Tax Evasion under Market Incompleteness

Marco Maffezzoli[1]
Università Commerciale "L. Bocconi" and IGIER

This draft: September 5th, 2011

[1]Department of Economics, Università Bocconi, Via Sarfatti 25, 20136 Milano (MI), Italy. E-mail: marco.maffezzoli@unibocconi.it. I am extremely grateful to Alberto Alesina, Christopher Carroll, Jean-Bernard Chatelain, Bruno Chiarini, Gianluca Femminis, Nicola Pavoni, Roberto Perotti, and seminar participants at Università di Cagliari (CRENoS DECA), Università Cattolica di Milano, AFI Workshop 2011 (Milan), 7th CSEF-IGIER Symposium (Anacapri), SHADOW 2011 (Münster), EEA-ESEM 2011 Congress (Oslo), for their useful comments and suggestions. All remaining errors are my own. Financial support from the Italian MIUR (grant COFIN 2008) is gratefully acknowledged.

Abstract

The available empirical evidence suggests that the distribution of income and its composition play an important role in explaining tax noncompliance. We address the issue from a macroeconomic point of view, building a dynamic general equilibrium Bewley-Huggett-Aiyagari model that jointly endogenizes tax evasion and income heterogeneity. Our results show that the model can successfully replicate the salient qualitative and quantitative features of U.S. data. In particular, the model replicates fairly well the shape of the cross-sectional distribution of misreporting rates over true income levels. Furthermore, we show that a switch from progressive to proportional taxation has important quantitative effects on noncompliance rates and tax revenues.

Keywords: Tax Evasion, Income Heterogeneity, Incomplete markets.

JEL codes: E13, E26, H26

1 Introduction

If one considers the economic history of taxation, as narrated for instance in Webber and Wildavsky (1986), then it clearly appears that tax evasion has been part of the picture from the very beginning: a substantial amount of tax evasion has always existed in the past, still exists in the present, and will probably exist in the foreseeable future. However, measuring the extent of tax evasion, or even exactly defining the dividing line between illegal evasion and legal avoidance, is far from a straightforward task.[1]

The U.S. Internal Revenue Service (IRS) provides careful and comprehensive estimates of the extent and nature of tax noncompliance since 1979. The IRS periodically estimates the "tax gap," i.e. how much tax should be paid, but is not paid voluntarily in a timely way, providing separate estimates of the failure to pay due to nonfiling, underreporting of tax due on tax returns, and nonpayment or late payment of taxes owed. These estimates are based on information from a program of random intensive audits,[2] combined with information obtained from enforcement activities and special studies about particular sources of income that can be difficult to uncover even in an intensive audit. The last available estimates, based on the data collected by the National Research Program (NRP) for the 2001 tax year, are extensively described in Slemrod (2007) and Johns and Slemrod (2008).

Table 1 summarizes the tax gap estimated in 2001, together with its main components.[3] The overall gross tax gap (gross of enforced and other late payments) is $345 billion, of which 83% can be attributed to underreporting, 8% to nonfiling, and 10% to underpayment. The overall tax gap amounts to 16% of estimated actual (paid plus unpaid) tax liability. Underreporting of the individual income tax is by far the most important component of the tax gap, accounting for about two-thirds of the total amount. Looking at the individual income tax in more detail, we realize that income underreporting, as opposed to the overreporting of offsets to income, accounts for 81% of total underreporting. Underreporting of business income, as opposed to underreporting of wages and salaries and investment income, accounts for about two-thirds of the understated individual income.

[1] Andreoni, Erard, and Feinstein (1998) survey the most important theoretical and empirical issues related to tax noncompliance. Another useful general reference is Slemrod and Yitzhaki (2002).

[2] The program of random intensive audits was originally known as the Taxpayer Compliance Measurement Program (TCMP); it began in 1968 and lasted until 1988, having been cancelled in 1995. A modified version of the program, the National Research Program (NRP), was implemented to examine individual income tax returns from the 2001 tax year. The IRS randomly selected about 46,000 returns for review, oversampling high-income returns as well as individual taxpayers who reported sole proprietorship income. All of these returns were given a manual review, and a subset of those returns has been selected for in-person audits. To correct for the errors potentially introduced by variability in auditor judgment, a modified version of the correction procedure developed in Feinstein (1991) was employed. Finally, the estimates made significant adjustments for undetected noncompliance that relayed on special studies of particular sources of income and deductions. See Slemrod (2007) for further details.

[3] The last column in Tables 1-2 refers to the percentage of the corresponding true income, offsets to income, credits, or tax depending on the row of the table.

	Tax Gap	Share of Tax Gap	Share of True Amount
	($Billion)	*(%)*	*(%)*
Gross Tax Gap	**345**	-	**16**
Underreporting	285	83%	13
Individual Income Tax	197	57%	18
Underreported Nonbusiness Income	*56*	*28%*	*4*
Wages and salaries	10	5%	1
Net capital gains	11	6%	12
Taxable pension annuities, IRA distributions	4	2%	4
Taxable interest and dividends	3	2%	4
Other	28	14%	38
Underreported Business Income	*109*	*55%*	*43*
Nonfarm proprietor income	68	35%	57
Partnership, S corp., estate and net trust inc.	22	11%	18
Rent and royalty net income	13	7%	51
Farm net income	6	3%	72
Overreported Offsets to Income	*15*	*8%*	*4*
Overreported Credits	*17*	*9%*	*26*
Employment Tax	54	16%	7
Corporation Income Tax	30	9%	17
Estate and Excise Taxes	4	1%	4
Nonfiling	27	8%	1
Underpayment	34	10%	2

Table 1: The U.S. Tax Gap in 2001.

Given this empirical evidence, the quantitative importance of tax evasion can hardly be overemphasized: a few simple back-of-the-envelope calculations[4] show that, under the conservative assumption of constant noncompliance rates, the annual tax revenues actually lost since 2001 have ranged between $257 and $376 billion a year, while cumulative losses during the 2001-2010 period are estimated at just over $3 trillion: this unrealized revenue is the equivalent of 39% of the new national debt accumulated by the federal government over the same period.

An interesting aspect of the data summarized in Table 1 is the remarkable variance of the rate of misreporting as percentage of true income by type of income. While only 1% of wages and salaries are misreported, and 4% of interest and dividends, the misreporting rate rises to 57% for nonfarm proprietor income, and skyrockets to 72% for farm net income. Wages and salaries, interest, and dividends are subject to extensive information reporting, i.e. they must all be reported to the IRS by those who pay them; furthermore, wages and salaries are subject to employer withholding. In contrast, self-employment business income is not subject to any kind of information reporting. Table 2 summarizes the Net Misreporting Percentages (NMP) for income types subject to different degrees of information reporting.[5] This casual evidence suggest that the absence of information reporting is positively, and dramatically, correlated to the rate of misreporting, and that this relationship is particularly evident in the case of self-employment business income. More sophisticated empirical work has actually confirmed this impression.[6]

The IRS estimates reported in Tables 1 and 2 can be compared to some aggregate measures provided by the BEA as part of the NIPAs: NIPA Table 7.14 compares the net profit of nonfarm proprietorships and partnerships reported by the IRS to nonfarm proprietors' income, Table 7.16 compares the total receipts less total deductions of corporations reported by the IRS to profits before taxes, while Table 7.18 compares total wages and salaries reported by the BLS to wage and salary disbursements. The adjustment for

[4] See D. Callahan, "*Tax Evasion: The Real Costs*," April 2011, available at `http://www.ourfiscalsecurity.org/taxes-matter/2011/4/15/tax-evasion-the-real-costs.html`.

[5] *Substantial information reporting and withholding:* wages and salaries; *substantial information reporting*: pensions & annuities, dividend income, interest income, unemployment compensation, social security benefits; *some information reporting*: deductions, partnership / s-corp. income, exemptions, capital gains, alimony income; *no information reporting*: nonfarm proprietor income, other income, rents and royalties, farm income, Form 4797 income, adjustments.

[6] Klepper and Nagin (1989) show that noncompliance rates are related to proxies for the traceability, deniability, and ambiguity of income items, which are in turn related to the probability of detection. Pissarides and Weber (1989) show that, conditional on household characteristics and reported incomes, the self-employed spend a higher proportion of their income on food, and they suggest that this reflects an underreporting of income, not a higher propensity to consume food. Combining data from the Bank of Italy and from SeCIT, the Tax Auditing Office of the Italian Ministry of Finance, Fiorio and D'Amuri (2005) find that tax evasion is consistently higher for self-employment income than for employment income. Feldman and Slemrod (2007), exploiting the relationship between reported charitable contributions and reported income from wages and salary as compared to alternative reported income sources, obtain qualitatively similar results for the U.S. More recently, Kleven, Knudsen, Kreiner, Pedersen, and Saez (2010), using data from a randomized tax enforcement experiment in Denmark, show that the tax evasion rate is very small for income subject to third-party reporting, but substantial for self-reported income.

	Tax Gap *($ Billion)*	Share *(%)*	NMP *(%)*
Underreported Individual Income Tax	**197**	**-**	**18**
Substantial information reporting and withholding	11	5	1
Substantial information reporting	9	5	5
Some information reporting	51	26	9
No information reporting	110	56	54
Credits	*17*	*9*	*26*

Table 2: Information reporting and tax noncompliance.

misreporting is explicitly reported, and this allows us to construct the implicit aggregate NMPs, equal respectively to 51.8%, 23.6%, and 1.4% in 2001: these independent estimates, while being obviously different from the IRS ones summarized in Table 1, are of the same order of magnitude, and this seems reassuring.

Johns and Slemrod (2008) assess the distributional consequences of income tax noncompliance using the 2001 NRP data, supplemented with IRS-calculated estimates of unreported income.[7] Figure 1 summarizes some of their findings: we plot the NMP versus the "true" Adjusted Gross Income (AGI), i.e. the reported AGI augmented with the estimated amount of noncompliance; we also report the share of AGI not subject to substantial information reporting. The NMP remains relatively constant for the income brackets below $\$100K$, slowly increasing from around 3% for the lowest income bracket to around 8%; then, it sharply increases when we consider income brackets ranging from $\$100K$ to $\$1M$, reaching a maximum value of 21%; finally, it decreases for the even higher income brackets, settling down on a 11% for incomes above $\$2M$. The share of AGI not subject to substantial information reporting remains relatively constant for incomes below $\$100K$, but then starts to rapidly and monotonically increase, reaching a maximum value of 55.4% at the highest income bracket. The two distributions essentially overlap for income brackets below $\$75K$, suggesting that for for those income brackets the NMP on income not subject to information reporting has to be substantial, possibly near 100%; they also seem to share the same "turning point," i.e. the $\$75K - \$100K$ income bracket: above this threshold, they tend to jointly increase until the $\$500K - \$1M$ bracket. This casual evidence, together with the previously described facts reported in Table 2, may suggest that the distribution of income and its composition play an important role in determining the amount of tax noncompliance.[8]

[7] The IRS employs a sophisticated econometric technique, denominated Detection Control Estimation (DCE), that has been initially introduced by Feinstein (1991). The DCE methodology estimates, via joint maximum likelihood, a noncompliance equation that models the total amount of unreported income and a detection equation that models the fraction of noncompliance detected by the IRS examiner. For further details, see also Johns and Slemrod (2008).

