# Mathematical control of income tax revenue distribution for the US joint tax returns: 1996 - 2008

**Misaina Navaloniana Andrianjafinandrasana**

September 25, 2012

School of Computational and Applied Mathematics
University of the Witwatersrand, Johannesburg

A dissertation submitted to the faculty of Science
in fulfillment of the requirements for the degree of Doctor of Philosophy

# Declaration

I declare that this thesis is my own, unaided work and it has not been submitted before for any degree or examination in any other university. It is being submitted for the degree of Doctor of Philosophy in the University of the Witwatersrand, Johannesburg.

On September 25, 2012

MN Andrianjafinandrasana

# Abstract

This thesis introduces time dimension into income taxation. Economic theory derives optimal income taxation by explaining individual behaviors through utility and social welfare maximization. We argue that observing collective behaviors through the distribution of number of returns and the revenue themselves offers a viable alternative. Our study is confined to US joint returns individual income tax from 1996 to 2008. We introduce a combination of cubic spline fit to data and Pareto distribution to estimate the revenue and number of returns distributions. Normalization admits comparative study of year to year changes in those distributions. We discover that the tax burden of low income taxpayers has decreased consistently, while the conventional rule of inflation indexation of tax bracket only is applied. This is an instability in the sense that the distribution of revenue transforms into a qualitatively different distribution over time. We derive the control of the instability. Control using brackets only implies optimal brackets are widened over the peak of the revenue and shifted to the right for higher brackets. Control using rates only implies an overall reduction of marginal tax rates. Control using both brackets and marginal rates combines these two features. Some brackets are removed and consequently marginal rates for some taxpayers are reduced. Dependency of the distribution of number of returns to tax liability and tax parameters proves to have insignificant effect on the optimal tax parameters. Fairness is defined in the literature as tax liability per unit income. Control of the instability together with a uniform change in fairness is addressed. There is a trade-off between controlling instability and ensuring constant change in the fairness of the tax system. We derive the optimal brackets and marginal rates that minimize this trade-off. As expected, the general rule for optimality is that marginal tax rates are barely changed to provide constant change in fairness. Yet, small adjustment of the brackets provides control of the instability.

*ho an'i dada sy neny,*
*ary izy fito mirahavavy,*

*ianareo no reharehako*

# Acknowledgements

I am $(\text{very})^{\infty}$ grateful to my supervisors Prof. David Sherwell and Colin Myburgh for their guidance and strong support over the years. Our many many brainstorming sessions and your hands-on advices (without hands-on management) really shaped me into a better scientist.

I would like to thank all friends and colleagues from the AIMS lab: Obi, Dario, Asha, Tumelo, Viren, Terry n Tanya, Frank, Giddy, Mboto and Sima. By just being you, you made a contribution to this work.

I would like to thank Son excellence Monsieur le Maire and Charles-Mokoto-le plus chretien, for keeping up the AIMS spirit. I might have finished earlier if I didn't know you but it's more likely that I wouldn't have finished at all.

I thankfully acknowledge financial supports from AIMS, the School of Computational and Applied Mathematics, The university of Witwatersrand, Johannesburg, The Bradlow foundation and the DAAD.

I would like to thank everyone I should have thanked but I have forgotten to thank already.

# Contents

# List of Figures

# List of Tables

# Chapter 1

# Introduction

This thesis investigates a new approach to income taxation by considering the dynamics in time of the actual tax revenue distribution from data. If we assume that the income tax revenue distribution is inherently a time varying process then the natural questions to ask are: what characterizes its dynamics? Are there any equilibrium states, stable or unstable, that the revenue converges to? More importantly, can we control the shape of the revenue distribution, given its current shape and in view of its trend and dynamics, optimally, so as to preserve desirable properties of the tax system? Those are the questions that this thesis answers, for the US joint returns individual income tax system between 1996 and 2008.

The general motivation of this thesis is to provide a clear practical procedure to obtain optimal tax parameters such as tax brackets or marginal tax rates to attain some predefined objectives. Economic theory gives two approaches in this regard. First, the theory of optimal income taxation following the seminal paper of Mirrlees (1971), provides a general framework to analyze efficiency and redistribution. This theory is very general and has not lead to implementable policy. The second approach is the theory of optimal tax reforms whereby taxpayer's behaviors are simulated, Creedy and Duncan (2002). Individual behaviors are obtained from the use of utility functions that quantify their labor supply choices. This leads to clear, implementable policy.

Both approaches are dependent on the choice of the utility and welfare functions. These are difficult to determine empirically and are modeled by appropriate functions. In this thesis, we avoid using utility functions by observing changes in the revenue distribution. Publicly available data from the IRS is used, that yields simple implementable optimal tax parameters. *In other words, our goal is to provide an empirical tool that only uses tax revenue data.*

Optimality is redefined in our analysis to reflect the emphasis on time dependence of the revenue distribution. As a result, our analysis is related to economic theory of income taxation but is not directly comparable. This chapter identifies those links to economic theory and currently applied policies where needed. We give motivations to our approach at greater length and define important concept used in the thesis, to set the reader in the right context. A summarized outline of the thesis is given at the end of this introduction.

