Review of Economic Dynamics 17 (2014) 559–581

Contents lists available at ScienceDirect

Review of Economic Dynamics

www.elsevier.com/locate/red

Income taxation of U.S. households: Facts and parametric estimates ☆

Nezih Guner [a,*], Remzi Kaygusuz [b], Gustavo Ventura [c]

[a] ICREA-MOVE, Universitat Autonoma de Barcelona and Barcelona GSE, Spain
[b] Faculty of Arts and Social Sciences, Sabanci University, Turkey
[c] Department of Economics, Arizona State University, USA

ARTICLE INFO

Article history:
Received 25 May 2011
Received in revised form 1 January 2014
Available online 22 January 2014

JEL classification:
E62
H24
H31

Keywords:
Taxation
Tax progressivity
Households

ABSTRACT

We use micro data from the U.S. Internal Revenue Service to document how Federal Income tax liabilities vary with income, marital status and the number of dependents. We report facts on the distributions of average taxes, properties of the joint distributions of taxes paid and income, and discuss how taxes are affected by marital status and the number of children. We also provide multiple parametric estimates of tax functions for use in applied work in macroeconomics and public finance.

1. Introduction

This paper has two goals. First, we aim to systematically describe how income taxes paid by a cross section of U.S. households vary according to their income, marital status and number of dependent children. Second, we provide estimates of *effective tax functions* that capture the observed heterogeneity in the data that can be readily used in applied work.

These goals are motivated by the large and growing body of literature that utilizes dynamic macroeconomic models with heterogeneous households; see Heathcote et al. (2009) for a recent survey. This literature has studied how existing models can account for properties of actual earnings, income and wealth distributions, and has used such models to address a host of policy questions.[1] As an input in this process, a large body of work, old and new and from related fields, documented the empirical properties of such distributions. However, the distribution of taxes effectively paid by households and the marginal tax rates that they face have received considerably less attention. This paper fills this void by systematically documenting basic properties of the structure of income taxation for a cross section of U.S. households.

☆ Guner acknowledges support from European Research Council (ERC) Grant 263600. Kaygusuz acknowledges The Scientific and Technological Research Council of Turkey (TUBITAK) Grant 112K075. We thank the editor, Gianluca Violante, an anonymous Associate Editor and two anonymous referees who provided excellent comments. We also thank Ezgi Kaya for excellent research assistance. Earlier versions of this paper circulated under the title “Income Taxation of U.S. Households: Basic Facts”. The usual disclaimer applies.

* Corresponding author.

E-mail addresses: nezih.guner@movebarcelona.eu (N. Guner), rkaygusuz@sabanciuniv.edu (R. Kaygusuz), gustavo.ventura@asu.edu (G. Ventura).

[1] There is a large literature that uses dynamic macroeconomic models with heterogeneous agents to study tax reforms. See Ventura (1999), Altig et al. (2001), Castañeda et al. (2003), Díaz-Giménez and Pijoan-Mas (2005), Nishiyama and Smetters (2005), Conesa and Krueger (2006), Erosa and Koreshkova (2007), Conesa et al. (2009) and Guner et al. (2012a, 2012b), among others. Guvenen et al. (forthcoming) study the effect of taxes on human capital accumulation and inequality.

1094-2025/$ – see front matter
http://dx.doi.org/10.1016/j.red.2014.01.003

The model economies studied in the above mentioned literature require, in accordance with data, a mapping of household's income to taxes paid conceivably depending on the household's marital status, the presence of children and retirement status. This naturally matters when asking how well models with heterogeneous households match distributional properties of data, as well as for the fruitful use of these frameworks to address policy questions. A *ready-to-use*, systematic representation of this mapping is not currently available for different types of households, and we offer it here. Therefore, we provide different parametric estimates of effective taxes as a function of household's income for different types of households; all, married, unmarried, with and without dependent children. We also provide estimates for special cases; with and without positive social security income, with exclusively labor income and with imputations for state and local taxes.

We aim at providing estimates of taxes *effectively* paid by households according to their income, family status and number of dependents. We use a large cross-sectional data set from U.S. Internal Revenue Service ('Public Use Tax File'), that is ideal for these purposes as it is representative of the entire set of U.S. taxpayers. For a notion of effective average tax rates out of Federal Income taxes, we find a substantial degree of heterogeneity generated by the U.S. income tax system and the underlying distribution of income. As we document, average rates increase non-trivially with income, and this is reflected in the distribution of average tax rates paid. For instance, if we rank married households by the average tax rates that they face, average taxes for married (unmarried) households at top 1% are in excess of 27.7% (23.0%), while the median tax rate is about 8.5% (6.1%). These facts, in conjunction with the substantial income dispersion that we document in this data, imply that households at the top of the income distribution account for the bulk of taxes paid; the top 5% accounts for nearly 55.1% of all taxes paid, whereas the top 1% accounts for about 35.7%.

Using this data, we estimate four functional forms for effective tax rates. In each case, we report estimates for all households, as well as for married and unmarried households with different numbers of dependent children. We first estimate a two-parameter specification, which we refer to as the *log* specification (used by Guner et al., 2012a, 2012b). Our second set of estimates, another two-parameter specification, pertains to the functional form in Heathcote et al. (2011), which we refer to as the *HSV* specification. Our third case is a three-parameter specification, which we refer to as the *Power* specification – a version of the power function is used by Guvenen et al. (forthcoming). Finally, we estimate the same functional form used in Gouveia and Strauss (1994), who provided estimates of tax functions for all taxpayers using the U.S. tax structure prevailing in 1989.[2] We find that while all specifications provide tax schedules for average rates that are similar and that track raw data quite well, the Gouveia and Strauss specification generates a better statistical fit than other specifications. We also compare the marginal taxes implied by these parametric specifications with the data. From the data, we compute both effective marginal rates, i.e. changes in tax liabilities associated with changes in household income, and statutory marginal tax rates. We find that effective marginal tax rates are below the statutory ones, especially for high income levels. We also find that the implied schedules of marginal tax rates, computed from the parametric tax functions, track the effective marginal tax rates from data reasonably well. In addition, the schedules become essentially flat after relatively low levels of income under the Gouveia and Strauss specification. The other specifications, in particular the *Power* and *log* specifications, in contrast, generate marginal rates that are closer to the effective marginal tax rates implied by data at high incomes.

The paper is organized as follows. Section 2 presents the basic data that we use in our calculations afterwards. Section 3 describes how effective average tax rates by households vary in cross section according to income, marital status and the number of dependent children. Section 4 reports facts on the distribution of tax rates, taxes paid and after-tax income distribution. Section 5 offers the parametric estimates for tax functions. Section 6 documents marginal tax rates from the data and from the parametric estimates. Section 7 concludes.

2. Data

We use data from the Internal Revenue Service *2000 Public Use Tax File*. With 145,663 records, it is a representative subsample of the universe of U.S. taxpayers who filed taxes in the year 2000. Since this data effectively contains no restrictions on income, either at the bottom or at the top, it allows for a comprehensive representation of income and tax liabilities across all income levels.[3]

The notion of *income* that we use is standard in cross-sectional studies and encompasses all income flows accruing to households; labor income, asset income from different sources and transfers. It corresponds closely to the notion employed by Gouveia and Strauss (1994). We define *income* to include

- Salaries and wages;
- Interest income (taxable and not taxable);

[2] Several papers estimated effective tax rates for the use in representative-agent models. See, for instance, Joines (1981), Seater (1982), Barro and Sahasakul (1983) for papers that used IRS data, and Mendoza et al. (1994) who estimated effective taxes for a large set of countries using national accounts and revenue statistics. Barro and Redlick (2010) use tax imputations from the TAXSIM program. Differently from these papers, Gouveia and Strauss (1994) used IRS data to estimate *tax functions* – average tax rates as a function of household income. Huggett and Parra (2010) estimate tax functions for retired and non-retired households from tabulated IRS data.

[3] For details on variable definitions, weights used and other technical details, see the *Individual Tax File Sample Description* booklet that accompanies the data.

- Dividends, interest income and royalties;
- Realized capital gains;
- Business or professional income;
- Total pensions and annuities received plus taxable IRA distributions;
- Unemployment compensation;
- Social Security benefits;
- State income tax refunds and alimony received.

It is worth emphasizing that the notion of income that we use is different from the legal notions *Adjusted Gross Income* and *Taxable Income*. In the 2000 tax forms, Adjusted Gross Income was computed by subtracting from all reported income flows IRA and Keough accounts contributions, moving expenses, student loans interest, alimony paid, contributions to medical income savings accounts, among other items. Taxable Income is obtained by subtracting personal exemptions and deductions from Adjusted Gross Income.[4]

Our notion of Federal Income taxes is comprehensive as well. It corresponds to total income taxes *owed* after credits (including the Earned-Income Tax Credit).[5] From this notion of tax liabilities, we calculate for our purposes *effective* average tax rates. Marginal tax rates that we report in the next section correspond to the *statutory* marginal rates for each household given their taxable income in the data.

Sample restrictions Households are included in the sample if (i) their income is strictly positive; (ii) their average tax rate is less than 40%. The second restriction eliminates those with reported taxes higher than the top *statutory marginal* tax rate, 39.5%. The resulting average level of income is US$ 53,063.[6]

2.1. Data limitations

Despite the wealth of information in the IRS data and its advantages, there are important limitations in this data that the reader should be aware of. Several issues are worth noting.

First, there is no detailed information on Social Security taxes paid, neither by households or individuals themselves, nor by employers.[7] Furthermore, social security taxes paid by an individual are not a linear function of earnings due to the cap on taxable earnings. As a result, since the unit of observation is the *household*, not an individual, and social security taxes are based on individual labor income, it is not possible to impute social security taxes for a majority of individuals in the sample (i.e. married ones). Altogether, knowing the labor income of a married household is not sufficient to infer completely the social security taxes of individual spouses. All this implies that from the IRS data we cannot construct a broad notion of Federal tax liabilities in cross section (income plus social insurance taxes). For this reason, we focus only on the relationship between household income and Federal Income taxes.

Second, as it is the case in commonly available data sets, there is no information on employer's contributions to health insurance and pension plans. This is arguably a substantial component of an individual's labor income. Since these contributions are not subject to taxation, they do not appear in individual tax forms and thus, they are absent in the data.

Finally, there is no information on the individual contributions to health insurance and pension plans within firms. Labor income reported in the data set is net of these contributions.[8] Hence, this point and the previous one imply that total labor income accruing to households is underestimated in the IRS data as it is in other commonly used data sets.

2.2. Statutory Federal Income taxes

Before presenting and discussing results on taxes paid by income, household structure and number of children, it is worth reporting the structure of statutory income taxes in 2000. Table 1 summarizes this information for three relevant categories: *married filing jointly, single* and *head of household*. Tax brackets are presented as defined in the law, according to the legal notion of *Taxable Income*.