[8] The empirical relationship between tax noncompliance and the distribution of income in the U.S. is also discussed in Bloomquist (2003). Matsaganis, Benedek, Flevotomou, Lelkes, Mantovani, and Nienadowska (2010) provide evidence for Greece, Hungary and Italy. Persson and Wissén (1984) discussed the

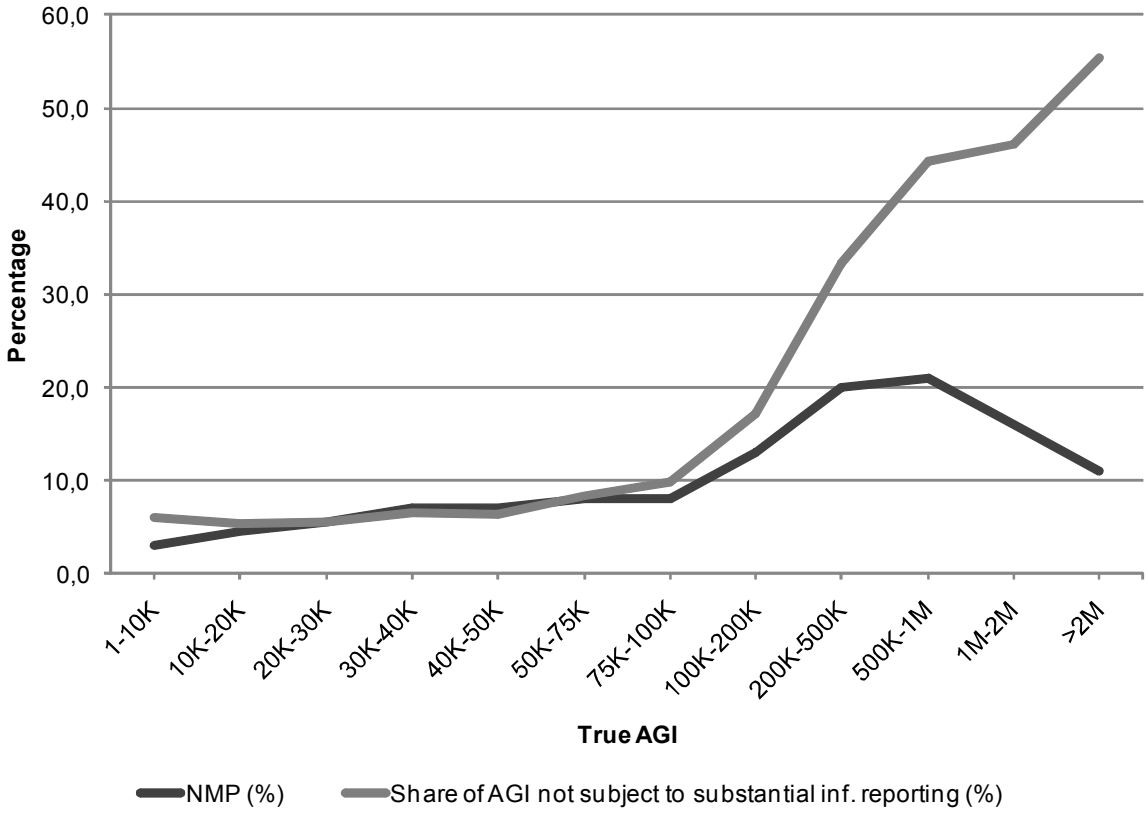

Figure 1: Distribution of NMP by true AGI in 2001.

We take this suggestion seriously, and address the issue from a macroeconomic and quantitative point of view. We construct a dynamic general equilibrium Bewley-Huggett-Aiyagari model that jointly endogenizes tax evasion and income heterogeneity. Our framework blends to distinct literatures: we combine the dynamic general equilibrium model of income heterogeneity pioneered by Bewley (1980), Huggett (1993), Aiyagari (1994) and Huggett (1997), with a dynamic version of the classical deterrence model of tax evasion, developed by Allingham and Sandmo (1972) and Yitzhaki (1974), which draws on Andreoni (1992).[9]

Our model is populated by a continuum of households that consume and save, subject to a credit constraint, and face idiosyncratic shocks to their income. They receive labor and capital income, possibly in different proportions, and are subject to a progressive tax schedule on total individual income. While labor income cannot be concealed, capital income can be misreported: this allows our households to at least partially evade taxation in the current period, but exposes them to the possibility of being audited, at an exogenous rate, in the next one. Audited households pay the previously evaded taxes

issue from a theoretical point of view.

[9]Heathcote, Storesletten, and Violante (2009) survey the recent literature on heterogeneous agents economies, while Andreoni et al. (1998), Cowell (2003), and Sandmo (2005) effectively present, discuss, and evaluate the classical deterrence model of tax evasion. The issue has been also studied using multiagent-based simulation models: for a survey, see Bloomquist (2006).

back, plus a substantial fine. The inability to fully insure against idiosyncratic shocks generates in steady state an endogenous stationary wealth distribution, which in turn implies a distribution of the capital income share in individual income that, thanks to the endogeneity of tax evasion, maps into a distribution of misreporting rates.

Our results suggest that the model is substantially able to reproduce the qualitative and quantitative features of the data. In particular, once we calibrate it to reproduce the estimated misreporting rate for total income, the model generates, in steady state, an average misreporting rate for income not subject to information reporting that seems in line with both the evidence summarized in Table 2 and the previosuly discussed BEA estimates. Furthermore, the model generates a distribution of misreporting rates over true income levels that mimics fairly well the estimated distribution represented in Figure 1: in particular, we are able to capture its single-peaked shape, due to the decrease of the misreporting rate that takes place for higher income brackets.

We perform some additional experiments. Our results suggest that: (i) an increase of the average tax rate, *ceteris paribus*, would unequivocally and significantly *increase* in steady state the noncompliance rates, without however preventing a substantial rise in tax revenues; (ii) moving from a progressive to a proportional tax system, while leaving the average tax rate on reported income unaffected, would drastically reduce, again in steady state, tax evasion, increase government revenues, and increase households' welfare, at the cost of higher income inequality; (iii) if the switch from progressive to proportional taxation is made revenue-neutral, via a contemporaneous decrease of the average tax rate, then the previously described effects are significantly enhanced. The results of the last two experiments seem broadly in line with the empirical evidence reported in Ivanova, Keen, and Klemm (2005) and Gorodnichenko, Martinez-Vazquez, and Peter (2009), who studied Russia's 2001 flat rate income tax reform and found large and significant effects of the reform on tax evasion and voluntary compliance.

The rest of the paper is structured as follows: section 2 discusses the model in some detail, Section 3 presents the benchmark calibration, Section 4 summarizes the results, and finally Section 5 concludes.

2 The model

There exists a continuum of *ex-ante* identical and infinitely lived households, with total mass equal to one. Households own both factors of production, capital and labor, and rent them to the firms on competitive factor markets. Firms buy factor services from households and produce a single homogenous good competitively, via a constant-returns-to-scale production function. The good can be used for both consumption and investment. As in Huggett (1993) and Aiyagari (1994), asset markets are incomplete: households are allowed to investment in physical capital accumulation only, and we assume that capital

holdings cannot be negative. Hence, households cannot fully insure themselves against idiosyncratic shocks to their income.

The government levies taxes on income and redistributes the revenues via lump-sum transfers. However, it cannot observe the entire amount of taxable income directly, but instead relies on households to self-report their incomes. Compliance is enforced through random audits and penalties levied on observed underreporting; however, as in Andreoni (1992), there is a delay between the time when the report is made and the time the audits occur. Hence, households can consume the benefits of tax evasion immediately, but they do not run the risk of being audited until the next period.

The next Sections will describe the model components more in detail. The recursive equilibrium is formally defined in Appendix A.

2.1 Households

Following Huggett (1997), we assume that, at the beginning of each period, households receive a fixed labor endowment, measured in efficiency units, and supply it inelastically to the labor market. The labor endowment is modelled as a finite-state Markov process, characterized by a transition matrix $\boldsymbol{\pi}_N$, which evolves independently across households.

Furthermore, households inherit a stock of evaded taxes from the past, denoted $e_t \geq 0$, and face the risk of being randomly selected for a tax audit: this individual status is modelled as an exogenous random variable, denoted $\varepsilon_t \in E = \{0, 1\}$, which determines whether the household is going to be audited in the current period, $\varepsilon = 1$, or not, $\varepsilon = 0$, and follows again a finite-state Markov process characterized by a transition matrix $\boldsymbol{\pi}_E$. If the household is audited, the government always learns the true size of e_t, and, if underreporting has been detected, the household is forced to pay a proportional fine $\mu > 1$ on the inherited stock of tax evasion.

The two stochastic processes are independent, and can be jointly represented by a finite-state Markov process, denoted $\sigma \in E \times L$, characterized by a transition matrix $\boldsymbol{\pi} = \boldsymbol{\pi}_E \otimes \boldsymbol{\pi}_N$ such that $\pi(\sigma_j, \sigma_i) \geq 0$ stands for the probability that $\sigma_{t+1} = \sigma_j$ if $\sigma_t = \sigma_i$, where, for the sake of notational convenience, $\sigma \equiv \{l, \varepsilon\}$.

During each period, households determine the optimal amount of underreporting. Denote $y_t \equiv r_t k_t + w_t l_t$ the total amount of taxable income in period t for a generic household: w_t represents the wage rate, r_t the interest rate, and k_t household's current capital stock. If $\varsigma_L \in (0, 1)$ and $\varsigma_K \in (0, 1)$ stand respectively for the shares of labor and capital incomes that the government cannot directly observe, i.e. income not subject to information reporting, then the total amount of income that the household can potentially underreport is $\bar{y}_t \equiv \varsigma_K r_t k_t + \varsigma_L w_t l_t$. Therefore, if $0 \leq \theta_t \leq 1$ denotes the share of taxable income that is voluntarily not reported, then the net after-tax income amounts to $y_t - \mathcal{T}(d_t)$, where $\mathcal{T}(x)$ is an effective tax function that expresses the total amount of

taxes paid by an individual with pre-tax income x, $d_t \equiv y_t - z_t$ denotes the amount of reported income, and $z_t \equiv \theta_t \bar{y}_t$ the amount of voluntarily unreported income.

The stock of evaded taxes evolves according to the following accumulation equation:

$$e_{t+1} = (1 - \varepsilon_t)(1 - \delta_E) e_t + \mathcal{T}(y_t) - \mathcal{T}(d_t), \tag{1}$$

where $\delta_E \in (0, 1]$ represents an exogenous depreciation rate. If $\delta_E = 1$, then the auditing process detects tax evasion occurred in the previous period only. If $\delta_E < 1$, instead, the government is able to partially recover also taxes evaded in later periods. This modelling device captures, quite crudely, the idea that a tax audit, if evasion is actually detected, may be extended to previous years, and therefore may allow to uncover further evasion.[10]

Households' preferences over stochastic consumption streams are given by:

$$v_t \equiv \mathbb{E}_t \left[\sum_{s=t}^{\infty} \beta^{s-t} \frac{\left(c_s - \frac{\omega}{\nu} z_s^{\nu}\right)^{1-\gamma} - 1}{1 - \gamma} \right], \tag{2}$$

where c_t is the consumption level, $\beta \in (0, 1)$ the intertemporal discount factor, $\gamma > 0$ the reciprocal of the elasticity of intertemporal substitution, $\omega \geq 0$ and $\nu > 0$ two preference parameters. In the spirit of Gordon (1989), we assume that the amount of concealed income affects negatively the current utility level: this can be interpreted in terms of tax morale and its determinants, like individual morality, reputation costs, and so on.[11] The particular functional form used in (2) is borrowed from Greenwood, Hercowitz, and Huffman (1988): the marginal rate of substitution between consumption and concealed income does depend on the latter only.[12]

The stock of physical capital, k_t, evolves over time according to the following accumulation equation:

$$k_{t+1} = (1 - \delta_K) k_t + y_t - \mathcal{T}(d_t) + G_t - \varepsilon_t \mu e_t - c_t, \tag{3}$$

where δ_K is a physical depreciation rate, and G_t denotes aggregate (i.e. *per capita*) lump-sum transfers, taken as given by the household. As already mentioned, households face a borrowing constraint: $k_{t+1} \geq 0$.