## 1.1 The US Federal Individual Income Tax System

Detailed individual income tax data is not freely available in most countries. Particularly in South Africa, where the minimum data required for our analysis has only been available since 2007. Meanwhile, the Internal Revenue Service in the US publishes yearly in the Statistics of Income (2008) the summary of their panel data: Continuous Work History Sample, Weber (2002). This sample is freely available for academics in America but not for others. Chapter 2 is devoted in working around this obstacle. We show that if the data is given with finer income bins, the original distribution is recovered successfully. In this section, we introduce the reader to the American individual income tax system. We describe briefly the general structure of the federal individual income tax in the US. We explain how to compute the tax liability of a taxpayer and what factors affect the final tax collection.

### 1.1.1 Adjusted Gross Income (AGI)

The AGI encompasses all sources of income for individual taxpayers. Namely, wage and salaries, dividend and interests, taxable refunds, alimony received, sales of capital assets (net gain or loss), business net gain or loss (self-employed) as well as net income or loss for farm, farm rental, royalty, rent, partnership, S corporation and estate or trust. This also includes unemployment compensation, depreciation in excess of a straight line depreciation, total pension income, other net income or loss, net operating loss. See Statistics of Income (2008) for more information.

There are deductions[1] to be subtracted to arrive at the reported AGI. Those are, disallowed passive losses, moving expenses, alimony paid and unreimbursed business expenses. Note that some of these deductions are different from the ones prior to 1987, making it incompatible to compare year to year AGI beyond that year. The IRS defined the 1979 AGI concept to bridge this

[1] Also known as adjustments, and thus the “*Adjusted*”in AGI.

difference. This is used in most of the literature that uses the IRS Statistics of Income data. We will not use this concept here since the focus of this thesis is from 1996 to 2007. The current AGI concept is appropriate for year to year comparison.

### 1.1.2 Filing Requirements

Taxpayers are not required to file their tax returns if their AGI is below a certain amount, unless they pay other taxes such as payroll taxes or claim some credits (explained below). In 1996, married taxpayers (both under 65 of age) whose combined AGI is below $11.800 do not have to file a return. This is considered to be the minimum cost to cover the necessities of life. Each filing status has its own amount. The amount for married taxpayers filing jointly is not double that of heads of household or single taxpayers. This amount is indexed for inflation every year.

### 1.1.3 Exemptions and Deductions

Exemptions and deductions are amounts to be subtracted from the AGI to obtain the Taxable Income. Exemptions consist of a minimum AGI, taxpayers below which are exempt from taxation. This is a fixed amount (2.550$ for 1996) for every taxpayer and each dependent. In a sense, exemptions are subsidies designed to help the poor and middle class taxpayers. Although the exemption amount is the same for every taxpayers and each dependent, phaseout levels (the amount of AGI a taxpayer is no longer exempted) are different according to the filing status. Both are indexed for inflation every year.

Deductions are expenses, such as charitable donations, mortgage interests, state and local taxes and medical expenses among other things, that taxpayers are entitled to deduct from their AGI. Taxpayers can choose between a standard and an itemized deduction. The former is a given amount (which varies according to filing status) that taxpayers can deduct regardless if they qualify for any of the allowed deductions. The latter is given by the sum of the deductions they qualify for. Taxpayers choose the larger value between both. Note that some items are only deductible for expenses above a given amount or percentage of the resulting AGI. Some items are also subject to an overall limit of the total itemized deduction. For high income taxpayers above a given income level known as limitation threshold, the overall itemized deduction is reduced by the lesser of 3 percent of the amount of AGI in excess of the limitation threshold or 80 percent of the total of deductions subject to the overall limit. This provision applied until 2005. In 2006 and 2007, only 2/3 of the amount of reduction is subtracted to the itemized deductions. In 2008, only 1/3 of the original reduction is subtracted. The overall limit is completely repealed in

2010. Both the standard deduction and the limitation threshold for the overall limit are indexed for inflation every year.

### 1.1.4 Taxable Income (TI)

Taxable Income is by definition the AGI subtracted by exemptions and deductions. Tax liabilities are computed by the size of the TI and the filing status of taxpayers. The tax system is progressive. That is, marginal tax rates increase with TI. The TI is subdivided in tax brackets whereby statutory marginal tax rates apply. Although the marginal tax rates structure are the same for every taxpayers, tax brackets are different according to their filing status. That is, single taxpayers, married filing jointly, married filing separately and head of household have their own set of tax brackets. Changes to the marginal rates and brackets are the core interest of this thesis. Historical changes to these will be discussed in section 1.2 below.

