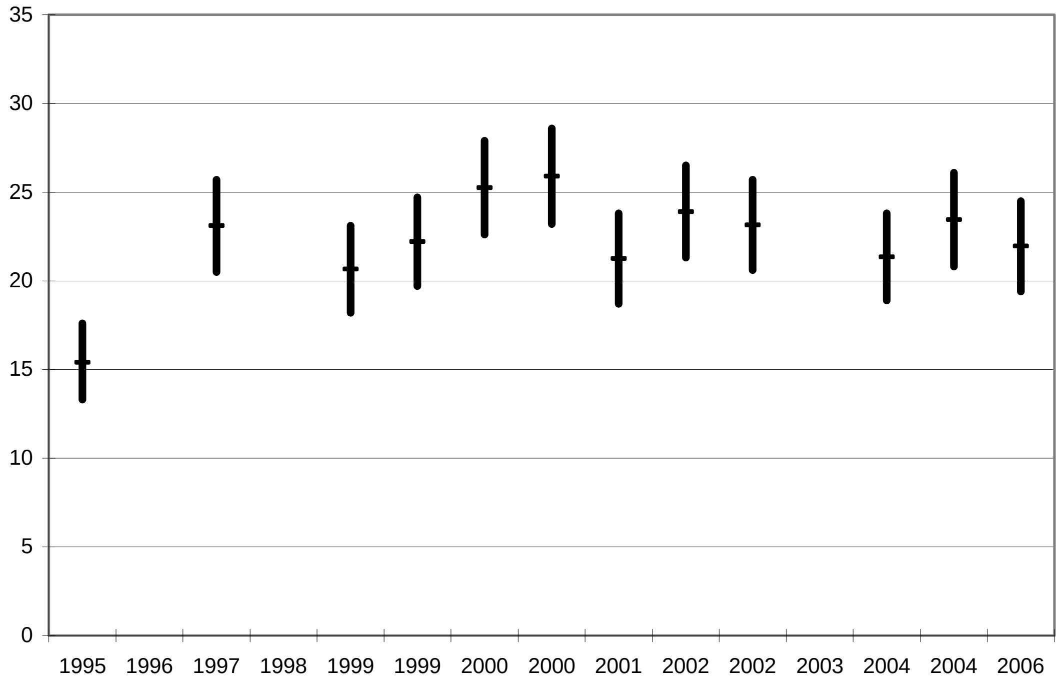
year	Estimation (and prediction) is based on survey conducted in:			
	2000	**2002**	**2004**	**2006**
2000	**0.25** (0.23, 0.28)			
2002	0.29 (0.26, 0.32)	**0.24** (0.21, 0.27)		
2004	0.32 (0.29, 0.36)	0.22 (0.19, 0.26)	**0.21** (0.19, 0.24)	
2006	0.36 (0.31, 0.40)	0.21 (0.17, 0.25)	0.2 (0.17, 0.23)	**0.22** (0.19, 0.25)

Note: As before, the first lines of each cell contain estimated probability expressed in percentages, the second lines give estimated 95% confidence interval. Shaded cells indicate predicted values statistically significantly different from actual values.

Table 7. Contingency table for test of stability of transition matrices.

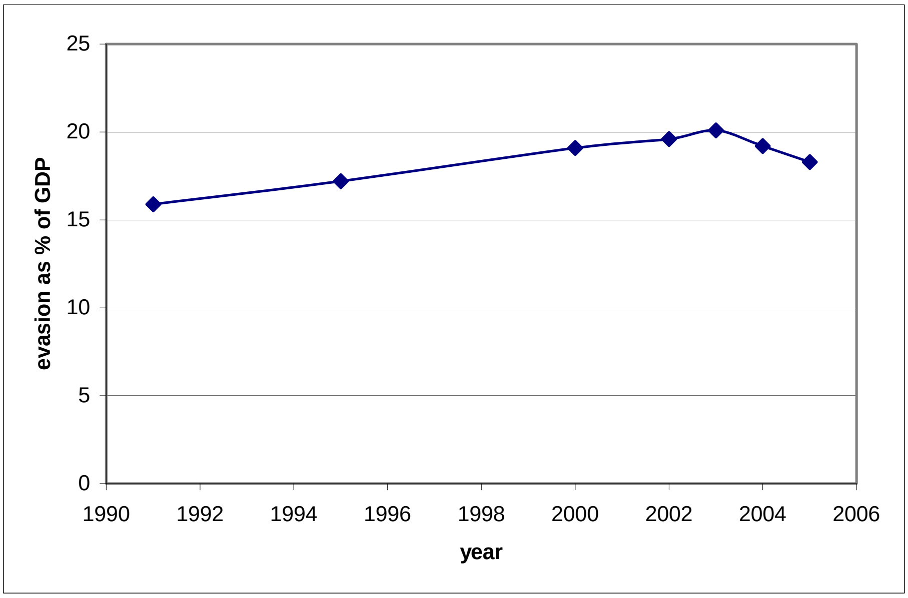
Sample (survey)	Change in evasion categories				Total
	1 [E->E]	2 [E->N]	3 [N->E]	4 [N->N]	
1 (2000)	n_{11} 168	n_{12} 0	n_{13} 100	n_{14} 795	$n_{1.}$ 1063
2 (2002)	n_{21} 194	n_{22} 44	n_{23} 73	n_{24} 720	$n_{2.}$ 1031
3 (2004)	n_{31} 148	n_{32} 78	n_{33} 86	n_{34} 741	$n_{3.}$ 1053
4 (2006)	n_{41} 188	n_{42} 40	n_{43} 30	n_{44} 733	$n_{4.}$ 991
	$n_{.1}$ 817	$n_{.2}$ 122	$n_{.3}$ 160	$n_{.4}$ 3043	n 4142