[10] Our approach is similar, at least in spirit, to Niepelt (2005).

[11] See Sandmo (2005) for further details. Bordignon (1993) discusses also the role of perceived fairness of the fiscal system, with respect to both governmental supply of public goods and the perceived behavior of other taxpayers.

[12] Introducing the disutility generated by concealed income has admittedly a very mundane goal: it allows us to easily calibrate the degree of tax evasion to the U.S. aggregate data. It is well known in the literature - see for instance Slemrod (2007) - that the classical variants of the Allingham and Sandmo (1972) framework generate, under sensible parametrizations, tax evasion rates that are far above the observed ones. A vast literature has developed alternative ways to solve this problem: we just chose the simplest one. An equally simple, and almost equivalent, approach is the one followed in Chen (2003), who introduces a quadratic monetary cost of tax evasion.

We can now put all the elements together; for given sequences of factor prices and transfers, the dynamic optimization problem of a generic household is as follows:

$$\begin{aligned}
\max_{\{c_s,\theta_s,k_{s+1},e_{s+1}\}_{s=t}^{\infty}} \quad & \mathbb{E}_t\left[\sum_{s=t}^{\infty}\beta^{s-t}\frac{\left(c_s-\frac{\omega}{\nu}z_s^{\nu}\right)^{1-\gamma}-1}{1-\gamma}\right], \\
\text{s.t.} \quad & k_{t+1}=(1-\delta_K)\,k_t+y_t-\mathcal{T}(d_t)+G_t-\varepsilon_t\mu e_t-c_t, \\
& e_{t+1}=(1-\varepsilon_t)(1-\delta_E)\,e_t+\mathcal{T}(y_t)-\mathcal{T}(d_t), \\
& \theta_t\geq 0, \\
& 1-\theta_t\geq 0, \\
& k_{t+1}\geq 0.
\end{aligned} \tag{4}$$

The first order conditions can be combined to obtain the following Euler equations:

$$\begin{aligned}
u_{c,t}-\varphi_{k,t} &= \beta\mathbb{E}_t\left(u_{c,t+1}\left\{1-\delta_K+\left[1-\mathcal{T}_y(d_{t+1})\right]r_{t+1}\right\}+\right. \\
&\quad \left.\left[\varkappa_{t+1}\theta_{t+1}\varsigma_K+\left(\psi_{t+1}-\varkappa_{t+1}\right)\digamma_{t+1}\right]r_{t+1}\right), \qquad (5) \\
\psi_t-\varkappa_t &= \beta\mathbb{E}_t\left[\varepsilon_{t+1}u_{c,t+1}\mu\mathcal{T}_y(d_t)+\right. \\
&\quad \left.(1-\varepsilon_{t+1})\left(\psi_{t+1}-\varkappa_{t+1}\right)(1-\delta_E)\frac{\mathcal{T}_y(d_t)}{\mathcal{T}_y(d_{t+1})}\right]. \qquad (6)
\end{aligned}$$

where $\mathcal{T}_y$ denotes the first-order derivative of $\mathcal{T}$, and:

$$u_{c,t} = \left(c_t-\frac{\omega}{\nu}z_t^{\nu}\right)^{-\gamma}, \tag{7}$$

$$\psi_t \equiv u_{c,t}\left[\mathcal{T}_y(d_t)-\omega z_t^{\nu-1}\right], \tag{8}$$

$$\digamma_t \equiv 1-\frac{\mathcal{T}_y(y_t)}{\mathcal{T}_y(d_t)}, \tag{9}$$

$$\varkappa_t \equiv \frac{\varphi_{\theta,2,t}-\varphi_{\theta,1,t}}{\bar{y}_t}. \tag{10}$$

Note that $\varphi_{k,t}$, $\varphi_{\theta,1,t}$, and $\varphi_{\theta,2,t}$ are the Lagrange multipliers associated with the non-negativity constraints $k_{t+1}\geq 0$, $\theta_t\geq 0$, and $1-\theta_t\geq 0$.[13]

2.2 Firms

The competitive firms are characterized by a constant-returns-to-scale "Cobb-Douglas" technology; let K_t and N_t stand for the per-capita capital stock and labor supply; per-capita output is then given by:

$$Y_t=\phi K_t^{\alpha}N_t^{1-\alpha}, \tag{11}$$

[13] Hence: $\varkappa_t>0$ iif $\theta_t>1$, $\varkappa_t=0$ iff $\theta_t\in[0,1]$, and $\varkappa_t<0$ iif $\theta_t<0$.

and factor prices by:

$$w_t = \phi \left(\frac{K_t}{N_t} \right)^{\alpha}, \tag{12}$$

$$r_t = \phi \left(\frac{K_t}{N_t} \right)^{\alpha - 1}. \tag{13}$$

2.3 Government

The government plays a minimalist role, collecting tax revenues and fines, and paying everything back to the households via lump-sum sum transfers:

$$G_t = \int_X \left[\mathcal{T}_y \left(y_t - z_t \right) + \mu \varepsilon_t e_t \right] d\lambda_t. \tag{14}$$

3 Calibration

The parameters that characterize the household's preferences are selected in the following way: the intertemporal discount factor and the reciprocal of the elasticity of intertemporal substitution are set to standard values in the literature, $\beta = 0.95$ and $\gamma = 2$. The parameters governing the disutility of concealed income, ω and ν, can hardly be separately identified, hence we arbitrarily set $\nu = 2$, and calibrate ω in order to reproduce in steady state the observed net misreporting percentage for total individual income, equal to 18% as shown in Table 2; the implied value for ω is 0.378.

Following Cooley and Prescott (1995), we set $\delta_K = 0.048$ and $\alpha = 0.4$; the productivity level ϕ is, without loss of generality, normalized to unity.

In order to parameterize the shares of labor and capital incomes not subject to information reporting, ς_L and ς_A, we proceed as follows. From Johns and Slemrod (2008), Table A4, p. 27, we take the composition of true income by source: the available decomposition accounts for Salaries and Wages, Interest, Dividends, Business (Sch. C, i.e. nonfarm sole proprietorships), Part. - S. Corp. - Estate and Trust, Capital Gains, and Other (income). Taking the actual classification to the extreme (see Footnote 5, p. 3), we consider Salaries and Wages, Interest, Dividends, and Other (income) as subject to information reporting. All items except Business (Sch. C) income can be univocally attributed to labor or capital income:[14] following again Cooley and Prescott (1995), we assume that the capital share in Business (Sch. C) income is equal to the aggregate one, in our case α. Hence, we compute the total share of capital and labor incomes not subject to information reporting, equal respectively to 85.1% and 4.7%.

[14] Business (Sch. C) income is roughly the counterpart of NIPA's Proprietor's Income, i.e. mainly income of self-employed individuals and sole proprietorships. This type of mixed income, as well know at least since Cooley and Prescott (1995), cannot be clearly attributed to labor or capital.

Following Conesa and Krueger (2006), we use a flexible functional form for the effective tax function that is theoretically motivated by the equal sacrifice principle, as discussed in Gouveia and Strauss (1994), and encompasses a wide range of progressive, proportional and regressive tax schedules:

$$\mathcal{T}(y) = a_0 \left[y - \left(y^{-a_1} + a_2 \right)^{-\frac{1}{a_1}} \right], \tag{15}$$

where $a_0 \geq 0$, $a_1 \geq 0$, and $a_2 \geq 0$.[15] Gouveia and Strauss (1994) estimate this tax function for the U.S., obtaining values of $a_0 = 0.258$ and $a_1 = 0.768$.[16] According to IRS's Statistics of Income, the ratio of total income tax revenues over AGI was 15.2% in 2001: the parameter a_2 is calibrated in order to reproduce this ratio in steady state; the implied value is 0.529.

We set the depreciation rate of previously accumulate evasion, δ_E, equal to one: this means that audited households pay the fine on taxes evaded in the previous period only; we will then discuss the effects of relaxing this assumption. The proportional fine, μ, is assumed to be equal to 1.75, which is in line with the discussion in Andreoni et al. (1998). Given that there is no clear evidence that having being audited in the past does *per se* change the probability of being audited by the IRS in the future, we assume that the probability of being audited is simply independent over time.[17] Under this assumption, the *ex-ante* probability of being audited corresponds to the *ex-post* share of households being audited in a given period. The IRS officially reports the examination coverage rates for several years: the average coverage rate for individual income tax returns, focusing on business returns only, was 1.5% for the Fiscal Year 2001. Hence, we set the transition matrix for the auditing process to:

$$\boldsymbol{\pi}_E = \begin{bmatrix} 0.985 & 0.985 \\ 0.015 & 0.015 \end{bmatrix}.$$

The log of the individual labor endowment is assumed to follow an autoregressive

[15] Note that if $a_1 \to 0$, then $\mathcal{T}(y) \to a_0 y$, i.e. the tax schedule collapses to a pure proportional system. If $a_1 > 0$, the system becomes progressive, and the overall progressivity increases with a_1.

[16] These estimates are for tax year 1989, the last year reported in Gouveia and Strauss (1994). We are currently not aware of any more recent estimates.

[17] Andreoni et al. (1998), par. 6.4, report that past audits do not seem to have any independent effect on the behavior of noncompliant tax payers. However, the available empirical evidence is clearly not conclusive.

process:[18]

$$\begin{aligned} \log l_{t+1} &= \bar{\iota} + \rho \log l_t + \epsilon_{t+1}, \\ \epsilon_t &\sim N\left(0, \sigma^2\right). \end{aligned}$$

Following Floden and Lindé (2001), we set $\rho = 0.9136$ and $\sigma = 0.2064$; we normalize the aggregate labor endowment in steady state to one, and set the parameter $\bar{\iota}$ accordingly. This process is approximated with a 5-state discrete Markov chain computed using Rouwenhorst's method, as suggested in Kopecky and Suen (2010).

As far as the solution method is concerned, our approach is fairly standard. At the household level, we have to solve a stochastic dynamic optimization problem with occasionally binding constraints: this is done using the time iteration algorithm described in Rendhal (2007). At the aggregate level, we compute the stationary distribution using a non-stochastic binning approach, extending the method described in Young (2010) to a bidimensional setting.[19] More details are provided in Appendix B.

4 Results

4.1 Benchmark parametrization

The main properties of the stationary equilibrium under our benchmark parametrization are summarized in the first column of Table 3. The aggregate capital stock is equal to 8.45, while its standard deviation across households is 6.98. The aggregate stock of taxes evaded in previous periods reaches 0.085, with a standard deviation of 0.039. The aggregate GDP level equals 2.35: this implies an aggregate capital-output ratio of 3.6, a figure broadly in line with the evidence reported in Cooley and Prescott (1995).