### 1.1.5 Alternative Minimum Tax (AMT)

The AMT, introduced in 1978 in the US, is a system designed to prevent taxpayers from using deductions to reduce substantially their tax liabilities or pay no tax at all. That is, taxpayers could have deductions that allowed them to have zero taxable income in a particular year. The AMT is an alternative system to compute tax liabilities. Under the AMT, taxpayers are neither entitled to exemptions nor standard deductions from the regular income tax. Some deductions under regular income tax such as the state, local and foreign tax deductions, home mortgage interest or tax preparations fees... are no longer deductible. Some deductions, albeit deductible, have stricter and tighter rules. For example, medical expenses are only deductible for the amount that exceeds 10% of the overall AGI, while it is 7% under regular income tax. The exemption is reduced from the resulting alternative minimum taxable income (AMTI) to arrive at the final AMTI. The exemption amount depends on the filing status and size of the AMTI of the taxpayers. For each filing status, the allowed exemption is maximal for incomes below a threshold, above which it is beginning to phase out and totally phased out above certain income level. The exemption amounts and phase out intervals are indexed for inflation.

The AMT structure is also progressive in marginal tax rates. It has only two brackets, and hence two marginal tax rates, per filing status. Tax brackets are not indexed for inflation. They are only changed explicitly through tax law changes. This is one of the reasons why as years went by, an increasing number of middle class taxpayers end up paying AMT whereas its original purpose was to target rich taxpayers legally avoiding the regular income tax. In practice, around 3% of taxpayers pay AMT, totaling around 20 billion dollars of AMTI in 2006.

Due to this difference in structure and the small relative size compared to the regular income tax, we will not concern ourselves with the AMT in our study.

### 1.1.6 Tax Credits

The taxpayer's tax liability is the greater of his tax liabilities under the regular income tax and his AMT. This is also known as income tax before credits. Tax credits such as earned income credits, education credits or retirement savings contributions credits etc are amounts taxpayers can subtract to the tax liability if they qualify for them. The resulting value is the taxpayer's final tax payment. Tax credits are designed to help advance government targeted goals such as education, child care, energy... For a complete list, see Statistics of Income (2008). The difference between deductions and credits is that deductions are removed on the AGI whereas credits are removed directly from tax liabilities. In other words, deductions decrease the TI and therefore lower tax payments by the corresponding marginal tax rate whereas credits directly lower tax payments, that is a dollar for dollar reduction.

## 1.2 Tax Law Changes Between 1996 and 2007

The US individual income tax system in 1996 was set by the omnibus budget reconciliation act (OBRA) of 1993. The individual tax structure in 1996 is depicted in Table A.1 below. There were five marginal tax rates and corresponding tax brackets indexed annually for inflation yearly.

Personal exemptions, standard and itemized deductions and tax credits are as explained in the previous sections. They are adjusted for inflation yearly. Note that they are not increased uniformly and have different inflation indexation. The precise values are given in Table A.2 and A.3 in appendix A.

### 1.2.1 EGTRRA 2001

The first major change in the individual tax law occurred in May 2001 with the Economic Growth and Tax Relief Reconciliation Act (EGTRRA). We briefly describe here the major changes brought by EGTTRA to help understand and rationalize intertemporal changes in the distribution in the following chapters. In the following, we give some details that reveal the microscopic adjustments that are typical of US tax law. We note however that we may work, as applied mathematicians, with raw data that includes the effects of these detailed actions. It will be clear that they can lead to distortions of the distribution of revenues. For a very detailed description of EGTRRA one can read EGTRRA Law (2001).

#### Rate Changes

EGTRRA introduced a new 10% marginal rate tax bracket. For single taxpayers the first 6000$ of the first bracket in 2000 is taxed at 10% and the remaining taxed at the usual 15% rate. The new bracket threshold 6000$ is indexed for inflation every year until 2008 where it is changed to 7000$ and then is indexed again for inflation. Similar adjustments are made for other filing status.

EGTRRA reduced all statutory marginal income tax rates above the 15% bracket. The 28%, 31%, 36% and 39.6% rates are phased-down to 25%, 28%, 33% and 35% over six years, respectively. The rates are reduced 1% every two years except for the top marginal rate which is reduced to 38.6% in 2001, then 37.5% in 2004 and finally 35% starting 2006.

The 15% tax bracket threshold for married taxpayers filing separately was increased to be the same as for single taxpayers and for those filing jointly to be double the rate. This was phased in over four years starting in 2005, see Statistics of Income (2008); EGTRRA Law (2001). Similarly the standard deduction for married taxpayers filing jointly is increased to be double of the rate for married taxpayers filing separately, which in its turn is increased to the rate for single taxpayers.

#### Exemptions and deductions

As mentioned earlier, the overall limit on itemized deductions is progressively eliminated. This is enacted under EGTRRA. In 2006 and 2007, one third of the reduction in itemized deduction is discarded. This becomes two-thirds in 2008 and 2009 and is completely repealed by 2010. Similarly, the personal exemption phase-out is reduced by one third in 2006 and 2007. That is, a taxpayer cannot lose more than two thirds of his exemption regardless of the size of his AGI. Prior to 2006 a taxpayer could lose all his personal exemption if his AGI was above a certain level. In 2008 and 2009 the exemption phase-out is reduced by two-third and is completely ruled out by 2010.