As the table shows, marginal tax rates range from 15% to 39.6%. The standard deduction for married people is not twice the standard deduction for singles. A similar remark applies to the width of the tax brackets. Very importantly, there is a wide range of income for which statutory marginal tax rates are unchanged; for instance, from $43,850 to $161,450 for married households, marginal rates change by only three percentage points (from 28 to 31 percent). Afterwards, statutory marginal rates increase non-trivially for high-income earners; to 36 and 39.6 percent, respectively. To calculate their taxable

[4] In terms of deductions, households can choose between a lump-sum *standard deduction* or multiple *itemized* deductions, the most common of which corresponds to mortgage interest paid.

[5] More specifically, we use the variable `TOTAL INCOME TAX` (E06500) in the 2000 Public Use Tax File.

[6] The corresponding average level of household income in the commonly used data from the Current Population Survey (CPS) data is somewhat higher: US$ 57,121.

[7] There is only information on social security taxes on tip income, and amounts of Social Security and Medicare (FICA) taxes paid in excess of those required.

[8] The entry for `Wages and Salaries` in the data corresponds to what households report in the first entry in their 1040 forms. The number entered is net of individual contributions to health and pension plans.

Table 1
2000 income tax schedule.

Marginal tax rate	Married filing jointly	Single	Head of household
	Tax brackets (taxable income)	Tax brackets (taxable income)	Tax brackets (taxable income)
15.0%	0–43,850	0–26,250	0–35,150
28.0%	43,850–105,950	26,250–63,550	35,150–90,800
31.0%	105,950–161,450	63,550–132,600	90,800–147,050
36.0%	161,450–288,350	132,600–288,350	147,050–288,350
39.6%	over 288,350	over 288,350	over 288,350
Standard deduction	$7350	$4400	$6450
Personal exemption	2800	2800	2800

Note: This table displays the income tax schedule in the year 2000 for different filing categories.

Table 2
Income distribution statistics.

Quantiles	Share of income	Share of adjusted gross income	Share of taxable income
Bottom			
1%	0.0%	0.0%	0.0%
1–5%	0.1%	0.1%	0.1%
5–10%	0.4%	0.4%	0.2%
Quantiles			
1st (bottom 20%)	2.0%	2.1%	1.4%
2nd (20–40%)	6.2%	6.2%	5.1%
3rd (40–60%)	11.3%	11.3%	10.4%
4th (60–80%)	19.2%	19.6%	18.2%
5th (80–100%)	61.4%	60.9%	65.0%
Top			
90–95%	10.6%	10.8%	10.5%
95–99%	15.0%	14.5%	15.4%
1%	20.9%	20.4%	24.5%
Other statistics			
Gini coefficient	0.59	0.58	0.63
Var–log income	1.50	1.46	2.04

Note: This table shows summary statistics for the distribution of income, adjusted gross income and taxable income in the sample.

income, households have the option of choosing the standard deduction, or a host of *itemized* deductions (e.g. mortgage-interest deduction) that naturally become more attractive at high income levels. Altogether, these features contribute to generate the substantial differences in average tax rates across income levels that we document later, the rapid rise of average rates with income at relatively low income levels, as well as the slow rise of average rates at relatively high income levels.

2.3. Descriptive income statistics

For a better understanding of the facts about tax liabilities in cross section, it is important to report on the properties of the distribution of income in the tax data. This is of interest, since there are no top-coding restrictions as in other commonly used data sets and as a result, the data is representative of the entire universe of U.S. taxpayers.

Table 2 summarizes the properties of the income distribution and highlights the substantial degree of concentration of income at the top. The richest 20% of households earns about 61.4% of total income, whereas the top 10%, 5% and 1% earn about 46.5%, 35.9% and 20.9%, respectively.[9] Table 2 also shows the differences between the notion of income that we consider and the legal notions of Adjusted Gross Income (AGI) and Taxable Income. The results show that the distribution of AGI is very close to the distribution of regular income. Nevertheless, as lump-sum deductions are concentrated at the bottom and have a large impact there, the distribution of taxable income is non-trivially more concentrated than the distribution of income. As a result, the Gini coefficient increases from 0.59 for regular income to about 0.63 for taxable income.

It is important to relate these summary distributional statistics to standard summary measures of income inequality. For instance, CPS data indicates that each quantile earned in 2000 about 3.6, 8.9, 14.8, 23.0 and 49.8 percent of income respectively, whereas the top 5% earned about 22.1 percent with a Gini coefficient of about 0.462.[10] The tax data shows that

[9] From the IRS data one can also assess the importance of the *very rich*. For instance, the top 0.5% earned in 2000 about 16.2% of income.

[10] See http://www.census.gov/hhes/www/income/data/historical/household/index.html.

Table 3
Sources of income.

Quantiles	Labor income	Capital income (I)	Transfer income	Capital income (II)
Bottom				
1%	88.7%	5.3%	6.1%	14.4%
1–5%	86.6%	12.6%	0.8%	14.3%
5–10%	89.2%	9.7%	1.1%	12.5%
Quantiles				
1st (bottom 20%)	88.0%	10.2%	1.8%	14.2%
2nd (20–40%)	88.6%	8.2%	3.2%	15.5%
3rd (40–60%)	89.2%	6.1%	4.9%	12.4%
4th (60–80%)	85.4%	8.8%	5.8%	16.6%
5th (80–100%)	81.6%	15.7%	2.6%	24.2%
Top				
90–95%	82.5%	15.1%	2.4%	22.7%
95–99%	74.0%	24.3%	1.7%	34.3%
1%	57.9%	41.4%	0.6%	54.6%

Note: This table shows the contribution of labor, capital and transfer income at different income levels in the sample. Both notions of capital income introduced in the text are presented.

each quintile earned about 2.0, 6.2, 11.3, 19.2 and 61.4 percent, respectively, whereas the top 5% earned about 35.9% with a Gini coefficient of about 0.59. Clearly, and as also emphasized by others (e.g. Piketty and Saez, 2003), the degree of income concentration from tax data is substantially higher than the degree of income emerging from standard data sets. However, this degree of concentration is quite close to the one found in the Survey of Consumer Finances (SCF). This is not surprising as the SCF does not censor the income of richer households. For instance, Díaz-Giménez et al. (2011) using the SCF found that for 2007 each quintile earned about 2.8, 6.7, 11.3, 18.3 and 60.9, whereas the top 5% earned about 35.9% with a Gini coefficient of about 0.575.[11]

Table 3 shows the varying composition of income as income increases. The third and fifth columns display the fraction of income corresponding to capital income at different quantiles for two concepts of capital income. The first concept of capital income, includes all interest income, dividends, 1/3 of business income, capital gains, rental + royalties income and 1/3 of farm income.[12] The more comprehensive second one, adds to the previous one all pension and annuity payments. In both cases, and as expected, capital income as a share of total income rises rapidly as income goes up. Note that at the very top, more than 40% of income accrues from capital income under the first definition (about 41.4%), whereas under the second notion about 54.6% of income derives from capital income. This is obviously relevant for economic purposes; high income households face much higher marginal tax rates (see below) and capital income is concentrated *there*. Transfers consist of unemployment insurance and social security payments, and labor income constitutes the remaining component of household income. For most households, labor income constitutes the major source of income, but at higher levels of income, income from capital becomes increasingly important. Transfers constitute a small fraction of household incomes and their contribution is hump-shaped, with households around the middle of the distribution of household income receiving the largest share from transfers.

3. Income taxes in cross section

In this section we report basic facts on how average and statutory marginal tax rates vary according to our broad notion of income, marital status and the number of children. For different levels of household income (quantiles as well as bottom and top percentiles), we calculate the *effective* average rate, defined as the average ratio of taxes paid to household's income in the corresponding income category. The marginal rate that we report corresponds to the one encountered by households in their actual tax filing, averaged across all households in the income category. Thus, marginal tax rates reported in this section correspond to mean *statutory* marginal rates.[13]

[11] Data sets such as the CPS and PSID underrepresent the top of the distribution of labor earnings and income. It is also known that there is underreporting in all income categories in the CPS, except in wage income. The data from the IRS is likely to exhibit underreporting of income in some categories (e.g. self-employed income). For instance, internal research from the IRS has found that the underreporting of individual income for tax purposes is the largest contributor to the difference between estimated tax liabilities and taxes effectively paid (tax gap). For instance, in 2006, underreporting of income accounted for 52% of the tax gap of *all* Federal taxes. See Black et al. (2012) for details. According to Johns and Slemrod (2010), business income, as opposed to wages or investment income, accounts for about two-thirds of the understated individual income. Furthermore, misreporting as a percentage of adjusted gross income is increasing in adjusted gross income (reaching about 19% in 0.99 to 0.995 percentiles of the AGI distribution).

[12] We allocate 1/3 of business and farm income according to standard estimates of the share of capital income in total income.

[13] In Section 5, we infer effective marginal tax rates from data.

3.1. Married and unmarried households

Table 4 shows the findings for all households, as well as for married and unmarried households as a group. Married households correspond to those filing as *married filing jointly*, whereas unmarried households encompass all those filing as *single* and as *head of household*. We explicitly include head of households in our unmarried group, as this category is designed for households headed with unmarried individuals with dependents. Tables 5 and 6 show in detail the married and unmarried groups, according to the number of dependent children present in the household. The income levels defining the percentiles in Tables 4, 5 and 6 are the same, and calculated using income data for all households, and to ease comparisons, we also report in these tables the mean income for each income group. Not surprisingly, married household in general have higher incomes than unmarried ones.

A central finding from Tables 4, 5 and 6 is that effective average rates increase substantially as income increases. Increasing household income from the central quintile (40% to 60%) to the top quantile, increases the mean, effectively-paid average tax rate, from about 3.9% to 14.0% for married households, and from 8.2% to 16.7% for unmarried households. In terms of statutory marginal rates, the increase goes from 13.5% to 27.7% for married households, and from 16.2% to 28.6% for unmarried ones. For households at the top 1%, average (marginal) rates are 23.1% (36.3%) for married households and 21.7% (34.6%) for unmarried ones. Hence, from these findings it is clear that there is a substantial degree of tax progressivity built into the tax system. This is the natural result of the observed degrees of income concentration, in conjunction with a tax schedule with increasing marginal tax rates as a function of income.

From Tables 4, 5 and 6 it also transpires that there are substantial differences in average rates between married and unmarried households. At low levels of income, effective rates are substantially higher for unmarried households, while these rates subsequently converge as income increases, and eventually become higher for married households at top levels of income. Fig. 1 illustrates these differences in tax rates. These differences are due to a host of factors; differences in the levels of income concentration between married and unmarried households, differences in standard deductions, and differences in the width of tax brackets and children. These latter factors are arguably more important in reducing average rates at lower levels of income. For instance, children are concentrated in married households and they lead to higher personal exemptions and tax credits, thereby reducing average rates for these households.

Finally, Tables 4, 5 and 6 also report standard deviation of average tax rates across households at different income levels. As Table 4 documents there is significant variation across household in taxes paid, especially at low levels of income, possibly reflecting the fact that a large number of these households do not pay any tax. For the middle quintile (40–60%), the standard deviation of average tax rates (4.3%) is more than half of average tax rate (7%). Not surprisingly, there is less variation at higher levels of income. For households who are at top 1% of income distribution, the standard deviation of taxes (9.7%) is less than half of their average tax rate (22.9%).