The average misreporting rate on total taxable income, being a calibration target, is exactly equal to 18%, the estimated NMP for the Individual Income Tax reported in Table 1.[20] The average misreporting rate on concealable income, instead, reaches a value of 48.7%, which is slightly lower but broadly consistent with both the 54% estimated NMP

[18] Two somehow conflicting views on the nature of idiosyncratic income processes have emerged in the literature: as discussed in Guvenen (2009), one view holds that individuals are subject to large and very persistent shocks, while facing similar life-cycle income profiles. The alternative view holds that individuals are subject to income shocks with low persistence, while facing individual-specific income profiles. See also Carroll (1997) for a detailed discussion. Given that currently the jury seems to be still out, our choice of a very persistent labor income process is mainly driven by comparability with the existing literature and numerical convenience.

[19] To solve for the policy functions, we discretize the state space using 200×200 nodes and employing linear multivariate interpolation to evaluate the functions at points that are not on the grid. To compute the stationary distribution, we increase the number of nodes to 400×400, using again linear interpolation. A further increase of the number of nodes does not substantially change the results. The grid points for capital and past evasion span respectively the intervals $[0, 40]$ and $[0, 0.18]$.

[20] For consistency with the data, average rates are computed as ratios of aggregate quantities.

	Experiments				
	(i) Bench.	(ii) Higher avg. tax rate $a_0 = 0.278$	(iii) Prop. taxation I $a_0 = 0.152$ $a_1 = 0$	(iv) Prop. taxation II $a_0 = 0.134$ $a_1 = 0$	(v) Flat-Tax Reform $a_0 = 0.152$ $a_1 = 0$ $a_3 = 0.189$
Physical cap.	8.45	8.20	9.34	9.70	9.34
(Std. dev.)	*(6.98)*	*(6.60)*	*(8.93)*	*(9.27)*	*(8.93)*
Evaded taxes	0.085	0.097	0.049	0.039	0.049
(Std. dev.)	*(0.039)*	*(0.044)*	*(0.018)*	*(0.013)*	*(0.018)*
GDP	2.35	2.32	2.44	2.48	2.44
NMP on conc. inc.	48.7%	52.5%	35.9%	31.8%	35.9%
NMP on total income	**18.0**%	19.3%	13.2%	11.7%	13.2%
Avg. tax evasion rate	22.4%	24.2%	13.2%	11.7%	14.3%
Government revenues	0.295	0.306	0.323	**0.295**	**0.295**
Avg. tax rate on true inc.	12.5%	13.1%	13.2%	11.8%	12.0%
Avg. tax rate on rep. inc.	**15.2**%	**16.2**%	**15.2**%	13.4%	13.8%
CEV	-	−0.64%	1.64%	2.30%	1.65%

Table 3: Selected steady-state features of the model.

for incomes not subject to information reporting (see Table 2) and the 51.8% reported by the BEA for nonfarm proprietors income. The average tax evasion rate, i.e. the ratio between the amount of taxes actually paid and the amount that should have been paid, equals 22.4%: in the model, more than a fifth of the potential government revenues are lost due to misreporting. The total amount of government revenues collected and then transferred back to the households in a lump-sum fashion, i.e. the amount of tax revenues plus the value of fines paid by audited households, reaches 0.295. Finally, the average tax rate effectively paid by households on their true income is equal to 12.5%; the average tax rate on reported income, being a calibration target, equals instead 15.2%, the observed ratio between the total income tax revenues and reported AGI.

Figure 2 compares the distribution of the NMP by true AGI, as reported in Johns and Slemrod (2008), to the distribution of the average misreporting rate by taxable income generated by the model. Note that we try to use comparable income brackets for the observed and simulated data: given that, according to the IRS *Statistics of Income* data for 2001, the households that report an income larger than $\$2M$ are just the 0.05% of the sample, we compute the corresponding threshold for the simulated data; hence, 100%

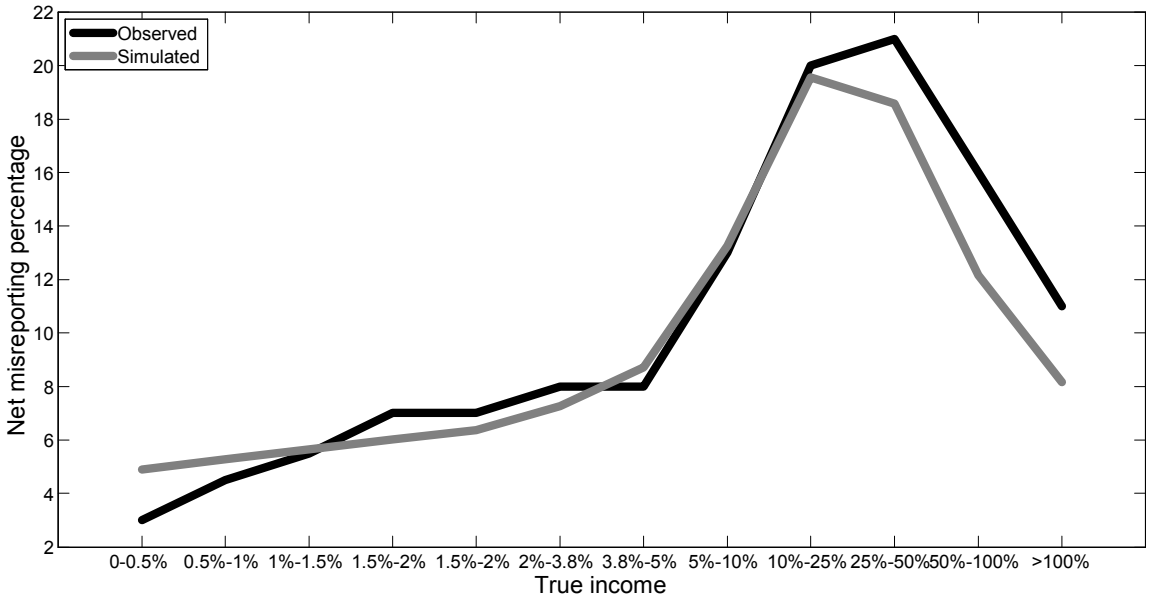

Figure 2: Distribution of NMP by true income: observed vs. simulated data.

in Figure 2 - and the following ones - corresponds to $\$2M$ in the observed data and to the $0.9995th$ income quantile in the simulated ones.[21] The match between the two curves seems relatively good from both a qualitative and a quantitative point of view: the shape of the two distributions is similar; in particular, the model is able to capture the sharp increase in the misreporting rate and its subsequent drop. However, there is somehow a quantitative mismatch at higher income brackets: for income levels above $\$200K - \$500K$, the simulated misreporting rate is lower ($2-3$ percentage points) than the observed one.

Figure 3 plots the distribution of the share of income not subject to information reporting, i.e. concealable income, by true AGI and its simulated counterpart. Evidently, the simulated distribution reproduces the observed one remarkably well: the fit is very good for incomes below $\$100K - \$200K$, while for higher brackets the simulated distribution tends to slightly underestimate the true one.

Figures 2 and 3 allows us also to disentangle the relative contribution of the two essential components of our model: endogenous tax evasion and income heterogeneity. Consider Figure 3, and assume we were able to kill the first component by imposing a constant and exogenous misreporting rate on concealable income equal, say, to 54%. This would leave the previously described average results substantially unaffected, but, given the simulated distribution of concealable income, it would make the model unable to reproduce the shape of the distribution of estimated misreporting rates on total income: the

[21] The procedure works as follows: first of all, we compute the distribution of the average misreporting rate per taxable income across all nodes in our state space. Then, having mapped the income brackets used for the observed data into comparable brackets for the simulated ones, we calculate the average misreporting rate for the extremes of each income bracket via linear interpolation on the simulated distribution. Finally, we compute the rate to report for each bracket as the average of the interpolated values at the extremes.

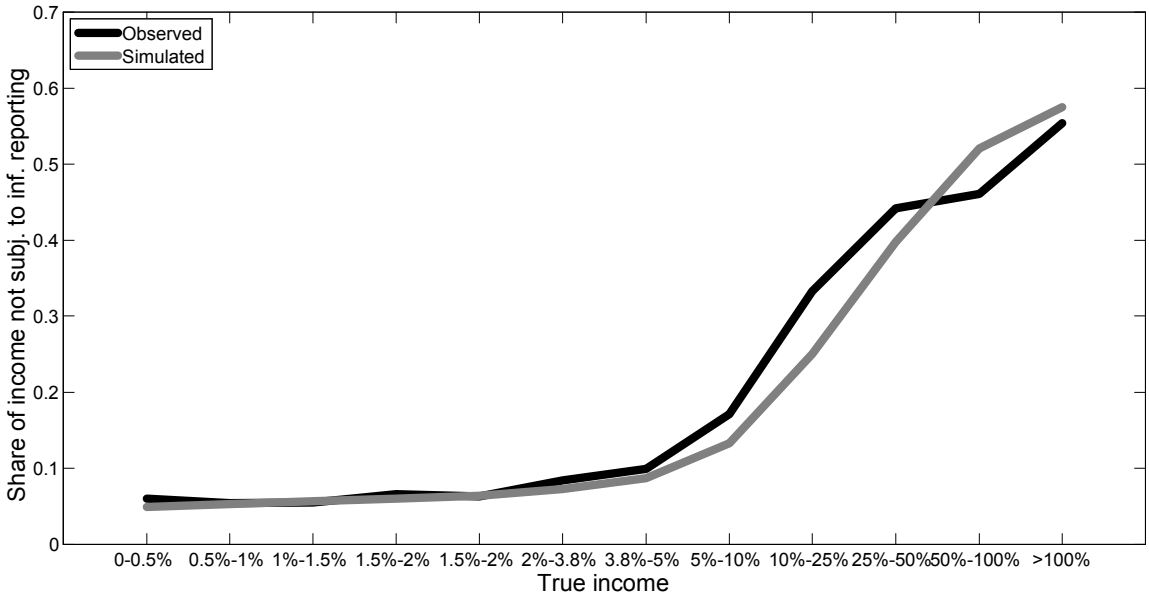

Figure 3: Distribution of the share of "concealable" income: observed vs. simulated.

simulated misreporting rate would always increase with income, completely missing the decrease observed for higher income brackets in the data. Here the endogenous reaction of underreporting becomes key: in our model, the marginal incentive to evade taxes decreases with consumption, and therefore indirectly with income, while the marginal utility cost of doing it increases with the stock of evaded taxes, and therefore again indirectly with income; as a result, the misreporting rate on concealable income tends to decrease with total income. However, given that for low income levels the share of concealable income is negligible, the misreporting rate on total income increases initially, following the evolution of the share of concealable income, until the latter becomes large enough to revert the mechanism and drive down the total misreporting rate again. Hence, the relative ability of our model to replicate the qualitative and quantitative features of the data hinges on the interaction of its two essential endogenous components.

4.2 Additional experiments

Table 3 summarizes the results for a set of additional experiments, performed changing the value of some relevant parameters but leaving everything else as in the benchmark calibration.