The phase-out level for the student loan interest deduction is increased to $65,000 for single taxpayers and $130,000 for married filing jointly. The provision that required the interest paid in the first 60 months is repealed. The tuition and fees deduction is $3,000 for 2002 and 2003 with the phase-out AGI level $65,000 for single taxpayers and $130,000 for married filing jointly. This is increased to $4,000 for taxpayers with less than $65,000 AGI and $2,000 for single taxpayers between $65,000 and $80,000 ($130,000 to $160,000 for married filing jointly).

**Tax Credits**

The child tax credit was increased progressively to 1000$ over 10 years. This is refundable to the extent of 10% (15% starting 2005) of the earned income in excess of 10,000$ (indexed for inflation yearly) and is allowed to the full extent of both the regular income tax as well as the AMT. EGTRRA also increases the maximum adoption credit and excludes employer-provided adoption assistance from AGI, each up to 10,000$ with the AGI phase-out level at 150,000$ for children other than special needs children. This is also granted for special needs children regardless of whether the taxpayer qualified adoption expenses benefits, starting 2002.

In 2002, EGTRRA allowed dependent employer-related expenses up to $3000 ($6000 if more than one dependent) and increases the credit up to 35% with the AGI phasing out level at $15,000. A tax credit for qualified child care expenses is allowed, 25% for employee child care and 10% for child care resources and referral services. The maximum child care credit is $150,000.

For married taxpayers filing joint, the income tax credit phase-out interval (that is the start and end of the phase-out level) is increased by $1000 for 2002-2004, by $2000 for 2005-2007 and by $3000 in 2008, this is indexed for inflation onwards. Starting 2002, the AMT no longer reduced the amount of earned income tax credit.

### 1.2.2 JGTRRA 2003

The Jobs and Growth Tax Reconciliation Act (JGTRRA) of 2003 accelerated some of the EGTRRA provisions to be enacted by 2003. The bracket threshold for the 10% tax rate is increased to $7,000 for single and $14,000 for married taxpayers by 2003 instead of 2008 as originally scheduled by EGTRRA. These are indexed for inflation from 2004. Moreover, The marginal tax rate reductions scheduled for 2004 and 2006 are enacted by 2003. Thus, from 2003, the new tax rates are 10%, 15%, 25%, 28%, 33%, 35%. The applied tax structure, that is the first part of EGTRRA combined with JGTRRA, is given in Table A.4 in Appendix A.3.

The standard deduction for married taxpayers filing jointly is set to be twice that of the single taxpayers from 2003. The standard deduction for single taxpayers is indexed for inflation every year. Similarly, the 15% tax bracket upper threshold is set to be twice that of single taxpayers from 2003. This is indexed for inflation yearly.

The child tax credit is increased to $1,000 by 2003 instead of 2010 as scheduled in EGTRRA. This is not indexed for inflation.

The AMT exemption for married filing jointly is set to $58,000 and $40,250 for single taxpayers from 2003. This is not indexed for inflation.

## 1.3 Time dependence in tax revenue modeling

The first step in obtaining an optimal tax formula to control the revenue is to model the growth in revenue. Creedy and Gemmell (2006) suggest that the central parameter in tax revenue growth modeling is the elasticity of tax liability, be it at the individual level or aggregated. This is the percentage change of the revenue with respect to the variable whose effect on revenue is being investigated as a percentage of the percentage change in the inquired-variable. In other words, in what percentage would the revenue respond to a percentage change in the variable. For example, the revenue elasticity of the income tax with respect to change in income, using the notation of Creedy and Gemmell (2006), is given by:

$$\eta_{T_{z,z_i}} = \frac{dT(z_i)/T(z_i)}{dz_i/z_i} = \frac{z_i}{T(z_i)}\frac{dT(z_i)}{dz_i}, \tag{1.1}$$

where $z_i$ is a particular level of income and $T(z_i)$ is the corresponding tax liability at that income.

Given (1.1), economic theory would look into various effects that the individual could respond to, say tax parameters changes. There is a controversial debate as to what those effects are and how to quantify them. The clear alternative is that to assume that those effects are not new and have history. Tax revenue changes with time and that is because all factors affecting it change with time. There are many effects. Also, it is very hazardous to estimate those time changes effects at the individual level as everyone is different in one way or another. This suggests that a simpler model than that provided by economic theory should be feasible by observing year to year changes of the distribution as an aggregate level of all economic and behavioral effects of individuals. This is the main philosophy of this thesis. We will observe the taxpayers aggregate group responses to changes in taxation.