We now try to illustrate the effects of the differential tax treatment of married and unmarried households in the United States. To isolate these effects, we use data that is *not* affected by the presence of children – the first panel in Tables 5 and 6. Consider for instance an average married household at the top 95–99% of the distribution (with nearly \$200,000 household income). If both wife and husband have individually the same income (about \$100,000 each), their tax liabilities would likely be higher when married. They would pay as a married household an effective average rate of about 17.6%, whereas as two single individuals, each would pay about 17% or less.[14] At another extreme, if only one of them earns all the household income, the average rate would be about 18.6% if each filed as a single individual, whereas it would be about 17.6% if they filed as a married couple. Other combinations can be constructed from these tables, reflecting the fact that two partners of similar incomes face a tax *penalty* if they marry, whereas those with sufficiently different incomes face a tax *bonus* or subsidy.[15]

Overall, the discussion above is driven by the fact that in the United States, the unit subject to taxation is the *household*, not the individual. The economic implications of this fact go beyond relative payments when married or not. Consider again a married household with no children with an income level at the top 95–99%. Table 5 indicates that this household faces on average a statutory marginal tax of about 30.3%. If all income is earned by one household member, the marginal tax on the first dollar of income earned by the other household member is also taxed at the same rate, 30.3%. This naturally creates large disincentives for labor supply of secondary earners. In contrast, if her/his income were treated as an individual income, the marginal tax rate would be substantially lower. If children are present, the same logic applies.[16]

---

[14] Unmarried households in the top 90–95% have an average income in excess of \$100,000 (\$112,663 according to Table 6) and the average tax rate amounts to 17.1%.

[15] See McCaffery (1997) for a detailed account of the U.S. tax system's treatment of married and single households. On the optimal taxation of couples, see Boskin and Sheshinski (1983), Alesina et al. (2011), Kleven et al. (2009) and Guner et al. (2012a).

[16] In Guner et al. (2012b), we show that these features of the U.S. tax law have large effects on the labor supply of married females. Kaygusuz (2010) studies how much changes in taxes contributed to the growth of married female labor supply in the U.S. since 1970s. Prescott (2004) studies how cross-country differences in taxes affect cross-country differences in hours per worker. Bick and Fuchs-Schundeln (2012) and Chakraborty et al. (2012) look at the relation between taxes and household labor supply across countries.

Table 4
Descriptive tax statistics.

Households	Income level										
	Bottom			Quantiles					Top		
	1%	1–5%	5–10%	20%	20–40%	40–60%	60–80%	80–100%	90–95%	95–99%	1%
All											
Average tax rate	0.023	0.015	0.009	0.017	0.041	0.070	0.096	0.145	0.148	0.179	0.229
(Standard deviation)	(0.065)	(0.048)	(0.033)	(0.039)	(0.041)	(0.043)	(0.045)	(0.059)	(0.043)	(0.059)	(0.097)
Marginal tax rate	0.018	0.020	0.032	0.049	0.121	0.155	0.195	0.279	0.277	0.309	0.360
Average income	396	1775	4006	5363	16,314	29,918	50,834	162,902	112,954	199,399	1,106,822
Married											
Average tax rate	0.008	0.000	0.000	0.001	0.011	0.039	0.075	0.140	0.144	0.177	0.231
(Standard deviation)	(0.055)	(0.000)	(0.002)	(0.011)	(0.020)	(0.034)	(0.032)	(0.056)	(0.037)	(0.055)	(0.095)
Marginal tax rate	0.003	0.001	0.002	0.001	0.076	0.135	0.161	0.277	0.275	0.308	0.363
Average income	300	1840	4038	6626	16,930	30,762	52,024	165,362	113,049	199,669	1,095,168
Unmarried											
Average tax rate	0.024	0.015	0.010	0.018	0.047	0.082	0.125	0.167	0.169	0.186	0.217
(Standard deviation)	(0.066)	(0.049)	(0.033)	(0.040)	(0.042)	(0.040)	(0.043)	(0.063)	(0.064)	(0.076)	(0.104)
Marginal tax rate	0.019	0.021	0.033	0.052	0.131	0.162	0.243	0.286	0.285	0.312	0.346
Average income	403	1772	4000	5258	16,153	29,576	49,094	148,179	112,392	197,627	1,094,423

Note: This table shows average tax rates and statutory marginal rates at different income levels. The statutory marginal rate reported is the average of the corresponding statutory marginal tax rate for each taxpayer unit within the income category.

Table 5
Descriptive tax statistics: Married households.

Households	Income level									
	Bottom			Quantiles					Top		
	1%	1–5%	5–10%	20%	20–40%	40–60%	60–80%	80–100%	90–95%	95–99%	1%
No children											
Average tax rate	0.000	0.000	0.000	0.000	0.020	0.056	0.087	0.149	0.154	0.176	0.211
(Standard deviation)	(0.000)	(0.000)	(0.003)	(0.004)	(0.023)	(0.034)	(0.033)	(0.053)	(0.039)	(0.059)	(0.103)
Marginal tax rate	0.000	0.002	0.003	0.002	0.098	0.131	0.166	0.278	0.274	0.303	0.351
Average income	331	1909	4032	6323	16,899	30,651	52,006	171,798	113,549	199,420	1,088,177
Two children											
Average tax rate	0.000	0.000	0.000	0.000	0.000	0.019	0.062	0.132	0.135	0.177	0.255
(Standard deviation)	(0.000)	(0.000)	(0.000)	(0.004)	(0.002)	(0.022)	(0.022)	(0.056)	(0.031)	(0.051)	(0.078)
Marginal tax rate	0.000	0.000	0.000	0.001	0.043	0.142	0.154	0.276	0.276	0.312	0.377
Average income	195	1399	4035	7327	16,836	30,775	52,436	160,145	112,767	199,406	1,078,971
Two+ children											
Average tax rate	0.000	0.000	0.000	0.000	0.001	0.005	0.040	0.117	0.121	0.172	0.254
(Standard deviation)	(0.000)	(0.000)	(0.000)	(0.000)	(0.009)	(0.011)	(0.024)	(0.066)	(0.034)	(0.054)	(0.082)
Marginal tax rate	0.000	0.000	0.000	0.000	0.011	0.122	0.149	0.265	0.275	0.311	0.377
Average income	238	1024	3784	6948	17,156	30,518	51,719	177,925	112,148	203,725	1,168,642

Note: This table shows average tax rates and statutory marginal rates at different income levels for *married* households. The statutory marginal rate reported is the average of the corresponding statutory marginal tax rate for each household within the income category.

Table 6
Descriptive tax statistics: Unmarried households.

Households	Income level										
	Bottom			Quantiles					Top		
	1%	1–5%	5–10%	20%	20–40%	40–60%	60–80%	80–100%	90–95%	95–99%	1%
No children											
Average tax rate	0.025	0.017	0.011	0.022	0.070	0.098	0.137	0.173	0.171	0.186	0.217
(Standard deviation)	(0.067)	(0.051)	(0.036)	(0.043)	(0.032)	(0.031)	(0.039)	(0.064)	(0.065)	(0.078)	(0.104)
Marginal tax rate	0.020	0.023	0.038	0.064	0.142	0.167	0.261	0.289	0.285	0.312	0.346
Average income	401	1751	3997	4926	16,119	29,752	49,165	153,406	112,663	198,070	1,100,938
Two children											
Average tax rate	0.000	0.000	0.000	0.000	0.000	0.019	0.074	0.135	0.125	0.193	0.247
(Standard deviation)	(0.000)	(0.000)	(0.000)	(0.000)	(0.004)	(0.024)	(0.026)	(0.054)	(0.053)	(0.064)	(0.080)
Marginal tax rate	0.000	0.000	0.000	0.000	0.088	0.148	0.181	0.271	0.276	0.311	0.353
Average income	398	2055	3980	6818	16,140	28,555	48,891	119,829	113,726	198,041	1,031,059
Two+ children											
Average tax rate	0.000	0.005	0.000	0.001	0.000	0.011	0.052	0.122	0.116	0.174	0.281
(Standard deviation)	(0.000)	(0.027)	(0.000)	(0.010)	(0.000)	(0.021)	(0.034)	(0.063)	(0.029)	(0.052)	(0.061)
Marginal tax rate	0.000	0.000	0.000	0.000	0.042	0.147	0.169	0.277	0.255	0.312	0.388
Average income	391	2019	4110	6323	16,344	29,001	49,293	160,838	103,602	192,705	1,595,545

Note: This table shows average tax rates and statutory marginal rates at different income levels for *unmarried* households. The statutory marginal rate reported is the average of the corresponding statutory marginal tax rate for each household within the income category.

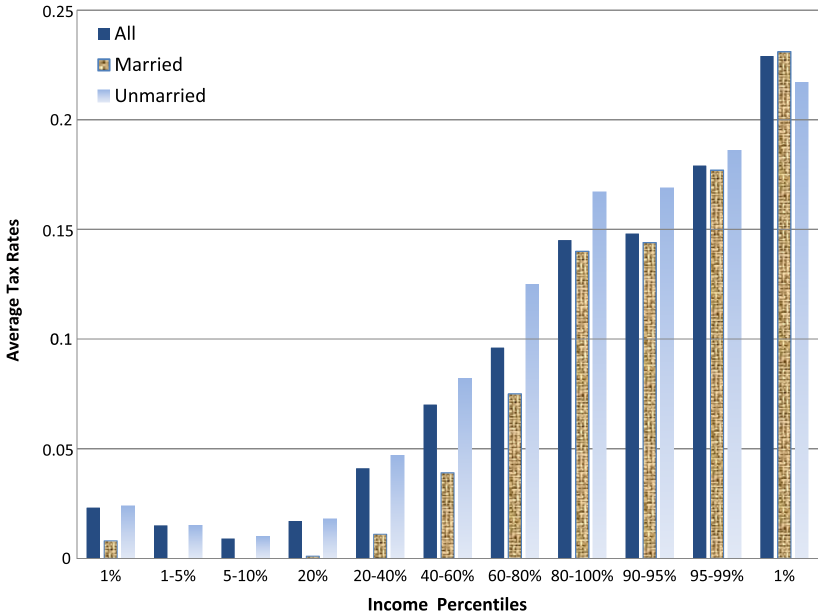


Fig. 1. Average tax rates.

3.2. The role of children

Tables 5 and 6 illustrate the quantitative importance of children in affecting average effective rates for households. As we mentioned earlier, for unmarried households, we use information from the *single filing* category for those without children, whereas for those with children we use information from the *head of household* category.

For married households, children reduce effective rates although the overall effect varies across income levels. Households with income at the top 20%, face an effective average rate of about 15.0% when no children are present, a rate of about 13.2% when two children are present, and a rate of 11.7% when more than two children are present. Therefore, for these households, at extremes, the reduction in effective rates driven by the presence of children is in the ballpark of three percentage points. At very high levels of income, the corresponding reduction in average rates is smaller. Meanwhile, for poorer married households the reduction is naturally much higher than at the top; nearly five percentage points at the central quantile. This is not at all surprising: children disproportionately affect tax liabilities of poorer households via lump-sum personal exemptions and tax credits.