4.2.1 Higher taxes

The first experiment, reported in column (ii), studies the effects of an overall increase in taxation: we rise the proportionality factor a_0 in the tax function from 0.258 to 0.278, generating a one-percentage-point increase in the average tax on reported income, from

0.152 to 0.162. The NMP on income not subject to information reporting increases by 3.8 percentage points, while the NMP on total income increases by just 1.3 percentage points. Figure 4 shows that this increase is more or less uniform across income brackets. The average tax rate on true income increases by 0.6 percentage points. Overall government revenues experience a 3.7% increase: this is not really surprising, since labor income essentially cannot escape the higher taxation, while noncompliance rates for capital income do not react enough, in this case, to counterbalance the increase in revenues coming from labor income taxation. This results contrast sharply with the predictions of the static deterrence model of tax evasion, as effectively summarized in Sandmo (2005): if the fine is imposed on the evaded tax, as in Yitzhaki (1974), the standard static model predicts a *negative* relationship between the tax rate and the amount of evasion. The economic intuition behind this result is straightforward: an increase in the tax rate rises the marginal incentive to evade taxes, but also its marginal cost; this kills the "substitution effect," leaving the "income effect" unaffected. In our model, an increase in the tax rate does not affect the two marginal effects proportionally, since the cost of evasion is eventually paid in the future and therefore has to be discounted. For the same reason, the "income effect," which in the static model depends on the degree of risk aversion, is quantitatively less relevant in our framework. In order to evaluate the overall welfare cost of this increase in taxation, we compute, and report in Table 3 as Consumption Equivalent Variation (CEV), the uniform percentage increase in consumption (for fixed non-compliance level) at each state of the world needed to make a household indifferent between being born into the steady state associated with higher overall taxation and being born into the benchmark steady state. A positive CEV reflects a welfare increase compared to the benchmark parameterization. As we can see, an overall increase in taxation will clearly decrease welfare, since it implies a negative CEV of -0.64%.

4.2.2 From progressive to proportional taxation

The second experiment, reported in column (*iii*), analyses the effect of a radical shift in fiscal policy: a switch from a progressive to a perfectly proportional tax system. This is achieved by setting the proportionality factor in the tax function to $a_0 = 0.152$ and the progressivity factor to $a_1 = 0$: in this way the average tax rate paid by the households on their reported income remains unchanged. The quantitative implications of this experiment are significant: the capital stock rises substantially in steady state, from 8.5 to 9.3, and so does the GDP level, even if the increase is less pronounced. More importantly, the level of evaded taxes shrinks considerably, by more than 42% (from 0.085 to 0.049). The NMP on income not subject to information reporting decreases from 48.7% to 35.9%, a remarkable 12.8 percentage points drop. Similarly, the overall NMP and the tax evasion rate fall both to 13.2%, being the two concepts identical under proportional taxation. The overall government revenues rise from 0.295 to 0.323, a 9.5% increase. Not

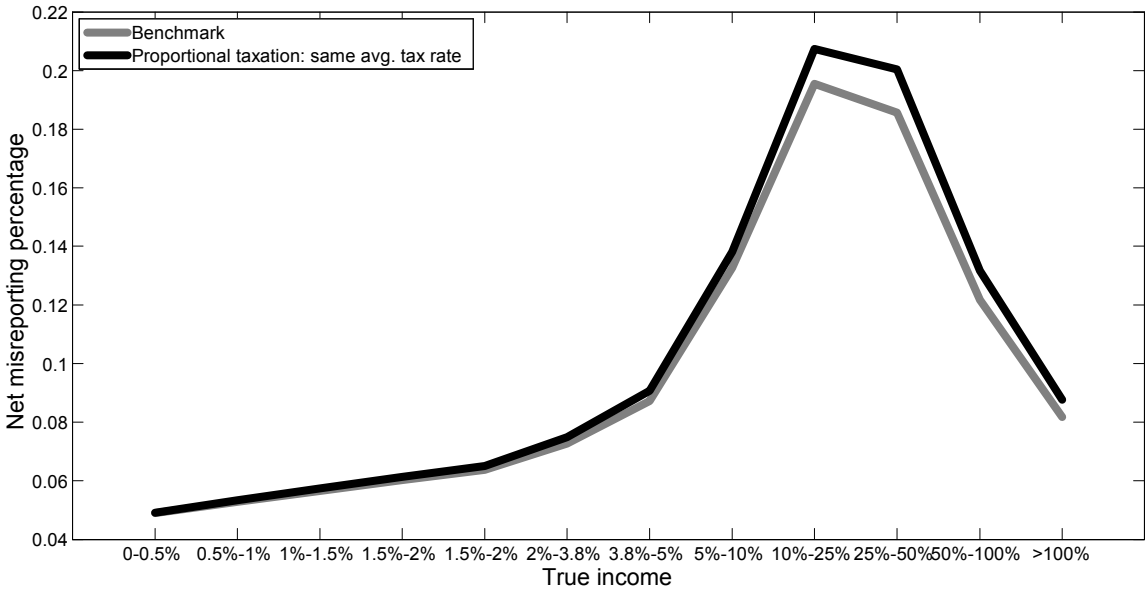

Figure 4: Distribution of NMP by true income: effect of overall higher taxation.

surprisingly, the cost to be paid for this policy shift is a clear increase in income and wealth inequality: the standard deviation of physical capital rises from 6.98 to 8.93, and the Gini coefficients for after-tax income and wealth increase respectively from 0.34 to 0.45 and from 0.44 to 0.50. However, switching from a progressive to a proportional tax system is clearly beneficial in terms of welfare: the CEV is 1.64%, and this implies that consumption under the benchmark parametrization has to be uniformly increased in order to make the individuals indifferent between the two tax systems.

The third experiment, reported in column (iv), is similar, but not identical: it analyses a *revenue-neutral* switch from a progressive to a proportional tax system. This implies that now the tax schedule has to be tilted, as in the previous experiment, and also shifted downwards: hence, the results reflect the changes in both the slope and the intercept of the tax function. We achieved this by setting the proportionality factor in the tax function to $a_0 = 0.134$ and, as before, the progressivity factor to $a_1 = 0$. The capital stock in steady state rises more than in the previous experiment, from the 8.5 obtained under the benchmark parameterization to 9.7. The amount of evaded taxes shrinks even more significantly, by more than 59%. The NMP on income not subject to information reporting drops by 16.9 percentage points, the one on total income by 6.3 points. The average tax rate on true income decreases by 0.7 percentage points, while the average tax rate on reported income by 1.8 points. The effects on inequality are essentially the same as in the previous experiment. As we can see, the revenue-neutral switch from progressive to proportional taxation has a larger effect on misreporting rates, and allows for a non-negligible reduction in the average tax rate. Evidently, our revenue-neutral tax reform is more beneficial in terms of welfare: the CEV reaches 2.3%, implying a further

0.7 percentage point increase in consumption to make the households indifferent between the two tax regimes.

This strong reaction of tax evasion to a switch from a progressive to a proportional tax system seems to be in line with the empirical findings of Ivanova et al. (2005) and Gorodnichenko et al. (2009), who studied Russia's 2001 flat rate income tax reform and found large and significant effects of the reform on tax evasion and voluntary compliance. In particular, Gorodnichenko et al. (2009) estimate a relatively large tax evasion response of households to changes in tax rates, a $10 - 11\%$ increase in reported income relative to consumption: this allows them to conclude that the adoption of a flat rate income tax can lead to significant reductions in tax evasion and to increased tax revenues, due to better reporting and increased compliance, in countries in which both tax rates and misreporting rates are high. Our approach is different, and so is our focus, but nonetheless we find somehow reassuring that our results are broadly compatible with their empirical evidence, at least in terms of sign and order of magnitude of the effects.

Figure 4 compares the distribution of NMPs under our proportional taxation scheme with the one obtained under our benchmark parameterization. A switch from a progressive to a proportional tax system essentially implies an increase in the average tax rate faced by low-income households, and a specular increase in the rate faced by high-income ones. Low-income households rely typically more on labor income than high-income ones: hence, low-income households can hardly increase their misreporting rate as their average tax rate increases, while high-income households can easily reduce it as their average tax rate decreases. As a result, the misreporting rates for high-income households drop significantly, while the ones for low-income households do not increase significantly, and therefore the overall average misreporting rate decreases sharply.

4.2.3 A flat-tax reform

Column (v), finally, reports the results for our last experiment: we tilt the tax schedule, setting $a_0 = 0.152$ and $a_1 = 0$, and introduce a fixed exemption equal to $a_3 = 0.189$, so that the tax schedule reduces to $\mathcal{T}(y) = a_0 \max(y - a_3, 0)$. Given that the exemption turns out to be lower than labor income in all states of the world, the previous tax structure is similar in spirit to a Hall and Rabushka (1995) flat-tax scheme characterized by an exemption for labor income and a common proportional tax rate on the remaining individual income. The results reported in column (iii) show that switching from a progressive to a purely proportional tax system, while leaving the average tax rate unaffected, reduces the misreporting rate, improves the social welfare, and significantly increases the overall tax revenues. If this increase in tax revenues is compensated for via a contemporaneous decrease in the average tax rate, as in the experiment reported in column (iv), then the reaction of the misreporting rate is amplified, and social welfare increases even more. If, instead, the increase in tax revenues is compensated for via an exemption on labor in-

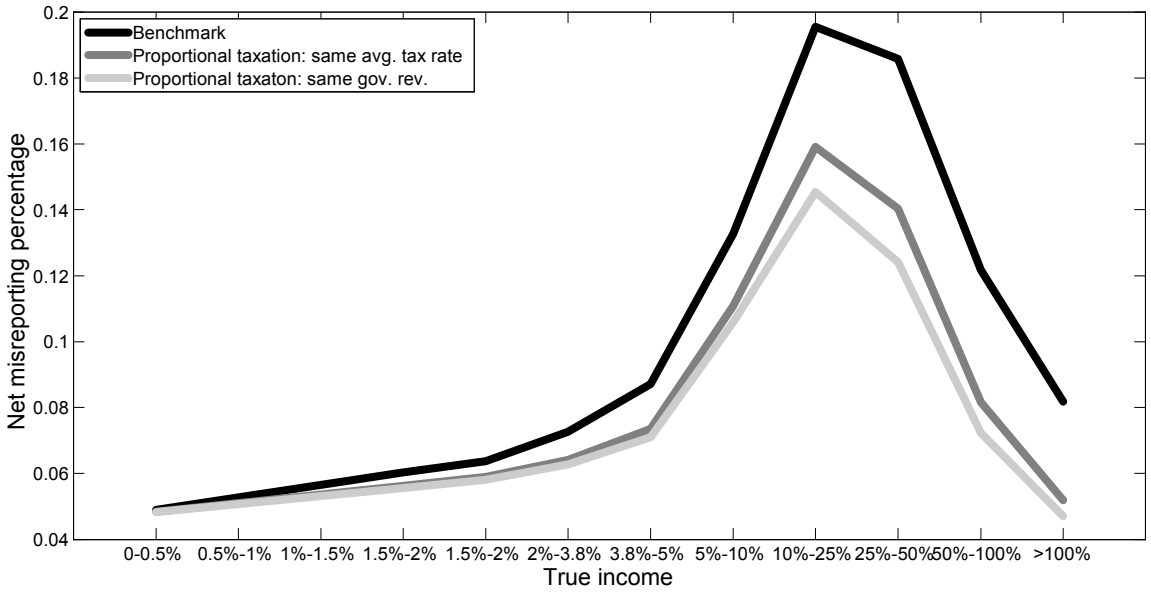

Figure 5: Distribution of NMP by true income: effects of a switch to proportional taxation.

come, then the results reported in column (v) show that neither the misreporting rate nor the social welfare level are significantly affected by the compensation itself. The intuition is actually straightforward: a fixed exemption that is consistently below labor income, which is exogenous from an individual point of view, will hardly affect the misreporting behavior of our households, since the poorer ones have already a limited ability to evade taxes on labor income, while the richer ones will essentially remain unaffected.