## 1.4 The Mirrlees Model

The Mirrlees model was introduced in the seminal paper: An exploration in the theory of optimal income taxation, in Mirrlees (1971) and it serves as the foundation for much of the modern economic theory of income taxation. In this model, each individual maximizes his utility $u(c,z)$ to choose the optimal combination of consumption $c$ and income $z$. Individuals differ in their skills level. With a number of hours worked $l$, an individual with a skill $n$ earns $z = nl$. Given the tax structure imposed by the government, namely by paying $T(z)$ of taxes, an individual consumes $c = z - T(z)$. The optimal utility level for an $n$-skilled individual follows the ODE:

$$\frac{du_n}{dn} = -\frac{lu_l}{n}, \tag{1.2}$$

where $u_n = u(c_n, nl)$ the utility of an $n$-skilled individual and $u_l$ is the derivative of the utility function $u$ with respect to its second argument.

With the density distribution of skill $f(n)$, the government imposes taxes on every individual by maximizing a social welfare function in such a way as to cover its expenses. The social welfare function $G(u_n)$ depends on the utility of each individual and formulates the government objective for redistribution. That is, increasing the utility of certain classes of the population by decreasing those of others. The resource constraints of the economy simply states that the aggregate production, $\int_0^\infty z_n f(n)dn$, has to be greater than the aggregate consumption, $\int_0^\infty c_n f(n)dn$, together with the government expenditures $R$. The problem is given in the following optimal control problem:

$$\begin{aligned} &\max_{T(z_n)} \int_0^\infty G(u_n) f(n) dn \\ &\text{subject to :} \\ &\begin{cases} \dfrac{du_n}{dn} = -\dfrac{lu_l}{n}, \text{and} \\ \displaystyle\int_0^\infty T(z_n) f(n) dn \geq R. \end{cases} \end{aligned} \tag{1.3}$$

As pointed out by Saez (2001), the optimal marginal rate is derived by maximizing the Hamiltonian $H(T(z_n)) = [G(u_n) - pT(z)n)] f(n) - \phi(n) \frac{l_n u_l(c_n, nl_n)}{n}$, with $p$ and $\phi$ the Lagrangian multipliers.

Tuomala (1990) reports most of the analytical results derived from the Mirrlees model. Mirrlees (1971) and Seade (1982) proved the interesting result that the marginal tax rate has to be between zero and hundred percent. Intuitively, above hundred percent marginal rate would discourage any working activity compared to leisure since no income is perceived either way.

Positive marginal rate implies that there should not be any subsidies in the optimal social welfare. As mentioned earlier, Sadka (1976) and Seade (1977) proved that while the distribution of skill is bounded above the most able individual should have a zero marginal tax rate. Not taxing additional income from the top skilled individual will not decrease the tax revenue but will make him better off.

Saez (2001) thoroughly assessed the various effects on the social welfare function in response to changes in the tax rate. Individual responses at each income level are categorized into mechanical, elastic and income effects. Those are expressed in terms of well-known economic quantities in the public finance arena and then aggregated to the overall distribution. One of the merits of his approach is that the optimal marginal rate, though equivalent to the Mirrlees result, involves the compensated and uncompensated elasticity of earnings. The tail of the income/skill distribution also plays an important role in his formula. However, the literature shows a wide range of values for the elasticity of earnings. From the early estimate of Lindsey (1987) to the recent studies of Saez (2004), the elasticity of earnings ranges from 1.6 to 0.2. In addition, the distinction between and compensated and uncompensated elasticity estimates was not made until Saez (2004). As a result, the Saez (2001) optimal marginal rate falls in the general guidance and academic development of the theory of income taxation but contributes little to implementable policy.

## 1.5 Mirrleesian optimal linear taxation

The theory of optimal linear income taxation is derived from the Mirrlees model by explicitly specifying the analytical form of the tax liability at the individual level to $T(z) = tz + k$ where $t$ is the linear marginal tax rate and $k$ is a lump sum tax or a subsidy depending on its sign. The optimal control problem in equation (1.3) is transformed into a simple optimization problem with $t$ and $k$ the unknown optimizers. A large literature is devoted into this setting as the original Mirrlees (1971) showed that the optimal nonlinear tax rate turns out to be approximately linear. To cite a few, Stern (1976) presents a thorough discussion and gives numerical results for a number of realistic elasticity of labor supply models. Tuomala (1985) derived the discrete version of the Mirrlees model applied to linear income taxation.

The linear income taxation has only one bracket and, as such, is an oversimplification of the actual bracket system used in practice. Very few researchers focused on the issue of optimal brackets. This is perhaps because of the results obtained by Sheshinski (1989) and Slemrod (1994) were complete and thus closed the debate. They discussed a two-bracket system and showed that optimal marginal rates were zero in one of the brackets. Thus, optimally a simple linear tax schedule is enough. This is expected as the nonlinear income tax rate in the Mirrlees model itself is approximately linear. Hence the optimal income tax theory goes against the public view of tax progressivity. That is to say, though progressive in the marginal rates and perhaps democratically accepted, the actual tax system applied does not necessarily imply optimal when balancing efficiency and equity. This is where the idea of optimal tax reforms comes in.