For unmarried households, the patterns just described above are similar but more pronounced; households at the top 20%, face an effective average rate of about 17.3% when no children are present, a rate of about 13.5% when two children are present, and a rate of 12.2% when more than two children are present. For households at the central quintile, the reduction associated to the presence of children in the ballpark of eight percentage points.

3.3. The role of income sources

How do tax liabilities vary according to income sources? We provide an answer to this question in Table A2 in the Online Appendix, by examining average tax rates at different quantiles of the joint distribution of labor and capital income, together with average capital and labor incomes in each quantile. As both capital and labor income increase along the diagonal, tax rates naturally increase. The data reveals that as capital income increases across quantiles, for a given labor income quintile, there is only a small increase in average tax rates. On the other hand, holding a capital income quintile constant and increasing labor income across quantiles, tax rates tend to increase. This finding is consistent with the importance of labor income as a source of income, as Table 3 shows.

3.4. The role of Earned Income Tax Credit (EITC) refunds

As we mentioned above, our notion of Federal Income taxes is the total income taxes owed after credits *including the EITC*. This variable does not take negative values, i.e. credits are subtracted from taxes owed up to zero tax liabilities. The IRS data set we use, however, contains another variable that reports the amount of EITC that is effectively reimbursed to households.[17] Hence, we proceed to subtract EITC refunds from taxes owed by each household and arrive at a measure of

[17] The particular variable is `Earned Income Credit Refundable Portion` (E59720).

taxes *net* of EITC refunds. The EITC, which is a refundable tax credit only available to individuals with positive earnings, is currently the largest anti-poverty program in the U.S. To get the credit, taxpayers must fill a tax return and complete Schedule EITC that gathers information about qualifying children. The EITC is refundable, so it is paid out to households even if the household does not have any federal income tax liability.[18]

In our data set, about 16.6 percent of households receive EITC payments from the government. The average payment is about $1728. The mean income of households that receive EITC payments is about $12,885. Table 7 shows the distribution of average tax rates (net of EITC refunds) for all, married and unmarried households. Not surprisingly, EITC refunds do not affect households at the top of the income distributions and average tax rates in Table 7 are similar to the ones in Tables 4, 5 and 6 above the 2nd quintile of the income distribution. However, the effect of EITC's refunds on poorer households is not trivial. Married household in the first quintile of the income distribution, for example, receives a transfer of 13.4% while the same number for unmarried households is about 4.3%. Overall, since unmarried households dominate at the bottom of the distribution, households in the bottom 20% receive a net transfer of about 4.8%. Households in the second quintile receive a near-zero transfer (0.2%). As the table demonstrates, since the EITC amount is increasing in the number of children, its effect is particularly strong for poorer households with children.

3.5. State and local taxes

How do state and local taxes vary as income changes? Our data allows to provide a partial answer to this question, as the I.R.S. data on state and local taxes is available only for those households who take itemized deductions in their filing of Federal Income taxes. Table A3 in the Online Appendix presents state and local taxes that households pay at different levels of household income. Since itemized deductions are rarely taken at low levels of income, there are essentially no observations of state and local taxes at the bottom income quintile as the table shows. On average, state and local taxes amount to about 4–5%. Poorer households face lower state and local taxes than richer households, but the differences are much smaller than the ones we observe for federal income taxes. The overall structure of state and local taxes is rather flat as a function of income, as Table A3 demonstrates.

4. Income taxes and inequality

In this section, we report on a series of facts related on the distribution of taxes paid, the relative contribution of different income groups, as well as the relationship between taxes and after-tax inequality. We aim at presenting a snapshot on the relationship between the structure of income taxes, and the underlying degree of income inequality.

Tax rates and taxes paid Table A4 in the Online Appendix describes the basic features of the distribution of average tax rates across households when we rank the households by their average taxes rather than by their income. As the table illustrates, a substantial fraction of households has *no* tax liabilities: this occurs for about 14.5% of the married group and for about 31.8% of the unmarried one. Median and mean effective tax rates are on the low side for both groups, with a median rate for married households of about 8.5% and a mean rate for married households of about 8.8%. For unmarried households, the median rate is of about 6.1% whereas the mean rate amounts to 6.4%. The bottom panel of Table A4 shows the tax rates defining the top percentiles. Households at the top of the distribution face significantly higher average rates than those around the middle: the ratio of tax rates defining the bottom 95% to the median is in excess of a factor of 2 for married households, and of a factor of nearly 3 for unmarried households.

How tax liabilities are distributed? Table 8 answers this question, by calculating the share of total taxes paid by different percentiles of the income distribution. The top 20% of households earns about 61.4% of total income and pays more than three quarters of total taxes (79.2%). Similarly, the top 1% earns about 20.9% of total income, yet it accounts for about 35.7% of total tax collections.[19]

After-tax income distribution How much do the before and after-tax income distributions differ? Table 9 shows income-distribution statistics before and after taxes, as well as when including EITC reimbursements.

In terms of the after-tax distribution of income (column 3), despite the vast heterogeneity we documented earlier in terms of income and tax payments, the results show that the before-tax and the after-tax income distributions effectively differ, albeit in relatively moderate amounts. The before-tax shares of the 60–80% quintile and the top quintile are, respectively, 19.2% and 61.4%, whereas the corresponding shares of the after-tax distribution become 20.2% and 58.5%. The largest changes occur for the top 1%, where shares decline from 20.9% to 18.4%. Overall, the changes in shares lead to the Lorenz curve associated to the before-tax income distribution to lie below the Lorenz curve of the after-tax distribution. These

[18] To qualify, the taxpayer must have income (labor and total) below a threshold. The threshold varies by family size and most EITC payments are received by taxpayers with at least one qualifying child. In 2000, taxpayers with two or more children could receive a credit of 40 percent of income up to $9720, for a maximum credit of $3888. Taxpayers (with two or more children) with earnings between $9720 and $12,680 received the maximum credit. Their credit was reduced by 21.1 percent of earnings between $12,680 and $31,152. Hotz and Scholz (2003) and Eissa and Hoynes (2006) provide excellent summaries of the program history and the extensive literature on its effects.

[19] The facts on the distribution of individual income tax liabilities are in line with estimates from the Congressional Budget Office (2012) for the year 2000. They estimate a share of taxes paid by the highest quintile of about 81.2%, and a share of taxes paid by the top 1% of about 36.6%.

Table 7
Descriptive tax statistics (with EITC refunds).

	Income level										
	Bottom			Quantiles					Top		
Average tax rates	1%	1–5%	5–10%	20%	20–40%	40–60%	60–80%	80–100%	90–95%	95–99%	1%
All	−0.045	−0.024	−0.047	−0.048	−0.002	0.068	0.096	0.145	0.148	0.179	0.229
Married (all)	−0.661	−0.140	−0.096	−0.134	−0.038	0.035	0.075	0.140	0.144	0.177	0.231
Married (no children)	−0.036	−0.021	−0.010	−0.014	0.019	0.056	0.087	0.149	0.154	0.176	0.211
Married (one children)	−5.501	−0.543	−0.231	−0.402	−0.136	0.010	0.062	0.132	0.135	0.177	0.255
Married (two+ children)	−0.508	−0.310	−0.279	−0.281	−0.121	−0.006	0.040	0.117	0.121	0.172	0.254
Unmarried (all)	−0.015	−0.020	−0.045	−0.043	0.005	0.080	0.125	0.167	0.169	0.186	0.217
Unmarried (no children)	0.018	0.009	−0.001	0.014	0.069	0.098	0.137	0.173	0.171	0.186	0.217
Unmarried (one children)	−1.248	−0.382	−0.384	−0.374	−0.184	0.002	0.074	0.135	0.125	0.193	0.247
Unmarried (two+ children)	−0.414	−0.376	−0.362	−0.350	−0.171	−0.009	0.052	0.122	0.116	0.174	0.281

Note: This table shows average tax rates at different income levels when EITC refunds are explicitly taken into account. See text for details.

Table 8
Distribution of tax liabilities.

Income level	Share of total taxes paid
Bottom	
1%	0.0%
1–5%	0.0%
5–10%	0.0%
Quantiles	
1st (bottom 20%)	0.3%
2nd (20–40%)	1.8%
3rd (40–60%)	5.7%
4th (60–80%)	13.1%
5th (80–100%)	79.2%
Top	
90–95%	11.2%
95–99%	19.4%
1%	35.7%

Note: This table shows the share of total taxes paid at different levels of income in the sample.

Table 9
After-tax distribution statistics.

Income level	Before tax share of total income	After tax share of total income	After tax and after EITC refunds share of total income
Bottom			
1%	0.0%	0.0%	0.0%
1–5%	0.1%	0.2%	0.2%
5–10%	0.4%	0.4%	0.5%
Quantiles			
1st (bottom 20%)	2.0%	2.3%	2.5%
2nd (20–40%)	6.2%	6.9%	7.1%
3rd (40–60%)	11.3%	12.2%	12.2%
4th (60–80%)	19.2%	20.2%	20.1%
5th (80–100%)	61.4%	58.5%	58.2%
Top			
90–95%	10.6%	10.6%	10.5%
95–99%	15.0%	14.3%	14.2%
1%	20.9%	18.4%	18.3%
Other statistics			
Gini coefficient	0.59	0.56	0.55
Var-log income	1.50	1.39	1.35

Note: This table shows statistics of the distribution of income before taxes, after taxes and after taxes and EITC refunds in the sample.

changes are summarized in the decline of the Gini coefficient due to the presence of income taxes, from about 0.59 to about 0.56.

Once we also include EITC refunds, the resulting changes are of second-order importance. The after-tax, after-refund income share of lowest quantiles increase slightly and the Gini coefficient declines from 0.56 to 0.55.

How progressive is the U.S. tax system? There are different answers to this question. On the one hand, a clear picture emerges from our findings. Effective tax rates on most households are relatively low (below 10%) and differ substantially from those at the top. For instance, married households around median income experience tax rates around 4%, while those at the top 1% face tax rates of around 23%. Furthermore, as we showed earlier, taxes paid are heavily concentrated at the top. In a nutshell, the provisions in the law, in conjunction with the observed dispersion in income lead to the finding that the bulk of tax payments are concentrated in upper income households and that a large fraction of U.S. households have effectively no tax liabilities. From this perspective, the answer to the question above is that there is substantial progressivity in the tax burden as measured by effective, average tax rates. Put differently, moving a hypothetical household along the income ladder implies substantial increases in average tax rates. These findings are reflected in the comparison between the distributions of income before and after taxes.