5 Conclusions

The available empirical evidence suggests that the level of tax evasion is strictly related to the distribution of income and its composition. We build a model that endogenizes both dimensions, calibrate it to U.S. data, and evaluate its ability to replicate their qualitative and quantitative features. Our results suggest that the model successfully captures the main properties of the estimated distribution of misreporting rates over true income levels. A policy experiment shows that moving from a progressive to a proportional tax system has significant quantitative implications: it would sharply decrease the amount of tax evasion, increase government revenues and increase income inequality; these implications seem in line with some recent empirical evidence.

Our framework is admittedly the simplest possible one able to fulfill our needs, and one can think of many extensions that may prove useful. For instance, the exogenous and fixed probability of being subject to an audit seems somehow in contrast with the evidence that the IRS is targeting relatively more higher-income households: endogenizing the auditing rate may be a way to solve this problem. Furthermore, introducing expected utility with

rank dependent probabilities may allow us to reduce the importance of the utility cost of tax evasion. We leave these extensions, and possibly others, to future research.

References

Aiyagari, S. R. (1994, August). Uninsured Idiosyncratic Risk and Aggregate Saving. *The Quarterly Journal of Economics 109*(3), 659–84.

Allingham, M. G. and A. Sandmo (1972, November). Income tax evasion: a theoretical analysis. *Journal of Public Economics 1*(3-4), 323–38.

Andreoni, J. (1992, October). IRS as loan shark: tax compliance with borrowing constraints. *Journal of Public Economics 49*(1), 35–46.

Andreoni, J., B. Erard, and J. Feinstein (1998, June). Tax Compliance. *Journal of Economic Literature 36*(2), 818–60.

Bewley, T. F. (1980). The Optimum Quantity of Money. In J. H. Kareken and N. Wallace (Eds.), *Models of Monetary Economies*, pp. 169–210. Federal Reserve Bank of Minneapolis.

Bloomquist, K. M. (2003). Income Inequality and Tax Evasion: A Synthesis. *OECD Papers 3*(10), 1–25.

Bloomquist, K. M. (2006, Winter). A Comparison of Agent-Based Models of Income Tax Evasion. *Social Science Computer Review 24*(4), 411–25.

Bordignon, M. (1993, October). A fairness approach to income tax evasion. *Journal of Public Economics 52*(3), 345 – 62.

Carroll, C. D. (1997, February). Buffer-Stock Saving and the Life Cycle/Permanent Income Hypothesis. *The Quarterly Journal of Economics 112*(1), 1–55.

Chen, B.-L. (2003, April). Tax Evasion in a Model of Endogenous Growth. *Review of Economic Dynamics 6*(2), 381–403.

Conesa, J. C. and D. Krueger (2006, October). On the optimal progressivity of the income tax code. *Journal of Monetary Economics 53*(7), 1425–50.

Cooley, T. F. and E. C. Prescott (1995, February). Economic Growth and Business Cycles. In T. F. Cooley (Ed.), *Frontiers of Business Cycle Research*, pp. 1–38. Princeton Univ. Press.

Cowell, F. A. (2003, March). Sticks and Carrots. STICERD - Distributional Analysis Research Programme Papers 68, Suntory and Toyota International Centres for Economics and Related Disciplines, LSE. http://sticerd.lse.ac.uk/dps/darp/darp68.pdf.

Feinstein, J. S. (1991, Spring). An Econometric Analysis of Income Tax Evasion and its Detection. *RAND Journal of Economics 22*(1), 14–35.

Feldman, N. E. and J. Slemrod (2007, March). Estimating tax noncompliance with evidence from unaudited tax returns. *The Economic Journal 117*(518), 327–52.

Fiorio, C. V. and F. D'Amuri (2005, November). Workers' Tax Evasion in Italy. *Giornale degli Economisti 64*(2-3), 247–70.

Floden, M. and J. Lindé (2001, July). Idiosyncratic Risk in the United States and Sweden: Is There a Role for Government Insurance? *Review of Economic Dynamics 4*(2), 406–37.

Gordon, J. P. P. (1989, April). Individual morality and reputation costs as deterrents to tax evasion. *European Economic Review 33*(4), 797 – 805.

Gorodnichenko, Y., J. Martinez-Vazquez, and K. S. Peter (2009, June). Myth and Reality of Flat Tax Reform: Micro Estimates of Tax Evasion Response and Welfare Effects in Russia. *Journal of Political Economy 117*(3), 504–54.

Gouveia, M. and R. P. Strauss (1994, June). Effective Federal Individual Income Tax Functions: An Exploratory Empirical Analysis. *National Tax Journal 47*(2), 317–39.

Greenwood, J., Z. Hercowitz, and G. W. Huffman (1988, June). Investment, Capacity Utilization, and the Real Business Cycle. *American Economic Review 78*(3), 402–17.

Guvenen, F. (2009, January). An Empirical Investigation of Labor Income Processes. *Review of Economic Dynamics 12*(1), 58–79.

Hall, R. E. and A. Rabushka (1995). *The Flat Tax, Second Edition*. Stanford: Hoover Press.

Heathcote, J., K. Storesletten, and G. L. Violante (2009, September). Quantitative Macroeconomics with Heterogeneous Households. *Annual Review of Economics 1*, 319–54.

Huggett, M. (1993, September). The risk-free rate in heterogeneous-agent incomplete-insurance economies. *Journal of Economic Dynamics and Control 17*(5-6), 953–69.

Huggett, M. (1997, August). The one-sector growth model with idiosyncratic shocks: Steady states and dynamics. *Journal of Monetary Economics 39*(3), 385–403.

Ivanova, A., M. Keen, and A. Klemm (2005, July). The Russian 'flat tax' reform. *Economic Policy 20*(43), 397–444.

Johns, A. and J. Slemrod (2008, September). The Distribution of Income Tax Noncompliance. Manuscript, http://www.bus.umich.edu/otpr/DITN%20091308.pdf.

Klepper, S. and D. Nagin (1989, Spring). The Anatomy of Tax Evasion. *Journal of Law, Economics and Organization 5*(1), 1–24.

Kleven, H. J., M. B. Knudsen, C. T. Kreiner, S. Pedersen, and E. Saez (2010, February). Unwilling or Unable to Cheat? Evidence from a Randomized Tax Audit Experiment in Denmark. Technical report, NBER Working Paper 15769.

Kopecky, K. and R. Suen (2010, July). Finite State Markov-chain Approximations to Highly Persistent Processes. *Review of Economic Dynamics 13*(3), 701–14.

Matsaganis, M., D. Benedek, M. Flevotomou, O. Lelkes, D. Mantovani, and S. Nienadowska (2010, February). Distributional implications of income tax evasion in Greece, Hungary and Italy. MPRA Paper 21465, University Library of Munich, Germany. http://ideas.repec.org/p/pra/mprapa/21465.html.

Niepelt, D. (2005, September). Timing tax evasion. *Journal of Public Economics 89*(9-10), 1611–37.

Persson, M. and P. Wissén (1984). Redistributional Aspects of Tax Evasion. *The Scandinavian Journal of Economics 86*(2), 131.

Pissarides, C. A. and G. Weber (1989, June). An expenditure-based estimate of Britain's black economy. *Journal of Public Economics 39*(1), 17–32.

Rendhal, P. (2007). Inequality Constraints in Recursive Economies. Technical Report 2006/06, European University Institute. http://sites.google.com/site/pontusrendahl/Inequalityconstr.pdf?attredirects=0.

Sandmo, A. (2005, December). The Theory of Tax Evasion: A Retrospective View. *National Tax Journal LVIII*(4), 643–63.

Slemrod, J. (2007, Winter). Cheating Ourselves: The Economics of Tax Evasion. *Journal of Economic Perspectives 21*(1), 25–48.

Slemrod, J. and S. Yitzhaki (2002). Tax avoidance, evasion, and administration. In A. J. Auerbach and M. Feldstein (Eds.), *Handbook of Public Economics*, Volume 3, Chapter 22, pp. 1423–70. Elsevier.

Webber, C. and A. B. Wildavsky (1986, August). *A history of taxation and expenditure in the Western world*. Simon and Schuster, New York.

Yitzhaki, S. (1974, May). Income tax evasion: A theoretical analysis. *Journal of Public Economics 3*(2), 201–2.

Young, E. R. (2010, January). Solving the incomplete markets model with aggregate uncertainty using the Krusell-Smith algorithm and non-stochastic simulations. *Journal of Economic Dynamics and Control 34*(1), 36–41.

A Appendix: the recursive equilibrium

The vector of individual state variables $s_t \equiv \{k_t, e_t, \sigma_t\}$ lies in $X = [0,\infty) \times [0,\infty) \times (E \times L)$. The distribution of individual states across agents is described by an aggregate state, the probability measure λ_t. More precisely, λ_t is the unconditional probability distribution of the state vector $\{k_t, e_t, \sigma_t\}$, defined over the Borel subset of X:

$$\lambda_t(\mathrm{k},\mathrm{e},\mathrm{z}) = \lambda_t(\mathrm{s}) = \mathrm{prob}(k_t = \mathrm{k}, e_t = \mathrm{e}, \sigma_t = \mathrm{z}). \tag{16}$$

For the *Law of Large Numbers*, $\lambda_t(\mathrm{s})$ can be interpreted as the mass of agents whose individual state vector is equal to s. Being λ_t a probability measure, the total mass of agents is equal to one.

In a recursive equilibrium, the time-invariant individual policy functions will depend on the exogenous state, σ, on the beginning of period capital stock, k, on the amount of past unreported income, e, and on the aggregate distribution λ. The aggregate prices w_t and r_t will depend on the distribution of individual wealth stocks. Hence, the exogenous Markov process for σ and the optimal policy functions $c(s;\lambda)$ and $d(s;\lambda)$ induce a law of motion for the distribution λ_t:

$$\lambda_{t+1}(\mathrm{s}) = \int\int \sum_{j=1}^{4} \mathcal{I}(\mathrm{k},\mathrm{e},k,e,\sigma_j)\,\pi(\mathrm{z},\sigma_j)\,\lambda_t(k,e,\sigma_j)\,dkde = \int_X \mathcal{I}(\mathrm{k},\mathrm{e},k,e,\sigma)\,\pi(\mathrm{z},\sigma)\,d\lambda_t, \tag{17}$$

where:

$$\mathcal{I}(\mathrm{k},\mathrm{e},k,e,\sigma) = \begin{cases} 1 & \text{if } k'(s;\lambda_t) = \mathrm{k} \text{ and } e'(s;\lambda_t) = \mathrm{e} \\ 0 & \text{otherwise} \end{cases}. \tag{18}$$

Given the absence of aggregate uncertainty, in the long run the economy will reach a stationary equilibrium, i.e. steady state characterized by a constant aggregate capital stock.