## 1.6 Non-welfarist optimal taxation

This theory departs from the Mirrlees framework by maximizing non-welfare *social evaluation* functions. This is a general function that depends explicitly on individual incomes and consumptions and possibly but not necessarily their utilities. This allows for non-individualistic preferences, unlike the Mirrlees approach where the objective function depends only on the individual utilities. However, the general structure of the resulting optimal tax is the same.

Seade (1980) pioneered this generalization of the Mirrlees setup. Other researchers followed with specific non-utilitarian objectives. For example, Kanbur et al. (1994) looked at poverty alleviation as its objective. Optimal non-welfarist mixed taxation is derived from the same structure where commodity taxation and income taxation are analysed altogether, Pirttila and Tuomala (2004); Kanbur et al. (2006).

## 1.7 Large-Scale Microsimulation Model

All four American government budget agencies, that is, the office of management and budget, the congressional budget office (CBO (2007b)), the joint committee on taxation (JCT (1992)) and the office of tax analysis in the treasury department (OTA (2008)), use a microsimulation model to evaluate proposed changes to the tax system or to identify the direction of improvement on the current system. Their model is not publicly available. However they all follow the same general principle as thoroughly explained in Creedy and Duncan (2002) and Creedy and Herault (2009) that we explain briefly.

A sample of cross-sectional taxpayer data representing the whole taxpayer population is used to simulate each taxpayer's change in labor supply to changes in tax rates. That is, for each taxpayer how many hours is he going to work given his new after tax wage. This assumes a particular utility function and a particular budget constraint. Once his new hours are worked, the corresponding change in total revenue automatically follows by aggregating the sample with the corresponding weights. The change in welfare can be computed as well. Again this assumes a particular type of welfare function representing value judgments of the government regarding redistribution.

Naturally, the general form would consider the source of income from all activities and not just labor. In addition, any changes to any tax parameters such as deductions, brackets and credits can be investigated. This is done by choosing the appropriate utility function (or elasticity) that depends on the desired variables. Optimality is achieved when a change in rates equates the changes in total revenue and the changes in welfare. That is, when the government is indifferent between taking one additional dollar from a taxpayer (with higher utility) to redistribute it to someone else (with lower utility).

## 1.8 Taxpayer's responses to changes in tax rates and the Laffer curve

Taxpayer's responses to changes in tax rates have long fascinated public economists. The parameter of interest here is the elasticity of taxable income with respect to marginal tax rates. This is the percentage change in taxable income due to a percentage change in the marginal tax rate. There is a vast empirical literature on the estimation of this parameter.

The elasticity of taxable income allows one to appreciate the precise strength of taxpayers responses. Low elasticity for some income range implies taxpayers over that range are insensitive to changes in tax rates and inversely. For labor income, the economic concept of elasticity divides taxpayer responses to change in rates in two effects, see Pencavel (1986), Saez (2001),

* Income effects: an increase in the tax rates reduces the taxpayer's after-tax income. Therefore, he has to work more to earn the same wages as before the increase in tax rates.

* Substitution effects: an increase in tax rates increases the price of labor and decreases the price of leisure. As his after-tax wage is decreased he is more reluctant to work additional hours as compared to before the tax increase.

These two effects are measured by the compensated and uncompensated elasticity of labor income. Whichever effect is stronger yields the taxpayer's responses to changes in tax rates.

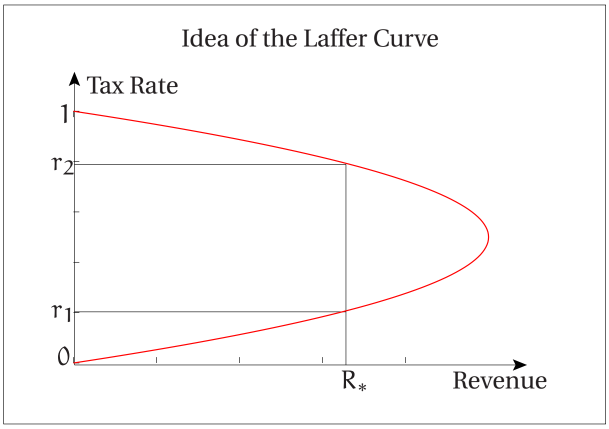



Figure 1.1: *Illustration of the idea of Laffer Curve. We can see that both 0% and 100% yield no revenue, both* $r_1$ *and* $r_2$ *yield the same revenue* $R_*$.

Either income effect is larger (usually the case for low income taxpayers) and the labor supply is increased and results in more revenues or the substitution effect is larger, decreasing labor supply and results in less revenues. The consequences of both responses on the revenue are captured in what is called the Laffer curve.

The Laffer curve stems from an observation of Laffer (2004, 1980) that both zero and hundred percent tax rates yield no revenue. Hence, for any achievable revenue, at least two tax rates leads to that same revenue. This is illustrated in Figure 1.1. As the tax liability is linearly proportional to tax rates, increasing tax rates increases the revenue up until the point where further increase in tax rates results in a decrease in revenue. High tax rates constitute a disincentive to earn income (higher substitution effect) and thus reduce revenue. This relationship is referred to as the *Laffer effect* in this thesis. The inverse, whereby high tax rates leads to high revenue, is referred to as the *anti-Laffer effect*.