On the other hand, tax rates at the very top of the income distribution are essentially constant as income changes. Once high income levels are reached, effective tax rates do not change and remain *flat*. We calculate that at *ten* times the level of household income in the sample (about $530,630 in 2000 dollars), the average tax rate for a married household was around 24%. At *fifteen* times the level of mean income (about $795,945) the rate was about 25%, while at twenty times mean household income (about $1,061,260), the rate was effectively unchanged at around 25%. In other words, income tax rates at the top are flat and do not approach the statutory marginal rate on income that these households faced (39.6%). This occurs due to a host of factors, and underlies debates and proposals on tax reform.

5. Parametric estimates

In this section, we provide estimates of tax *functions* for applied use. Specifically, we posit parametric functional forms for effective average tax rates, and estimate the relevant parameters for all households, married and unmarried households, distinguishing by the number of dependent children. We also provide estimates for a number of special cases.[20]

In our choice of functional forms for average tax rates, we are guided by the basic, concave shape of tax rates as a function of income that was evident in our earlier description of tax rates in Section 3. Average rates start at near zero, and grow rapidly as income increases. The growth of average tax rates eventually stabilizes, and rates become nearly constant at high levels of income. All the specifications we present and discuss below are consistent with these patterns.

Functional forms We estimate four specifications for average tax rates. The first two specifications have two parameters while the last two require the estimation of three parameters. In the first case, we posit that

$$t(\tilde{y}) = \alpha + \beta \log(\tilde{y}), \tag{1}$$

where t is the average tax rate, and the variable $\tilde{y}$ stands for multiples of mean household income in the data. That is, a value of $\tilde{y}$ equal to 2.0 implies an average tax rate corresponding to an actual level of income that is twice the magnitude of mean household income in the data. This specification was used by Guner et al. (2012a, 2012b). We refer to it as the *log* specification.

Notice that for this specification, marginal tax rates, m, are given by

$$m(\tilde{y}) = \alpha + \beta \log(\tilde{y}) + \beta = t(\tilde{y}) + \beta. \tag{2}$$

That is, marginal tax rates differ from average tax rates by the constant factor β. In macroeconomic terms, this specification is consistent with *balanced growth*: if all incomes increase by a given factor, average and marginal tax rates are unchanged, and total taxes paid increase by the same factor.

Our second and third specifications are also consistent with balanced growth. The second one corresponds to the function used in Benabou (2002) and Heathcote et al. (2011). We refer to it as the *HSV* specification. It is given by

$$t(\tilde{y}) = 1 - \lambda \tilde{y}^{-\tau}, \tag{3}$$

with corresponding marginal tax rate

$$m(\tilde{y}) = 1 - \lambda(1-\tau)\tilde{y}^{-\tau}. \tag{4}$$

In this specification, the parameter λ controls the level of the tax rate, whereas the parameter τ controls the curvature, or degree progressivity in the tax schedule. If $\tau = 0$, average and marginal tax rates are constant as income changes (flat-rate tax), whereas $\tau > 0$ implies a progressive tax.

Our third specification is given by

$$t(\tilde{y}) = \delta + \gamma \tilde{y}^{\varepsilon}, \tag{5}$$

and

$$m(\tilde{y}) = \delta + (1+\varepsilon)\gamma \tilde{y}^{\varepsilon}. \tag{6}$$

We refer to this specification as the *power* specification. A version of this power function is used by Guvenen et al. (forthcoming).

Finally, we also estimate the same functional form used by Gouveia and Strauss (1994):

$$t(y) = b\left[1 - \left(sy^{p} + 1\right)^{-1/p}\right]. \tag{7}$$

In this case, the variable y stands for the level of household income in the data set. We refer to this as the *GS* specification. The corresponding marginal tax function is

[20] Our strategy is to estimate parametrically average (effective) tax rates for different types of households by their filing status and the number of children, capturing in this way the complex nature of the tax legislation in a parsimonious way. An alternative strategy would be to estimate from the data a relation between household income and household taxable income, and use the statutory tax rates to figure out taxes paid for each household. We briefly report in the Online Appendix how income and taxable income are related in the data, and corresponding implications for tax liabilities.

Table 10
Parameter estimates: All and married households.

Estimates	All	Married (all)	Married No child.	Married One child	Married Two child.	Married Two+ child.
Log						
α	0.099	0.085	0.096	0.089	0.073	0.058
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
β	0.035	0.058	0.054	0.061	0.067	0.060
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
AIC	−458,238.82	−341,358.66	−168,293.14	−56,948.98	−80,694.80	−47,433.30
HSV						
λ	0.902	0.913	0.903	0.910	0.925	0.940
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
τ	0.036	0.060	0.058	0.064	0.070	0.058
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
AIC	−454,289.41	−336,293.15	−166,921.73	−55,710.24	−78,843.81	−45,796.08
Power						
δ	−0.089	−0.451	−0.829	−0.415	−0.495	−0.266
	(0.002)	(0.011)	(0.053)	(0.020)	(0.020)	(0.009)
γ	0.186	0.534	0.923	0.500	0.565	0.320
	(0.002)	(0.011)	(0.053)	(0.020)	(0.020)	(0.009)
ϵ	0.236	0.108	0.059	0.124	0.116	0.186
	(0.002)	(0.002)	(0.003)	(0.005)	(0.004)	(0.004)
AIC	−472,252.24	−344,987.74	−168,759.00	−57,955.38	−81,968.21	−49,976.70
GS						
b	0.264	0.247	0.227	0.251	0.270	0.278
	(0.002)	(0.001)	(0.001)	(0.002)	(0.001)	(0.002)
s	0.012	0.001	0.001	0.001	0.000	0.000
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
p	0.964	1.850	1.842	1.844	2.070	2.602
	(0.006)	(0.014)	(0.023)	(0.029)	(0.025)	(0.042)
AIC	−482,426.01	−360,639.42	−175,942.65	−61,049.20	−87,249.78	−56,009.23

Note: This table shows the parameter estimates for all households as well as for married households for the three specifications considered. Standard errors are in parentheses.

$$m(y) = b\left[1 - \left(sy^p + 1\right)^{-1/p-1}\right]. \tag{8}$$

Some comments about the specifications are in order. First, the *HSV* specification is a special case of the *Power* one. The two coincide when $\delta = 1$, $\gamma = -\lambda$ and $\epsilon = -\tau$. Second, the *log*, the *HSV* and the *GS* specifications imply that the ratio of marginal to average rates approaches 1 from above. These properties suggest that if average tax rates become relatively flat at high levels of income, then effective marginal tax rates will become close to average rates. Instead, the *power* specification implies that as income grows, the ratio of marginal to average rates approaches $1+\epsilon$.[21] Thus, if estimates for the *power* specification dictate a high value of ϵ, this specification may have problems in reproducing the levels of marginal tax rates at high income levels. We return to these issues later, with a discussion of effective marginal tax rates implied by the different parametric estimates and their raw data counterparts.

Parameter estimates Tables 10 and 11 show the parameter estimates for all households and for married and unmarried households (with and without children present in the household), for all the specifications that we consider. As the tables demonstrate, in all cases parameters are estimated quite precisely. We estimate the *log* and the *HSV* specification using Ordinary Least Squares (OLS) and estimate the *Power* and the *GS* formulation using Nonlinear Least Squares (NLS).[22] Both with OLS and NLS estimations, we use the household weights provided by the IRS data.[23]

For illustration purposes, Fig. 2 plots the resulting average tax rates under all specifications for the universe of married households, alongside with data averages at each bin.[24] The figure shows that the resulting shape of average tax rates is similar under all cases; all track the shape of average rates at most income levels. The data shows that a married household around mean income (three times mean income) faces an average rate of about 7.7% (16.5%). The *log* specification implies

---

[21] This requires that ϵ be positive. If instead ϵ is negative, then the ratio converges to one. Our estimates discussed later are always such that $\epsilon > 0$.

[22] Since Gouveia and Strauss (1994) report their estimates for incomes measured in $1000, for consistency we divide household income by 1000 in the estimation.

[23] Nonlinear Least Squares (NLS) estimates minimize the sum of square residuals $\sum_{i=1}^{N} w_i(y_i - x_i(\beta))^2$, where y_i are the data points, $x_i(\beta)$ is a non-linear function of parameters β, and w_i are sample weights. Ordinary Least Squares (OLS) estimates are obtained when $x_i(\beta)$ is a linear function.

[24] From the data we calculate average tax rates at 0.2, 0.6, 1.2, 1.6, etc. times the mean household income. The value for 0.2 corresponds to the average tax rate for households in interval of 0 to 0.4 times the mean income, the value for 0.6 corresponds to the average tax rate for households in interval of 0.4 to 0.8 times the mean income, etc. The parametric estimates are evaluated at 0.2, 0.6, 1.2, 1.4, etc. Note that for the *GS* specification, we simply multiply normalized income by the mean household income in the data (53,063/1000); see also Footnote 22.

Table 11
Parameter estimates: Unmarried households.

Estimates	Unmarried (all)	Unmarried No child.	Unmarried One child.	Unmarried Two child.	Unmarried Two+ child.
Log					
α	0.105	0.121	0.077	0.048	0.037
	(0.000)	(0.000)	(0.001)	(0.001)	(0.002)
β	0.034	0.035	0.042	0.028	0.022
	(0.000)	(0.000)	(0.001)	(0.001)	(0.001)
AIC	−141,800.11	−116,720.40	−16,514.01	−9944.23	−2909.50
HSV					
λ	0.897	0.882	0.926	0.954	0.965
	(0.000)	(0.000)	(0.001)	(0.001)	(0.002)
τ	0.034	0.036	0.042	0.027	0.021
	(0.000)	(0.000)	(0.001)	(0.001)	(0.001)
AIC	−140,705.17	−115,646.18	−16,309.34	−9877.80	−2899.00
Power					
δ	−0.068	−0.086	−0.101	−0.056	−0.049
	(0.002)	(0.003)	(0.005)	(0.003)	(0.005)
γ	0.180	0.212	0.183	0.114	0.092
	(0.002)	(0.003)	(0.005)	(0.003)	(0.006)
ϵ	0.296	0.243	0.345	0.468	0.422
	(0.004)	(0.004)	(0.010)	(0.013)	(0.024)
AIC	−147,304.40	−120,627.61	−17,633.08	−10,887.58	−3100.75
GS					
b	0.238	0.226	0.170	0.197	0.221
	(0.002)	(0.002)	(0.002)	(0.003)	(0.010)
s	0.008	0.019	0.000	0.000	0.000
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
p	1.366	1.192	9.545	7.317	4.078
	(0.018)	(0.016)	(0.153)	(0.479)	(0.368)
AIC	−152,448.50	−124,590.48	−20,396.16	−12,891.46	−3509.09

Note: This table shows the parameter estimates for unmarried households for the three specifications considered. Standard errors are in parentheses.