Figure 1. **Trends in estimated confidence intervals for percentage of tax evaders.**



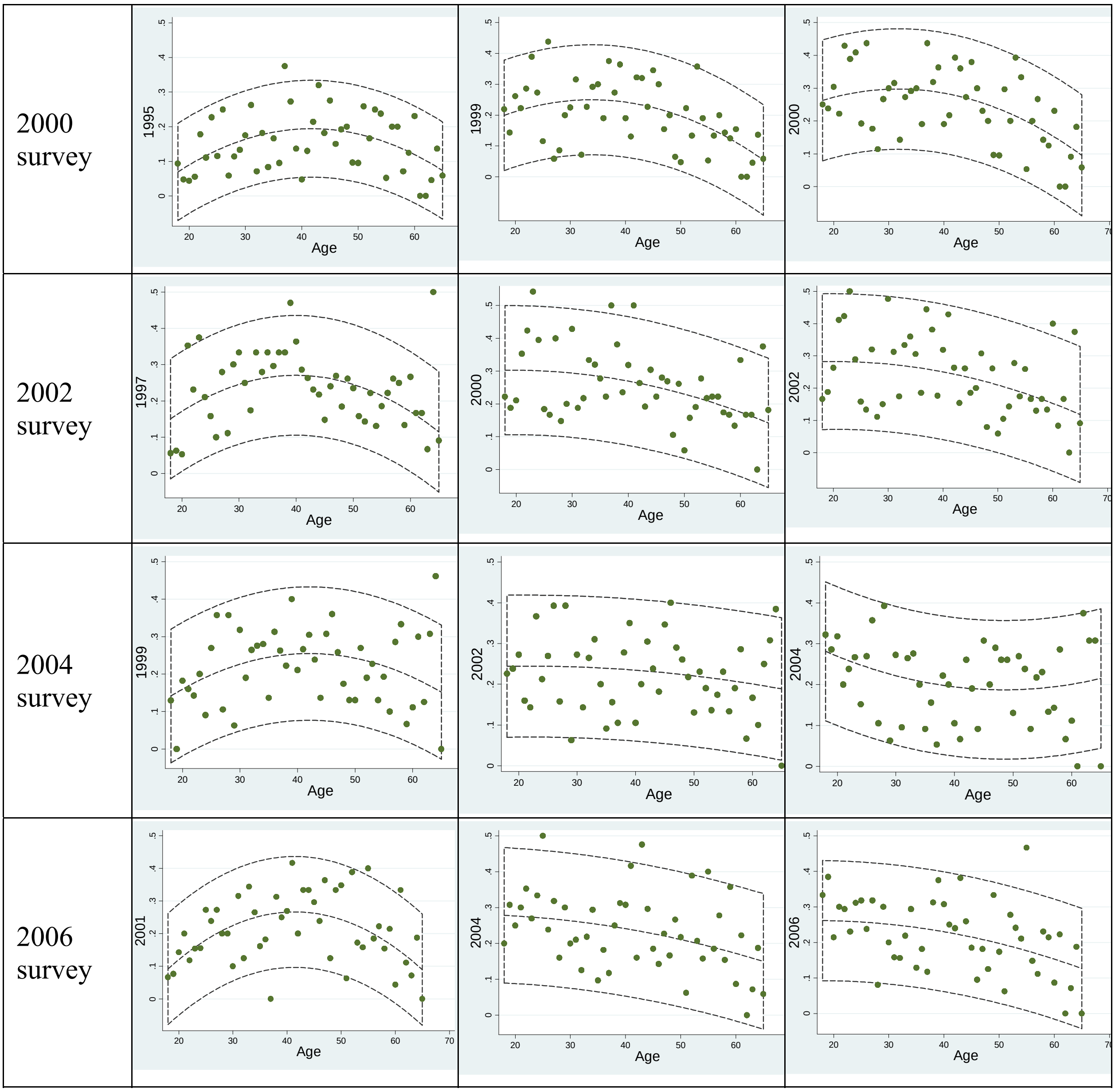
Source: Table 3.

Figure 2: Evasion as a percentage of GDP in the Czech Republic 1991-2003 according to Schneider et al.