Definition 1 *A stationary recursive equilibrium is a couple of policy functions $c(s;\lambda)$ and $\theta(s;\lambda)$, a couple of values $\{w, i\}$, and a probability distribution λ such that:*

1. *The policy functions $c(s;\lambda)$ and $\theta(s;\lambda)$ solve the individual optimization problem (4).*
2. *The factor prices $\{w, r\}$, together with $K = \int_X k d\lambda$ and $N = \int_X l d\lambda$, satisfy the first order conditions for the firm.*
3. *The market for the final good clears:*

$$C + K' = (1-\delta)K + \phi K^{\alpha} N^{1-\alpha}.$$

4. The distribution satisfies the induced law of motion:

$$\lambda(\mathrm{s}) = \int_X \mathcal{I}(\mathrm{k}, \mathrm{e}, k, e, \sigma)\, \pi(\mathrm{z}, \sigma)\, d\lambda, \quad \forall \mathrm{s} \in X.$$

B Appendix: the solution algorithm

B.1 Solving for a stationary equilibrium

The iterative solution method used to compute the stationary equilibrium adapts to our needs the standard approach outlined in Aiyagari (1994) and Huggett (1997):

Algorithm 2 *Choose an initial guess for G, say $G_0 > 0$. Then, for $z \geq 0$:*

1. *Choose an initial guess for K_z, say $K_{z0} > 0$. Then, for $j \geq 0$:*
   (a) *Given K_{zj}, compute w_{zj} and r_{zj} from (12) and (13).*
   (b) *Solve the household problem for the individual policy functions.*
   (c) *Compute the implied stationary distribution $\lambda_{zj}(s)$.*
   (d) *Compute the implied aggregate capital stock:*

$$\hat{K}_{zj} = \int_X k'_z d\lambda_{zj}.$$

   (e) *Given $\hat{K}_z$, compute a new estimate of K_z:*

$$K_{zj+1} = \upsilon \hat{K}_{zj} + (1 - \upsilon) K_{zj}$$

   where $\upsilon \in (0, 1)$is a relaxation parameter.
   (f) *Iterate $(a) - (f)$ until convergence.*
2. *Compute the amount of implied government transfers per capita:*

$$G_{z+1} = \int_X \left[\mathcal{T}_y(y - z) + \mu\varepsilon e\right] d\lambda_z.$$

3. *Iterate on $(1) - (2)$ until convergence.*

From a practical point of view, the fixed point problem described in the previous Algorithm can be efficiently solved using bisection, or any other univariate solution algorithm like Ridder's or Brent's ones. The next Sections will describe the methods used for points (2) and (3) in the above algorithm more in detail.

B.2 Solving for the individual policy function

Our stochastic dynamic optimization problem with endogenous occasionally binding constraints is solved using the time iteration algorithm described in Rendhal (2007):

Algorithm 3 *Choose suitable univariate grids for the individual capital stock and the amount of past evaded taxes on* R_+*, say* $\hat{\mathbf{k}} = \{k_i\}_{i=1}^{h_k}$ *and* $\hat{\mathbf{e}} = \{e_i\}_{i=1}^{h_e}$*, with* $k_1 = 0 < k_2 < ... < k_{h_k} = k_{\max}$*, and* $e_1 = 0 < e_2 < ... < e_{h_e} = e_{\max}$.[22] *Define matrices* $\mathbf{k} \equiv \hat{\mathbf{k}}^T \otimes \mathbf{1}_{h_k}$ *and* $\mathbf{e} \equiv \hat{\mathbf{e}} \otimes \mathbf{1}_{h_e}^T$*. Choose initial guesses for* c*,* θ *and* $\varkappa$ *at each grid point, i.e. matrices* $\mathbf{c}_{z,0}$*,* $\boldsymbol{\theta}_{z,0}$*, and* $\boldsymbol{\varkappa}_{z,0}$*,where* z *denotes the exogenous state. Compute* $\mathbf{y}_z = r\mathbf{k} + wl_z$ *and* $\bar{\mathbf{y}}_z = \varsigma_K r\mathbf{k} + \varsigma_L w l_z$*. Then, for* $j \geq 0$*:*

1. *Given the current guesses* $\mathbf{c}_{z,j}$ *and* $\boldsymbol{\theta}_{z,j}$*, compute:*

$$\begin{aligned}
\mathbf{z}_{z,j} &= \boldsymbol{\theta}_{z,j}\bar{\mathbf{y}}_z, \\
\mathbf{d}_{z,j} &= \mathbf{y}_z - \mathbf{z}_{z,j}, \\
\boldsymbol{\xi}_{z,j} &= \left(\mathbf{c}_{z,j} - \frac{\omega}{\nu}\mathbf{z}_{z,j}^{\nu}\right)^{-\gamma}, \\
\boldsymbol{\psi}_{z,j} &= \boldsymbol{\xi}_{z,j}\left[\mathcal{T}_y\left(\mathbf{d}_{z,j}\right) - \omega\mathbf{z}_{z,j}^{\nu-1}\right],
\end{aligned}$$

   and:

$$\begin{aligned}
\mathbf{k}'_{z,j} &= \left(1-\delta_K\right)\mathbf{k} + \mathbf{y}_z - \mathcal{T}\left(\mathbf{d}_{z,j}\right) - \varepsilon_z\mu\mathbf{e} + G - \mathbf{c}_{z,j}, \\
\mathbf{e}'_{z,j} &= \left(1-\varepsilon_z\right)\left(1-\delta_E\right)\mathbf{e} + \mathcal{T}\left(\mathbf{y}_z\right) - \mathcal{T}\left(\mathbf{d}_{z,j}\right).
\end{aligned}$$

2. *Compute the future policy variables* $\mathbf{c}'_{q,z,j}$*,* $\boldsymbol{\theta}'_{q,z,j}$*, and* $\boldsymbol{\varkappa}'_{q,z,j}$*, via bivariate cubic interpolation (or extrapolation, if needed) on* $\mathbf{k}$*,* $\mathbf{e}$*,* $\mathbf{c}_{z,j}$*,* $\boldsymbol{\theta}_{z,j}$*, and* $\boldsymbol{\varkappa}_{z,j}$*.*

3. *Compute* $\bar{\boldsymbol{\psi}}_{z,j}$ *as:*[23]

$$\bar{\boldsymbol{\psi}}_{z,j} = \beta\sum_{q=1}^{2}\pi\left(\sigma_q,\sigma_z\right)\left(1-\delta_E\right)\left(\boldsymbol{\psi}'_{q,z,j} - \boldsymbol{\varkappa}'_{q,z,j}\right)\frac{\mathcal{T}_y\left(\mathbf{d}_{z,j}\right)}{\mathcal{T}_y\left(\mathbf{d}'_{q,z,j}\right)} + \beta\sum_{q=3}^{4}\pi\left(\sigma_q,\sigma_z\right)\boldsymbol{\xi}'_{q,z,j}\mu\mathcal{T}_y\left(\mathbf{d}_{z,j}\right),$$

[22] The grid has not necessarily to be uniformly distributed. Given that the policy function is particularly non-linear near the point where the credit constraint starts to bite, one may concentrate more nodes near that region.

[23] With a slight abuse of notation, let $\mathbf{c}_{\varepsilon,j}^{-\gamma}$ represent the matrix obtained by raising each element of $\mathbf{c}_{\varepsilon,j}$ to the power $-\gamma$.

where:

$$\begin{aligned}
\mathbf{z}'_{q,z,j} &= \boldsymbol{\theta}'_{q,z,j}\left(\varsigma_K r\mathbf{k}'_{z,j} + \varsigma_K w l_q\right), \\
\mathbf{d}'_{q,z,j} &= \mathbf{y}'_{q,z,j} - \mathbf{z}'_{q,z,j}, \\
\boldsymbol{\xi}'_{q,z,j} &= \left[\mathbf{c}'_{q,z,j} - \frac{\omega}{\nu}\left(\mathbf{z}'_{q,z,j}\right)^{\nu}\right]^{-\gamma}, \\
\boldsymbol{\psi}'_{q,z,j} &= \boldsymbol{\xi}'_{q,z,j}\left[\mathcal{T}_y\left(\mathbf{d}'_{q,z,j}\right) - \omega\left(\mathbf{z}'_{q,z,j}\right)^{\nu-1}\right].
\end{aligned}$$

4. *Compute* $\bar{\mathbf{d}}_{z,j}$ *as:*[24]

$$\hat{\boldsymbol{\theta}}_{z,j} = \frac{1}{\bar{\mathbf{y}}_z}\left\{\frac{1}{\omega}\left[\mathcal{T}_y\left(\mathbf{d}_{z,j}\right) - \frac{\bar{\boldsymbol{\psi}}_{z,j}}{\boldsymbol{\xi}_{z,j}}\right]\right\}^{\frac{1}{\nu-1}},$$

and $\hat{\boldsymbol{\theta}}_{z,j}$ *as:*

$$\hat{\boldsymbol{\theta}}_{z,j} = \min\left[\max\left(\bar{\boldsymbol{\theta}}_{z,j}, 0\right), 1\right].$$

5. *Compute* $\hat{\varkappa}_{z,j} = \hat{\boldsymbol{\psi}}_{z,j} - \bar{\boldsymbol{\psi}}_{z,j}$*, where:*

$$\begin{aligned}
\hat{\boldsymbol{\psi}}_{z,j} &= \left(\mathbf{c}_{z,j} - \frac{\omega}{\nu}\hat{\mathbf{z}}^{\nu}_{z,j}\right)^{-\gamma}\left[\mathcal{T}_y\left(\hat{\mathbf{d}}_{z,j}\right) - \omega\hat{\mathbf{z}}^{\nu-1}_{z,j}\right], \\
\hat{\mathbf{z}}_{z,j} &= \hat{\boldsymbol{\theta}}_{z,j}\bar{\mathbf{y}}_z, \\
\hat{\mathbf{d}}_{z,j} &= \mathbf{y}_z - \hat{\mathbf{z}}_{z,j}.
\end{aligned}$$

6. *Compute* $\bar{\mathbf{c}}_{z,j}$ *as:*

$$\bar{\mathbf{c}}_{z,j} = \frac{\omega}{\nu}\mathbf{z}^{\nu}_{z,j} + \left(\boldsymbol{\Lambda}'_{q,z,j}\right)^{-\frac{1}{\gamma}},$$

where:

$$\begin{aligned}
\boldsymbol{\Lambda}'_{q,z,j} &= \sum_{q=1}^{4}\pi\left(\sigma_q, \sigma_z\right)\beta\left(\boldsymbol{\xi}'_{q,z,j}\left\{1 - \delta_K + \left[1 - \mathcal{T}_y\left(\mathbf{d}'_{q,z,j}\right)\right]r\right\} + \right. \\
&\quad \left.\left[\varkappa'_{q,z,j}\boldsymbol{\theta}'_{q,z,j}\varsigma_K + \left(\boldsymbol{\psi}'_{q,z,j} - \varkappa'_{q,z,j}\right)F'_{q,z,j}\right]r\right), \\
F'_{q,z,j} &= 1 - \frac{\mathcal{T}_y\left(\mathbf{y}'_{q,z,j}\right)}{\mathcal{T}_y\left(\mathbf{d}'_{q,z,j}\right)},
\end{aligned}$$

and $\hat{\mathbf{c}}_{z,j}$ *as:*

$$\hat{\mathbf{c}}_{z,j} = \min\left(\bar{\mathbf{c}}_{z,j}, \left(1 - \delta_K\right)\mathbf{k} + \mathbf{y}_z - \mathcal{T}\left(\mathbf{d}_{z,j}\right) - \varepsilon_z\mu\mathbf{e} + G\right).$$

[24] Or, if $\omega = 0$ or $\nu = 1$, as:

$$\bar{\boldsymbol{\theta}}_{z,j} = \frac{\mathbf{k}'_{z,j} - \left(1 - \delta_K\right)\mathbf{k} - \mathbf{y}_z + \mathcal{T}\left(\mathbf{d}_{z,j}\right) + \varepsilon_z\mu\mathbf{e} - G + \frac{\omega}{\nu}\mathbf{z}^{\nu}_{z,j} + \left(\frac{\bar{\boldsymbol{\psi}}_{z,j}}{\mathcal{T}_y\left(\mathbf{d}_{z,j}\right) - \omega\mathbf{z}^{\nu-1}_{z,j}}\right)^{-\frac{1}{\gamma}}}{\tau\bar{\mathbf{y}}_z} \tag{19}$$

7. Update the guesses for $\mathbf{c}_{z,j}$, $\mathbf{d}_{z,j}$, *and* $\varkappa_{z,j}$, *as follows:*

$$\begin{aligned}
\mathbf{c}_{z,j+1} &= \upsilon\hat{\mathbf{c}}_{z,j} + (1-\upsilon)\,\mathbf{c}_{z,j}, \\
\mathbf{d}_{z,j+1} &= \upsilon\hat{\mathbf{d}}_{z,j} + (1-\upsilon)\,\mathbf{d}_{z,j}, \\
\varkappa_{z,j+1} &= \upsilon\hat{\varkappa}_{z,j} + (1-\upsilon)\,\varkappa_{z,j},
\end{aligned}$$

where $\upsilon \in (0,1)$, *and iterate on points* $(1)-(7)$ *until convergence.*

B.3 Computing the stationary distribution

Building on Young (2010), we compute the stationary distribution using a non-stochastic "binning" approach. This allows us to avoid the small sample bias that plagues more traditional simulation methods.