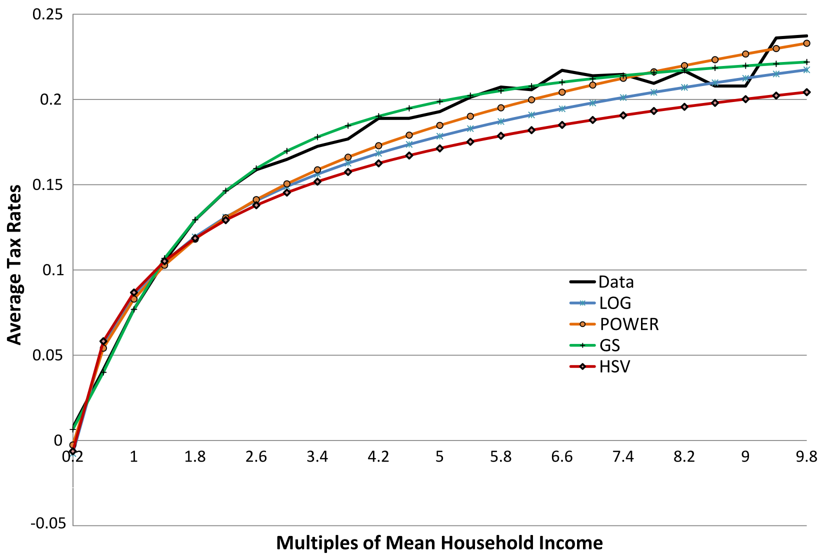


Fig. 2. Average tax rates for married households (data and the parametric estimates).

that a married household around mean income (three times mean income) faces an average rate of about 8.5% (14.9%). The corresponding values under the *HSV* specification are 8.7% (14.5%), under the *power* specification are 8.3% (15.1%), and under the *GS* specification are 7.7% (17.0%), respectively. Overall, the fit of all tax functions is very good. Indeed, it is good even at high levels of income.

The role of children and marital status are straightforward; average tax rates tend to be lower for married households, and tend to decrease with the presence of children in the household. This is straightforward to see for the *log* and *HSV* specification. Note that when $\tilde{y}$ equals 1.0, household income equals mean income, and the average tax rate equals α in the *log* case and $1 - \lambda$ in the *HSV* case.

Tables 10 and 11 also provide a more formal statistical comparison between different specifications. For each case, we report the Akaike information criteria (AIC), originally proposed by Akaike (1973), and widely used to compare non-linear models. The AIC basically favors the model with the maximum log-likelihood criteria with an adjustment for simplicity. A smaller value for the AIC indicates a better fit.[25] The results in Tables 10 and 11 provide a consistent story. Across all specifications, three-parameter specifications perform better than two-parameter specifications and the *GS* specification performs best. Between two-parameter specifications *log* performs betters than the *HSV* case.[26]

Special cases We present in Table 12 the parameter estimates for a number of special cases, as they can be useful in different research applications. We consider the cases of (i) negative taxes by subtracting EITC refunds from households' tax liabilities; (ii) only households with positive social security income; (iii) households without social security income; (iv) households with only labor income, and (v) when state and local income taxes are included alongside federal income taxes. We present results for all households, as well as for all married and unmarried ones. To save space, we focus on one two-parameter specification (the *log* case) and one three-parameter specification (the *GS* case) with better AIC. We present the results for the *HSV* and *power* functions in Table A5 in the Online Appendix.

Fig. 3(a) shows tax functions for married households and for the benchmark economy when we consider EITC refunds. Not surprisingly, the tax functions are steeper when we allow for transfers. When we only consider taxes paid, a married household with mean household income faces and average tax rate of 8.5%, while the tax rate declines to 7.4% once we take EITC refunds into account. The figure naturally displays the negative rates that emerge at very low levels of income.

Fig. 3(b) shows tax functions for married households with positive social security income as well as those without any social security income for the *log* case. As the figure shows, considering households with only social security income implies a counter-clockwise shift in the tax function. Households that receive positive transfers from the social security system face lower taxes at lower levels of income. At very low levels of incomes this difference is significant. At 60% of mean household income, for instance, households with positive social security income face average tax rates that are about 3 percentage points *lower* than for the case of all married households. This difference declines, however, at higher levels of income: it is about 2.2 percentage points at the mean level of household income and reverses around four times mean household income. Overall, these patterns are not surprising. At low levels of income, social security transfers constitute the bulk of income of these households and social security transfers receive preferential tax treatment; at higher levels of income, retired households, who constitute the bulk of social security recipients, have access to less deductions than more typical households and the contribution of social security transfers to household income declines. Not surprisingly, the picture for the households without any social security income is exactly the opposite. They pay higher (lower) taxes at lower (higher) levels of income.

In Fig. 3(c), we show, as an illustration, how the *log* tax function for married households is affected if we include as tax liabilities the sum of federal, state and local taxes. Consistent with Table A3, state and local taxes imply an almost parallel, upward shift of about 5 percentage points after the mean income.

Finally, Fig. 3(d) shows how the *log* tax function is affected when households have only labor income. The figure displays a flatter tax function than in the benchmark case. For comparison purposes, Fig. 3(c) plots the function for households who do not have any social security income as this is a natural group to compare with those who only have labor income. Two functions appear similar, especially at low levels of income. Households who have only labor income face on average lower tax rate than all households, possibly reflecting different characteristics of these households in terms of household size, etc.[27]

Comparisons with previous estimates It is of interest to compare the estimated tax functions with the existing ones from Gouveia and Strauss (1994), who provided estimates for effective rates using data from 1980 and 1989 for all households. This comparison is displayed in Fig. 4, where the corresponding average rates are plotted for these three years.[28]

The figure indicates that there are only minor differences in the resulting average tax functions between 1989 and 2000. Differences occur only at higher income levels and are in the ballpark of one percentage point. The results largely suggest that changes in taxes that took place in 1991 and 1994 did not affect effective average rates significantly. In contrast, as the figure demonstrates, the changes in the tax structure that took place in the 1980s, affected the shape of average rates significantly. For higher income households, the differences are quite substantial; for instance, at five times mean income levels, the differences between 2000 and 1980 are in excess of eleven percentage points.

---

[25] Formally, $AIC = -2\ln L + 2q$, where L is likelihood value and q is the number of parameters.

[26] The basic picture remains the same if we used other commonly used statistics, such as the Bayesian information criteria (BIC), calculated as $BIC = -2\ln L + (\ln N)q$, where N is the number of observations or the squared correlation between data and the predicted values.

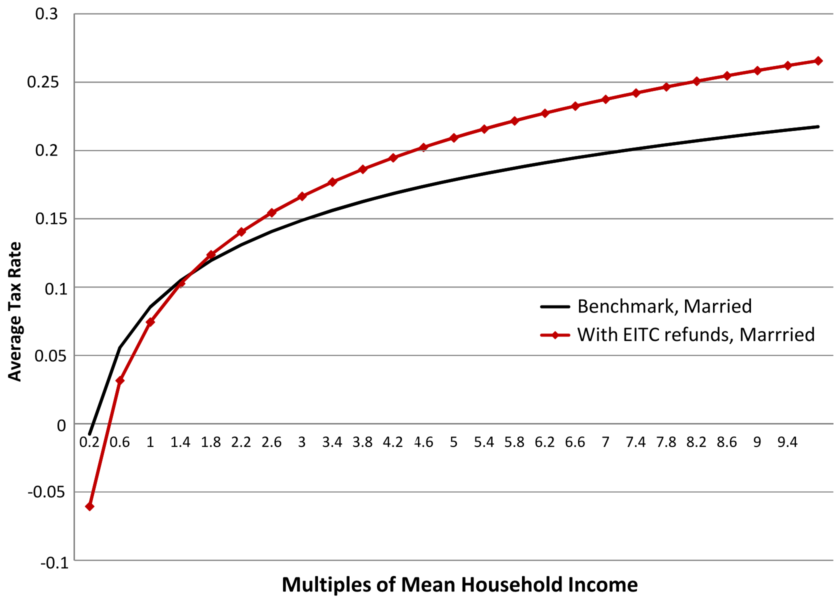
[27] One caveat with these calculations is that given the nature of the tax data, observations on labor income refer to the labor income of households and *not* of individuals.

[28] Gouveia and Strauss (1994) estimates for 1989 and 1980 are for real (in 1990 US dollars) household income, while we use nominal household income for 2000 estimates. Simple algebra shows that parameter s for 1980 and 1989 needs to be adjusted to $\frac{s}{f^{-p}}$; where f is the adjustment for price level ($CPI1980/CPI1990$ and $CPI1989/CPI1990$) to arrive at estimates that can be used with nominal income levels. For Fig. 4, we use mean household income data from the Census Bureau: \$57,135, \$36,520 and \$21,063 for 2000, 1989 and 1980, respectively. See also Footnote 22.

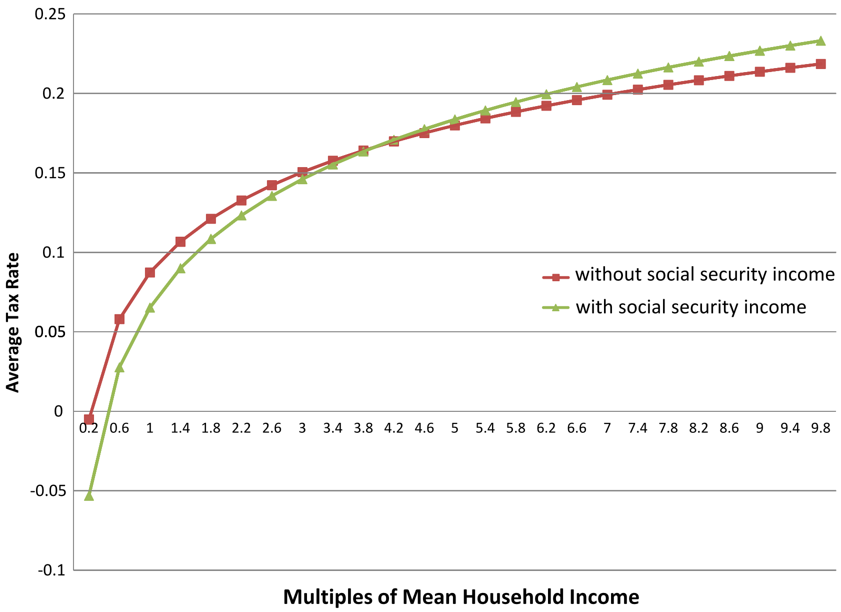
Table 12
Parameter estimates: Special cases.