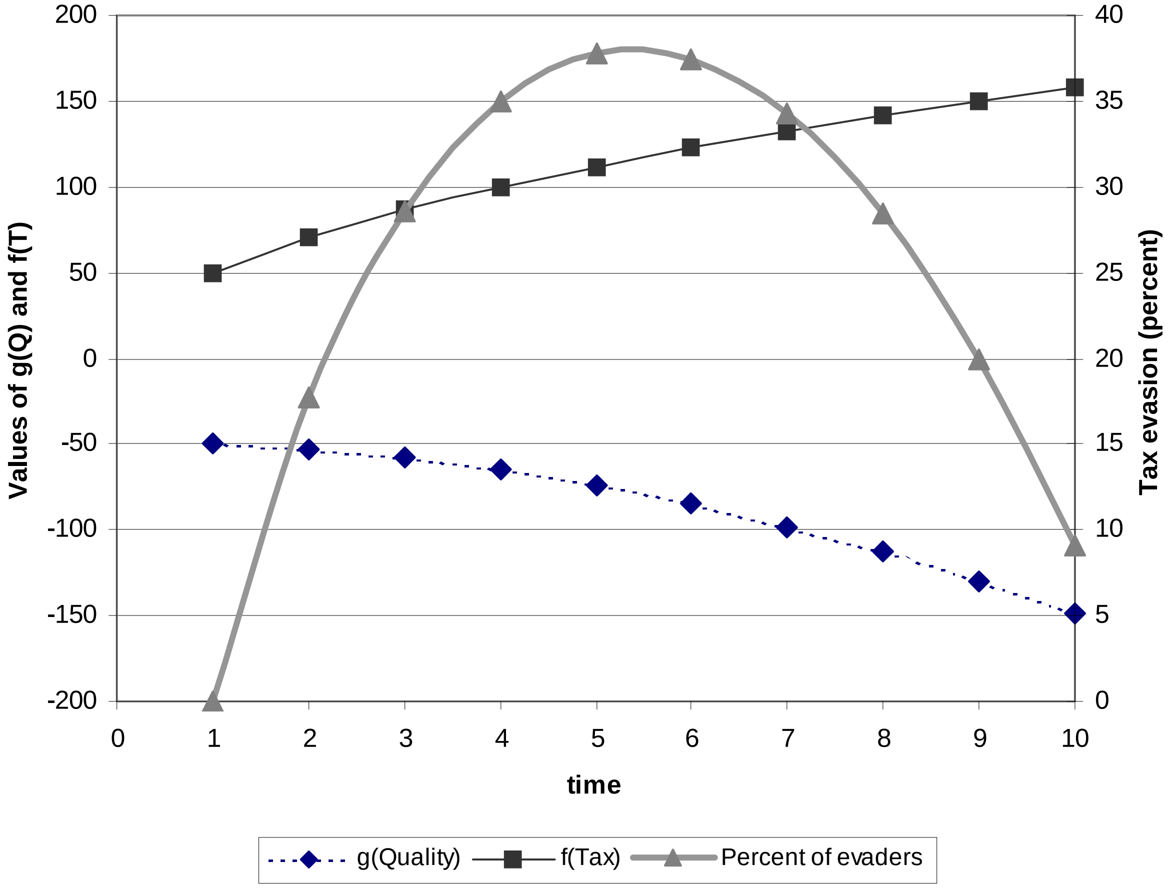
Source : Schneider (2005, 2006, 2007), Schneider and Klinglmair (2004)

Figure 3: Evasion rates by age groups for all four surveys



Source: Our 2000, 2002, 2004, 2006 surveys of tax evasion in the Czech Republic. Upper and lower lines indicate standard deviation bandwidths.

Figure 4: The evasional Kuznets curve as the sum of the influence of perceived quality of government services and index of taxes



Individual researchers, as well as the on-line and printed versions of the CERGE-EI Working Papers (including their dissemination) were supported from the following institutional grants:

- Economic Aspects of EU and EMU Entry [Ekonomické aspekty vstupu do Evropské unie a Evropské měnové unie], No. AVOZ70850503, (2005-2010);
- Economic Impact of European Integration on the Czech Republic [Ekonomické dopady evropské integrace na ČR], No. MSM0021620846, (2005-2011);

Specific research support and/or other grants the researchers/publications benefited from are acknowledged at the beginning of the Paper.

Published by
Charles University in Prague, Center for Economic Research and Graduate Education (CERGE)
and
Economics Institute ASCR, v. v. i. (EI)
CERGE-EI, Politických vězňů 7, 111 21 Prague 1, tel.: +420 224 005 153, Czech Republic.
Printed by CERGE-EI, Prague
Subscription: CERGE-EI homepage: http://www.cerge-ei.cz

Editors: Directors of CERGE and EI
Managing editors: Deputy Directors for Research of CERGE and EI

ISSN 1211-3298
ISBN 978-80-7343-160-0 (Univerzita Karlova. Centrum pro ekonomický výzkum a doktorské studium)
ISBN 978-80-7344-149-4 (Národohospodářský ústav AV ČR, v. v. i.)

CERGE-EI
P.O.BOX 882
Politických vězňů 7
111 21 Praha 1
Czech Republic
http://www.cerge-ei.cz