The distribution is approximated with an histogram over a fixed and uniformly distributed grid on $[0, k_{\max}] \times [0, e_{\max}] \times E$, say $\{k_i\}_{i=1}^{m_k} \times \{e_i\}_{i=1}^{m_e} \times E$, with $k_1 = 0$, $e_1 = 0$, $k_m = k_{\max}$, and $e_m = e_{\max}$. The histogram can be summarized by a $(m_k \times m_e \times m_z)$ 3-dimensional array $\boldsymbol{\lambda}_t$, whose element $\lambda_t(i,j,z)$ represents the share of households with wealth i, evasion stock j, and exogenous state z at the beginning of period t. This implies that the aggregate capital stock can be approximated by:

$$K'_t \approx \sum_{i=1}^{m_a}\sum_{j=1}^{m_e}\sum_{z=1}^{m_z} \mathsf{k}'_z(i,j)\,\lambda_t(i,j,z), \tag{20}$$

where $\mathsf{k}'_z(i,j)$ can be obtained via interpolation.

Suppose that a strictly positive mass of households, say υ, saves an amount k' such that $k_i \leq k' \leq k_{i+1}$ for some $i \in \{1,2,...,m_k\}$ and underreports an amount e' such that $e_j \leq e' \leq e_{j+1}$ for some $j \in \{1,2,...,m_e\}$. The key step in our discrete approximation is to allocate the mass υ to the nodes $\{k_i, e_j\}$, $\{k_{i+1}, e_j\}$, $\{k_i, e_{j+1}\}$, and $\{k_{i+1}, e_{j+1}\}$ in such a way that the aggregate variables remain unaffected. If $\omega_{i,j}$ denotes the share of households that end up at node $\{k_z, e_j\}$, then the previous requirement boils down to the following constraints:

$$\omega_{i,j} + \omega_{i,j+1} = \frac{k_{i+1} - k'}{k_{i+1} - k_i}, \tag{21}$$

$$\omega_{i,j} + \omega_{i+1,j} = \frac{e_{j+1} - e'}{e_{j+1} - e_j}. \tag{22}$$

One way to achieve the result is to set:

$$\omega_{i,j} = \frac{(k_{i+1}-k')(e_{j+1}-e')}{(k_{i+1}-k_i)(e_{j+1}-e_j)}, \tag{23}$$

$$\omega_{i,j+1} = \frac{(k_{i+1}-k')(e'-e_j)}{(k_{i+1}-k_i)(e_{j+1}-e_j)}, \tag{24}$$

$$\omega_{i+1,j} = \frac{(k'-k_i)(e_{j+1}-e')}{(k_{i+1}-k_i)(e_{j+1}-e_j)}, \tag{25}$$

$$\omega_{i+1,j+1} = 1-\omega_{i,j}-\omega_{i,j+1}-\omega_{i+1,j}. \tag{26}$$

Hence, the mass υ is distributed according to the following rule:[25]

$$\omega(i,j,k',e') = \begin{cases} \frac{(k'-k_i)(e'-e_j)}{(k_i-k_{i-1})(e_j-e_{j-1})} & \text{if } k' \in [k_{i-1},k_i] \text{ and } e' \in [e_{j-1},e_j] \\ \frac{(k'-k_i)(e_{j+1}-e')}{(k_i-k_{i-1})(e_{j+1}-e_j)} & \text{if } k' \in [k_{i-1},k_i] \text{ and } e' \in (e_j,e_{j+1}] \\ \frac{(k_{i+1}-k')(e_j-e')}{(k_{i+1}-k_i)(e_j-e_{j-1})} & \text{if } k' \in (k_i,k_{i+1}] \text{ and } e' \in [e_{j-1},e_j] \\ \frac{(k_{i+1}-k')(e_{j+1}-e')}{(k_{i+1}-k_i)(e_{j+1}-e_j)} & \text{if } k' \in (k_i,k_{i+1}] \text{ and } e' \in (e_j,e_{j+1}] \\ 0 & \text{otherwise} \end{cases}. \tag{27}$$

Note that $\omega(i,j,k',e') \geq 0$, and $\omega(i,j,k',e') > 0$ for at most four nodes $\{i,j\}$.

The law of motion for the wealth distribution described in (17) boils down to the following relationship:

$$\begin{aligned} \lambda_{t+1}(u,q,l) = \sum_{i=1}^{m_k}\sum_{j=1}^{m_e}\sum_{z=1}^{m_z} \omega\left[u,q,\mathsf{k}'_z(k_i,e_j),\mathsf{e}'_z(k_i,e_j)\right]\pi(\sigma_l,\sigma_z)\lambda_t(i,j,z) = \\ \pi(\sigma_l,1)\sum_{j=1}^{m_e}\sum_{i=1}^{m_k}\omega\left[u,q,\mathsf{k}'_1(k_i,e_j),\mathsf{e}'_1(k_i,e_j)\right]\lambda_t(i,j,1) + \\ \pi(\sigma_l,2)\sum_{j=1}^{m_e}\sum_{i=1}^{m_k}\omega\left[u,q,\mathsf{k}'_2(k_i,e_j),\mathsf{e}'_2(k_i,e_j)\right]\lambda_t(i,j,2) + \\ \pi(\sigma_l,3)\sum_{j=1}^{m_e}\sum_{i=1}^{m_k}\omega\left[u,q,\mathsf{k}'_2(k_i,e_j),\mathsf{e}'_2(k_i,e_j)\right]\lambda_t(i,j,3) + ... \\ ... + \pi(\sigma_l,m_z)\sum_{j=1}^{m_e}\sum_{i=1}^{m_k}\omega\left[u,q,\mathsf{k}'_2(k_i,e_j),\mathsf{e}'_2(k_i,e_j)\right]\lambda_t(i,j,m_z) \end{aligned} \tag{28}$$

Let us now denote $\mathbf{p}_{j,u,q,z}$ the $m_k \times 1$ column vector whose ith element is:

$$\mathbf{p}_{j,u,q,z}(i) \equiv \omega\left[u,q,\mathsf{k}'_z(k_i,e_j),\mathsf{e}'_z(k_i,e_j)\right]. \tag{29}$$

Furthermore, let us define a $(m_k m_e) \times 1$ vector $\mathbf{p}_{u,q,z}$, a $m_k \times (m_k m_e)$ matrix $\mathbf{P}_{q,z}$, and

[25] Note that the two special cases $z=1$ and $z=m$ have to be taken care of separately: if $z=0$, then $\omega(1,k') = 1-(k'-k_1)/(k_2-k_1)$ if $k_1 \leq k' \leq k_2$ and $\omega(1,k')=0$ otherwise; if $z=0$, then $\omega(m,k') = (k'-k_{m-1})/(k_m-k_{m-1})$ if $k_{m-1} \leq k' \leq k_m$, $\omega(m,k')=1$ if $k'>k_m$, and $\omega(m,k')=0$ otherwise.

a $(m_k m_e) \times (m_k m_e)$ square matrix $\mathbf{P}_z$:

$$\mathbf{p}_{u,q,z} \equiv \begin{bmatrix} \mathbf{p}_{1,u,q,z} \\ \mathbf{p}_{2,u,q,z} \\ \vdots \\ \mathbf{p}_{m_e,u,q,z} \end{bmatrix}, \quad \mathbf{P}_{q,z} \equiv \begin{bmatrix} \mathbf{p}_{1,q,z}^T \\ \mathbf{p}_{2,q,z}^T \\ \vdots \\ \mathbf{p}_{m_k,q,z}^T \end{bmatrix}, \quad \mathbf{P}_z \equiv \begin{bmatrix} \mathbf{P}_{1,z} \\ \mathbf{P}_{2,z} \\ \vdots \\ \mathbf{P}_{m_e,z} \end{bmatrix} \tag{30}$$

Hence, we can rewrite (28) as:

$$\lambda_{t+1}(u,q,l) = \left[\pi(\varepsilon_l,1)\,\mathbf{p}_{u,q,1}^T | \pi(\varepsilon_l,2)\,\mathbf{p}_{u,q,2}^T | ... | \pi(\varepsilon_l,m_z)\,\mathbf{p}_{u,q,m_z}^T\right] \mathrm{vec}(\boldsymbol{\lambda}_t) \tag{31}$$

In matrix notation:

$$\mathrm{vec}(\boldsymbol{\lambda}_{t+1}) = \mathbf{P}\,\mathrm{vec}(\boldsymbol{\lambda}_t), \tag{32}$$

where:

$$\mathbf{P} \equiv (\boldsymbol{\pi} \otimes \mathbf{I}_{m_k m_e}) \begin{bmatrix} \mathbf{P}_1 & 0 & 0 & \cdots & 0 \\ 0 & \mathbf{P}_2 & 0 & \cdots & 0 \\ 0 & 0 & \mathbf{P}_3 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & 0 \\ 0 & 0 & 0 & 0 & \mathbf{P}_{m_z} \end{bmatrix}. \tag{33}$$

Finally, the approximated stationary distribution $\boldsymbol{\lambda}$ can be computed as the ergodic distribution of the Markov chain implied by (32):

$$\mathrm{vec}(\boldsymbol{\lambda}) = \mathbf{P}\,\mathrm{vec}(\boldsymbol{\lambda}). \tag{34}$$

To efficiently compute $\boldsymbol{\lambda}$, define $\hat{\mathbf{A}} \equiv \left(\mathbf{A}^T\mathbf{A}\right)^{-1}\mathbf{A}^T$, where:

$$\underset{(m_k m_e+1)\times(m_k m_e)}{\mathbf{A}} \equiv \begin{bmatrix} \mathbf{I}_{m_k m_e} - \mathbf{P} \\ \mathbf{1}^T \end{bmatrix}, \tag{35}$$

and $\mathbf{1}$ is a $(m_k m_e) \times 1$ vector of ones: the ergodic distribution $\boldsymbol{\lambda}$ corresponds to the $m_k m_e + 1$ column of $\hat{\mathbf{A}}$. Unfortunately, the previously describe strategy is computationally unfeasible if the grids are relatively dense: in this case, we can simply iterate until convergence on:

$$\mathrm{vec}(\boldsymbol{\lambda}_{j+1}) = \mathbf{P}\,\mathrm{vec}(\boldsymbol{\lambda}_j). \tag{36}$$