Estimates	w/EITC			w/SS income			wo/SS income			Only labor income			w/ State and local		
	All	Married	Unmarried	All	Married	Unmarried	All	Married	Unmarried	All	Married	Unmarried	All	Married	Unmarried
Log															
α	0.090	0.074	0.095	0.085	0.074	0.099	0.101	0.087	0.107	0.089	0.077	0.095	0.135	0.117	0.158
	(0.000)	(0.000)	(0.001)	(0.000)	(0.000)	(0.001)	(0.000)	(0.000)	(0.000)	(0.001)	(0.001)	(0.001)	(0.000)	(0.000)	(0.001)
β	0.053	0.084	0.050	0.057	0.065	0.059	0.035	0.057	0.034	0.031	0.047	0.033	0.062	0.075	0.068
	(0.000)	(0.000)	(0.000)	(0.000)	(0.001)	(0.001)	(0.000)	(0.000)	(0.000)	(0.000)	(0.001)	(0.000)	(0.000)	(0.000)	(0.001)
GS															
b	0.170	0.225	0.194	0.196	0.200	0.194	0.276	0.261	0.251	0.182	0.204	0.276	0.317	0.305	0.287
	(0.001)	(0.002)	(0.002)	(0.001)	(0.000)	(0.002)	(0.002)	(0.001)	(0.003)	(0.005)	(0.007)	(0.011)	(0.002)	(0.001)	(0.003)
s	0.000	0.000	0.000	0.000	0.000	0.000	0.013	0.001	0.008	0.008	0.000	0.005	0.016	0.002	0.006
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.001)	(0.000)	(0.000)	(0.000)	(0.000)	(0.001)
p	9.145	3.084	10.389	3.068	6.365	5.389	0.926	1.644	1.336	1.496	2.322	1.444	0.940	1.604	1.514
	(0.821)	(0.134)	(0.280)	(0.096)	(0.376)	(0.468)	(0.006)	(0.012)	(0.018)	(0.048)	(0.135)	(0.041)	(0.011)	(0.020)	(0.051)

Note: This table shows the parameter estimates for all, married and unmarried households under two specifications, *log* and *GS*. Four alternative cases are considered: with only social security income, without social security income, with only labor income and including state and local taxes. Standard errors are in parentheses.



(a)



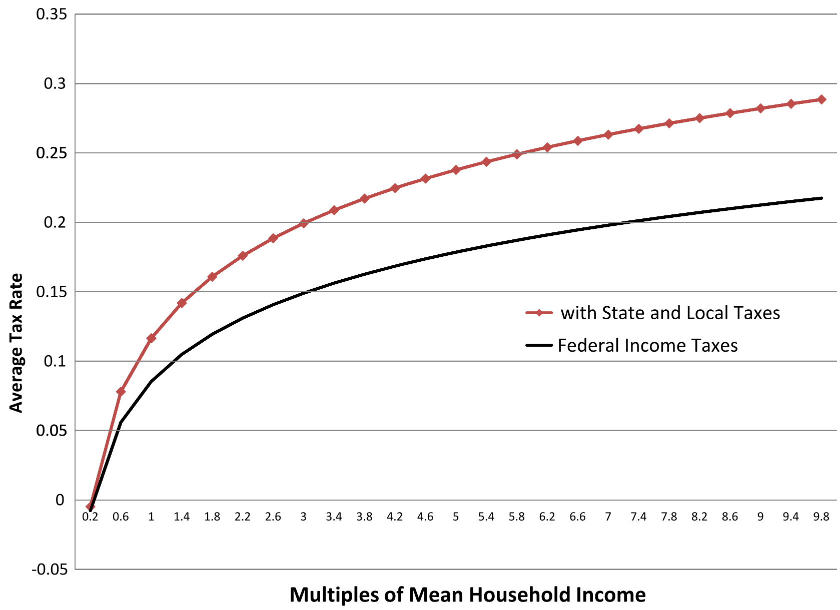
(b)

Fig. 3. Log tax function: (a) married households, with and without EITC refunds; (b) married households, with and without social security income; (c) married households, with and without state and local taxes; (d) all households, with labor income only.

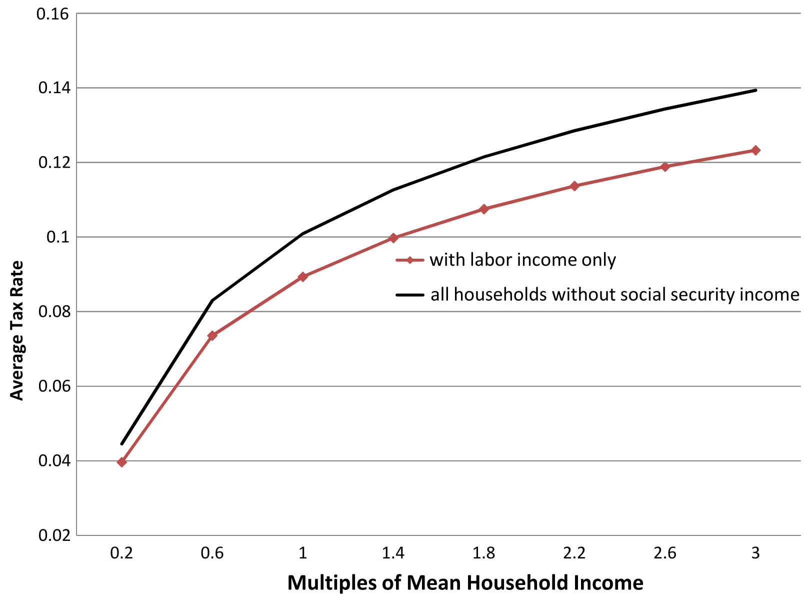
5.1. Caveats

To close this section, it is important to mention two caveats regarding the estimated tax functions, and their use in applied work. First, as we discussed earlier in Section 2.1, an arguably important portion of labor income is not considered in the I.R.S. data nor in other data sets; wage and salary income recorded is net of individual contributions to pension and health plans, and employers contributions to these are not recorded. Researchers using the estimated tax functions should be aware of this fact, and adjust their notion of income or taxable income appropriately if they deem it necessary.

Secondly, our benchmark parametric estimates pertain only to Federal Income Taxes, and for the reasons we explained earlier, they do not take into account state income taxes as well as Social Security and Medicare taxes. Of course, these other forms of taxation also distort behavior in economic models. Hence, researchers using our estimates should, if necessary and possible in the context of their analyses, impute these forms of taxation to get a more accurate notion of distorting taxes.



(c)



(d)

Fig. 3. (*continued*)

6. Marginal tax rates

We now turn our attention to the marginal tax rates emerging from the data and our parametric estimates. We compare the *effective* marginal tax rates implied by our parametric estimates with measures of marginal tax rates from the data. For these purposes, we compute directly effective marginal tax rates from the data. We also compare these effective marginal rates with the statutory ones that households encounter in their tax filing.

There is debate on whether effective or statutory marginal rates are the relevant measures of distortions. Effective rates reflect the inframarginal exemptions, deductions, etc., that reduce average rates. Yet, it can be argued that for many economic decisions the relevant marginal rates are those from the actual tax schedule (statutory rates), as they are the operative ones for decisions on the margin; e.g. to work overtime or not, labor force participation decisions of secondary earners, buying or selling extra units of assets, make charitable contributions, etc. We do not take sides on this debate here. Instead, we report on the effective marginal rates emerging from our data and parametric estimates, and compare them with the actual statutory rates at different levels of income.

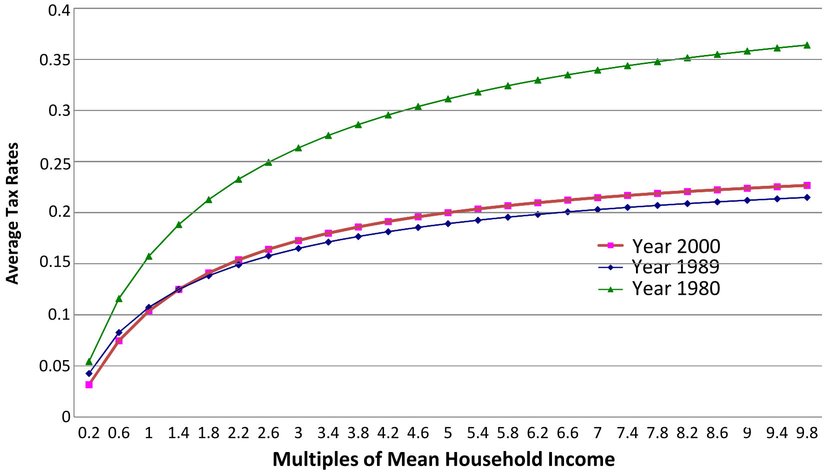


Fig. 4. *GS* tax functions (all households).

Computation of effective marginal rates Our approximation to marginal tax rates in the data is as follows. For a given level of income, say y_0, we compute the variation in tax liabilities when income increases to $y_0 + \Delta y$ and when income decreases to $y_0 - \Delta y$. Let $m(y_0^+)$ be the marginal tax rate when income increases from y_0, and let $m(y_0^-)$ be marginal tax rate when income decreases from y_0. Let average tax rates be given by $t(y)$ as before. Hence $m(y_0^+)$ and $m(y_0^-)$ are given by

$$m(y_0^+) \equiv \frac{(y_0 + \Delta y)t(y_0 + \Delta y) - y_0 t(y_0)}{\Delta y},$$

and

$$m(y_0^-) \equiv \frac{(y_0 - \Delta y)t(y_0 - \Delta y) - y_0 t(y_0)}{-\Delta y}.$$

Simple algebra implies that

$$m(y_0^+) = \left[t(y_0 + \Delta y) - t(y_0)\right]\frac{y_0}{\Delta y} + t(y_0 + \Delta y), \tag{9}$$

and

$$m(y_0^-) = \left[t(y_0) - t(y_0 - \Delta y)\right]\frac{y_0}{\Delta y} + t(y_0 - \Delta y). \tag{10}$$

Not surprisingly, Eqs. (9) and (10) show that marginal tax rates exceed average rates by a factor that is proportional to the change in average rates as income changes. Given these marginal rates, we calculate our approximation to the marginal tax rate at income level y_0 by averaging out $m(y_0^+)$ and $m(y_0^-)$.

We operationalize the calculation of marginal tax rates as follows. We create a range of income levels, defined as multiples of mean household income, that correspond to the income levels y_0 above. We set the minimum income level to 0.25 times the mean household income and the maximum income to 10 times the mean income and the distance (mesh size) between income levels, Δ, is given by 0.25 times mean income. Given a particular y_0, we calculate $t(y_0)$ as the average tax rate in interval $[1.2y_0, 0.8y_0]$. Given values for y_0, $t(y_0)$ and Δ, we compute the marginal tax rates using Eqs. (9) and (10), and average $m(y_0^+)$ and $m(y_0^-)$ to find the marginal tax rate at income level y_0.

Findings Our findings are summarized in Figs. 5 and 6 for the case of all married households. Fig. 5 shows the effective marginal tax rates computed as discussed above, alongside the corresponding statutory rates. For any particular income level, the statutory marginal tax rates simply reflect the average value of the statutory marginal tax rates around that income level (similar to Tables 4–6). From the figure, it is clear that statutory marginal rates can exceed effective marginal rates by substantial amounts, and that the gap grows with household income. For instance, at twice mean income, the statutory is above 25%, whereas it is around 22.5% in effective terms. At seven times mean income, the statutory rate is about 35%, whereas the effective marginal rates are 25–26%. Overall, Fig. 5 illustrates the fact that as effective average tax rates flatten out at high levels of income, effective marginal rates follow, despite the fact that statutory marginal tax rates increase with income to the highest possible level of 39.6%. We conclude from these findings that the avenues for reducing households' tax liabilities in practice lead to much lower marginal tax rates in effective terms. For households at high levels of income, the gap relative to statutory rates is large and of about ten percentage points.