## 1.9 Thesis outline

In chapter 2 we explain the technique we used to estimate the distribution of revenue given its data requirements. We use a cubic spline in the low and middle income range. A Pareto distribution is fitted at the upper tail, following Pareto's theory and Saez (2001)'s confirmation of the Pareto tail for the US income distribution. We derive the fundamental equations such that the total revenue per income bin is constrained by the data and the overall curve is smooth. Nondimensionalization and normalization is introduced to allow year to year comparisons of estimated curves as well as improve numerical accuracy of the estimation. We call this procedure the extended cubic splines. A bound of the relative error using this technique is given.

In chapter 3, we apply the theory of extended cubic splines to the US joint returns tax data from 1996 to 2008. We show that the peak of the normalized revenue distribution consistently decreased during the years before the Economic Growth and Tax Relief Reconciliation Act of 2001 (EGTRRA) and after it was fully phased in. Normalization shows that the share burden of the high income taxpayers increased year after year. The opposite trend is observed during the phase-in period of EGTRRA. The distribution of number of returns on the other hand, does not display such a drastic and systematic trend.

In chapter 4, we provide a rational control of the consistent drop of the normalized revenue at the peak. We derive the optimal tax liability and corresponding tax parameters, brackets and marginal rates, that keep future normalized revenue to the current one with a total revenue target. The simple case where the distribution of number of returns does not depend on the tax liability is derived first. We show for the US data from 1996 to 2008 that a reasonable control is achievable. The general case where the number of returns depends on the tax liability is considered as well. This is the case where there are responses in the number of returns distribution to changes in tax liability. Reasonable control is achieved for the US data.

In chapter 5, we give consideration to fairness of the tax system. We derive the optimality conditions for the year to year control of the instability in the normalized revenue in such a way that it keeps the change in average tax rate constant for all income. Both cases where the number of returns is independent and dependent on the tax liability will be derived. Application on the US joint returns from 1996 to 2008 that a reasonable and improved control can be achieved.

In the final chapter, conclusions from the thesis are given with discussions of future work.

# Chapter 2

# Estimating Personal Income Tax Revenue Distribution Using Cubic Splines and a Pareto Distribution

## 2.1 Introduction

The maximum likelihood (Peck et al. (2001)) and Bayesian methods (Lee (2004)) are the conventional approach to statistical fitting of distributions. The former assumes the data, corrupted with noise, stem from a known distribution with unknown parameters whereas the latter treats the unknown parameters as random variables with prior probabilities. The general idea for both is to use the data to estimate the parameters and consequently recover the underlying distribution. A major problem therefore arises when there are no consistent theories as to which known statistical distribution describes best the distribution to be fitted. This is the case for the distribution of tax revenues.

There has not been much discussion about the shape of the distribution of individual income tax revenue. As introduced in the previous chapter, major parts of the literature in economics of income taxation followed the seminal work of Mirrlees (1971) that used a lognormal distribution. However, Mirrlees in his derivation merely assumed a well-known distribution that could describe the distribution of skills and income. There was no serious attempt to derive, estimate and interpret the observed revenue distribution. There are two main reasons for this. First, the overall theory of income taxation focused exclusively on the behavioral responses of taxpayers, through utility and welfare functions, to changes in the tax structures and the distribution of income was just used to aggregate the overall effect. The distribution of revenue has been analyzed in the applied world but has not attracted attention from theorists. A new alternative that will be explored in this chapter is to estimate the overall response to taxation by observing the dynamics of the distribution of tax revenues directly. We arrive at a phenomenological revenue distribution that is its own natural explanation of revenues.

Second, the scarcity of the revenue data and the structures it is provided with, made it difficult if not impossible to give more than an understanding and a conventional compromise on the nature of the distributions of income and revenue. The statistics of Income that is the source of data provided by the Internal Revenue Service (IRS) for the US data only gives the revenue collected per intervals of income. Using statistical estimation on this type of data is heavily dependent on the assumed distribution and is computationally expensive if no prior knowledge of the parameters are assumed.

The contribution of this chapter is to provide a general technique to estimate the normalized tax revenue distribution given data requirements and constraints. We will have no interest in probability and will not refer to probability distribution. However, for the purpose of comparing distributions, the areas under the observed revenue (and returns) distributions are normalized to unity. We fit cubic splines on the middle income range of the distribution by satisfying the revenue constraints given by the data. No statistical assumption on the data is required in this case. The idea here is to use well-known and computationally cheap numerical techniques to let the data reveal the shape of the underlying distribution. Special care has to be taken on both ends of the distribution to bypass adverse drawbacks of numerical techniques. After many attempts, we found it best to use a Pareto distribution at the upper tail and a polynomial fit at the lower tail of the distribution. This mixture of distributions and techniques we name extended cubic spline estimation. We explore the merits of this approach in this chapter. Applications to the actual US data will be done in the next chapter.