Fig. 6 illustrates that the marginal tax rates emerging from our parametric specifications track the concave-shaped empirical estimates well. A few observations of the Fig. 6 are in order. First, the marginal rates emerging from the *GS* specification

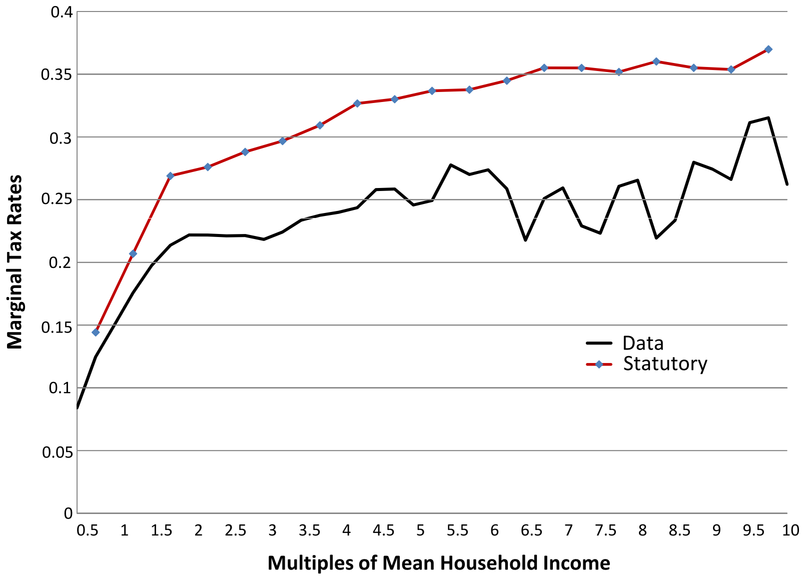


Fig. 5. Marginal tax rates for married households (data and statutory rates).

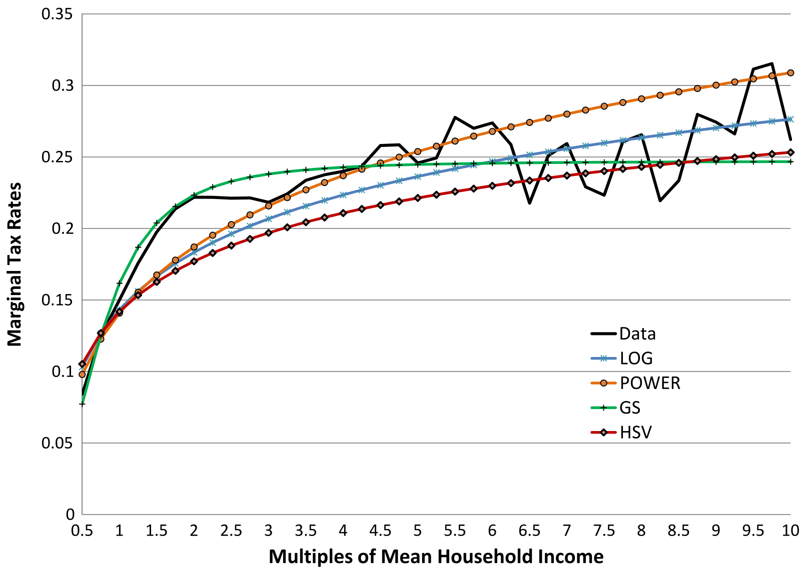


Fig. 6. Marginal tax rates for married households (data and the parametric estimates).

become essentially constant at relatively low levels of income (about twice mean income). Second, the gap between the *log*, *HSV* and *Power* specification grows with income. Indeed, the three-parameter *Power* specification leads to the largest marginal rates at high levels of income (in excess of 30% at ten times mean income), and is the one closest to the statutory marginal tax rates.

7. Conclusion

We presented basic facts on the effective taxation of U.S. households in cross section, distinguishing them by their marital status, the number of dependent children, and other characteristics. We have done so by exploiting the rich cross-sectional data from the U.S. Internal Revenue Service for the year 2000. This allowed us to document the substantial degree of heterogeneity observed in income and taxes paid across U.S. households.

A central contribution of our paper is the estimation of parametric estimates of effective tax *functions* that can be readily used in applied work. We estimated four specifications for different household categories (e.g. married households). All these specifications account for the patterns of average taxes as a function of income quite well.

We conclude the paper by mentioning one caveat in interpreting our results. The caveat is that they pertain to the structure of federal income taxation prevailing in the year 2000. Naturally, the temporary changes that occurred in 2001 and 2003 (*Economic Growth and Tax Relief Reconciliation Act of 2001* and the *Jobs and Growth Tax Relief Reconciliation Act of 2003*) are not captured in our analysis. Nonetheless, we view the snapshot presented of the relationship of taxes and income in cross section as a very good approximation of the nonlinearity (and potential distortions) present in the current system. Indeed, as we write, the tax structure in 2013 under recent amendments is much closer to the 2000 structure than in previous years. For instance, top marginal rates are back to 2000 levels.

Appendix A. Supplementary material

Supplementary material related to this article can be found online at http://dx.doi.org/10.1016/j.red.2014.01.003.

References

Akaike, H., 1973. Information theory as an extension of the maximum likelihood principle. In: Petrov, B.N., Csaki, F. (Eds.), Second International Symposium on Information Theory. Akademiai Kiado, Budapest, pp. 267–281.
Alesina, A., Ichino, A., Karabarbounis, L., 2011. Gender based taxation and the division of family chores. American Economic Journal: Economic Policy 3 (2), 1–40.
Altig, D., Auerbach, A., Kotlikoff, L., Smetters, K., Walliser, J., 2001. Simulating fundamental tax reform in the United States. American Economic Review 91 (3), 574–595.
Barro, R., Redlick, C., 2010. Macroeconomic effects from government purchases and taxes. Mimeo.
Barro, R.T., Sahasakul, C., 1983. Measuring the average marginal tax rate from individual income tax. Journal of Business 56 (4), 419–452.
Benabou, R., 2002. Tax and education policy in a heterogeneous-agent economy: what levels of redistribution maximize growth and efficiency?. Econometrica 70 (2), 481–517.
Bick, A., Fuchs-Schundeln, N., 2012. Taxation and labor supply of married women across countries: A macroeconomic analysis. Mimeo.
Black, T., Bloomquist, K., Emblom, E., Johns, A., Plumley, A., Stuk, E., 2012. Federal tax compliance research: Tax year 2006 tax gap estimation. IRS Research, Analysis & Statistics Working Paper.
Boskin, M.J., Sheshinski, E., 1983. Optimal tax treatment of the family: Married couples. Journal of Public Economics 20 (3), 281–297.
Castañeda, A., Díaz-Giménez, J., Ríos-Rull, J.V., 2003. Accounting for the US earnings and wealth inequality. Journal of Political Economy 111 (4), 818–857.
Chakraborty, I., Holter, H.A., Stepanchuk, S., 2012. Marriage stability, taxation and aggregate labor in the US vs. Europe. Mimeo.
Conesa, J.C., Krueger, D., 2006. On the optimal progressivity of the income tax code. Journal of Monetary Economics 53 (7), 1425–1450.
Conesa, J.C., Kitao, S., Krueger, D., 2009. Taxing capital? Not a bad idea after all!. American Economic Review 99 (1), 25–48.
Congressional Budget Office, 2012. The Distribution of Household Income and Federal Taxes, 2008 and 2009. Statistical Supplement. Washington, DC.
Díaz-Giménez, J., Pijoan-Mas, J., 2005. Flat tax reforms in the U.S.: A boon for the income poor. Mimeo.
Díaz-Giménez, J., Glover, A., Ríos-Rull, José-Víctor, 2011. Facts on the distributions of earnings, income, and wealth in the United States: 2007 update. Federal Reserve Bank of Minneapolis Quarterly Review 34 (1).
Eissa, N., Hoynes, Hilary W., 2006. Behavioral responses to taxes: Lessons from the EITC and labor supply. Tax Policy and the Economy 20, 73–110.
Erosa, A., Koreshkova, T., 2007. Progressive taxation in a dynastic model of human capital. Journal of Monetary Economics 54 (3), 667–685.
Gouvieia, M., Strauss, R., 1994. Effective federal individual income tax functions: An exploratory analysis. National Tax Journal 47 (2), 317–339.
Guner, N., Kaygusuz, R., Ventura, G., 2012a. Taxation and household labor supply. The Review of Economic Studies 79 (3), 1113–1149.
Guner, N., Kaygusuz, R., Ventura, G., 2012b. Taxing women: A macroeconomic analysis. Journal of Monetary Economics 59 (1), 129–132.
Guvenen, F., Kuruscu, B., Ozkan, S., forthcoming. Taxation of human capital and wage inequality: a cross-country analysis. Review of Economic Studies. http://dx.doi.org/10.1093/restud/rdt042.
Heathcote, J., Storesletten, J., Violante, G.L., 2009. Quantitative macroeconomics with heterogeneous households. Annual Review of Economics 1, 319–354.
Heathcote, J., Storesletten, J., Violante, G.L., 2011. Redistributive taxation in a partial-insurance economy. Mimeo.
Hotz, V. Joseph, Scholz, John Karl, 2003. The earned income tax credit. In: Moffitt, Robert A. (Ed.), Means-Tested Transfer Programs in the United States. University of Chicago Press.
Huggett, M., Parra, J.C., 2010. How well does the US social insurance system provide social insurance?. Journal of Political Economy 118 (1), 76–112.
Johns, A., Slemrod, J., 2010. The distribution of income tax noncompliance. National Tax Journal 63 (3), 397–418.
Joines, D.H., 1981. Estimates of effective marginal tax rates on factor incomes. Journal of Business 54 (2), 191–226.
Kaygusuz, R., 2010. Taxes and female labor supply. Review of Economic Dynamics 13 (4), 725–741.
Kleven, H., Kreiner, C.T., Saez, E., 2009. The optimal taxation of couples. Econometrica 77 (2), 537–560.
McCaffery, E.J., 1997. Taxing Women. The University of Chicago Press, Chicago, IL.
Mendoza, E.G., Razin, A., Tesar, L.L., 1994. Effective tax rates in macroeconomics: Cross-country estimates of tax rates on factor incomes and consumption. Journal of Monetary Economics 34 (3), 297–323.
Nishiyama, S., Smetters, K., 2005. Consumption taxes and economic efficiency with idiosyncratic wage shocks. Journal of Political Economy 113 (5), 1088–1115.
Piketty, T., Saez, E., 2003. Income inequality in the United States, 1913–1998. Quarterly Journal of Economics 118 (1), 1–39.
Prescott, E., 2004. Why do Americans work so much more than Europeans?. Federal Reserve Bank of Minneapolis Quarterly Review, 2–13.
Seater, E.M., 1982. Marginal federal personal and corporate income tax rates in the U.S., 1909–1975. Journal of Monetary Economics 10 (3), 121–135.
Ventura, G., 1999. Flat tax reform: A quantitative exploration. Journal of Economic Dynamics & Control 23 (9–10), 1425–1458.