This chapter is organized as follows. Section 2 presents the assumptions and requirements from the US data. Fundamental equations defining the extended cubic spline estimation is derived in section 3. The treatment of both the lower tail with a polynomial fit and a Pareto distribution for the upper tail of the distribution follow respectively in section 4 and 5. In section 6, links to the well-known theory of clamped cubic spline is discussed. We will be interested year to year comparison of distribution and to this end we introduce nondimensionalization and rescaling of variables in section 7. General curve fitting performance of the extended cubic spline approximation using lognormal test functions are presented in section 8. In section 9, we investigate the sensitivity of the approximation with respect to changes in the Pareto tail of the distribution. The chapter ends up with a brief conclusion in section 10.

## 2.2 Description and Preparation of Data

For the case that we investigate in this thesis, individual income tax data comes in histogram bins for the period 1996 to 2008. We have the number of returns per income interval per year together with the corresponding tax revenue. Table 2.1 depicts the US joint returns, tax revenue (in thousands of \$) and the corresponding number of returns, classified per level of adjusted gross income (in \$), for the year 1998, 2002 and 2007. For example in 2002, there were 2005 taxpayers with adjusted gross income between 1\$ and 5000\$ who paid \$433.000 in total in income tax. The figures in this table are based on a sample from the actual taxpayer population. The data for year 1998 is less detailed than those of 2002 and 2007. The revenues and number of returns for taxpayers above \$1 million is combined in 1997 (total revenue in this range is approximately equal to \$ 120.033 millions) whereas for 2002 and 2007, it is broken down into more income thresholds.

We wish to develop a continuous revenue curve $R(z)$ and the number of returns curve $N(z)$ as functions of income $z$. Thus, in each year, we seek functions $f(z), g(z)$ such that revenues $R(z) = \int_0^z f(z)\,dz$, $N(z) = \int_0^z g(z)\,dz$. Income and revenue are given in discrete integer values and its range is large (from 0 to hundreds of millions), it is usual to assume a continuous distribution. The same argument can be used for the number of returns and the tax revenue distribution. Hence, the problem is to estimate a continuous distribution given the histogram data.

The data has to be further transformed before we determine $f$. Inflation indexation is necessary as the revenue amounts given in the data are in nominal terms. Thus, any year to year comparison of the distribution will not be carried on the same money values. The data has to be expressed in terms of a constant dollar.

In this thesis, we express the money dimension in terms of the 1996 US dollar. That is, we set the 1996 figures as base 100 and we index the remaining values to the 1996 CPI-U rate. We have, for example in Table 2.2 that one dollar in 1996 is worth 96.26 cents in 1998. All money value in Table 2.1 is deflated using the above table.

Furthermore, we find undersampling on low income together with the total deductions on joint returns and therefore we discard intervals below the 10.000\$ AGI. Thus, combining inflation indexation and low income adjustment we arrive at Table 2.3.

The first column has to be deflated per year as well. Full deflated income ranges for all years are given in Table A.13 in the appendix. The raw and adjusted data for the full period 1995-2008 is given in the appendix as well.

| AGI Range | Income Tax After Credit | | | Number of Returns | | |
|---|---|---|---|---|---|---|
| | **1998** | **2002** | **2007** | **1998** | **2002** | **2007** |
| 1 | 0 | 433 | 1714 | 0 | 2005 | 1652 |
| 5000 | 0 | 484 | 3005 | 0 | 1046 | 1138 |
| 10000 | 63749 | 5887 | 364 | 436181 | 88780 | 752 |
| 15000 | 885217 | 335875 | 25622 | 1577940 | 1222199 | 253851 |
| 20000 | 2040473 | 848895 | 307318 | 1836516 | 1287690 | 902130 |
| 25000 | 3443385 | 1452265 | 674514 | 2196305 | 1413266 | 906688 |
| 30000 | 12869065 | 6814913 | 3368608 | 5349047 | 3933032 | 2490405 |
| 40000 | 20178531 | 12634169 | 6843230 | 5595343 | 4728996 | 3132765 |
| 50000 | 71667361 | 56428611 | 40518383 | 11716060 | 11494860 | 10052491 |
| 75000 | 68807911 | 70504909 | 61005220 | 6130908 | 7536872 | 8752815 |
| 100000 | 119263879 | 144238598 | 181369966 | 5367156 | 7187168 | 11161429 |
| 200000 | 94186359 | 106048961 | 164078817 | 1383004 | 1634831 | 2946286 |
| 500000 | 49055531 | 54534220 | 87510328 | 258898 | 289023 | 546882 |
| 1000000 | 120033097 | 23123307 | 40761589 | 143371 | 65254 | 138388 |
| 1500000 | - - | 13486253 | 24614555 | - - | 26373 | 58464 |
| 2000000 | - - | 32385619 | 64779947 | - - | 36735 | 89523 |
| 5000000 | - - | 16442409 | 36047029 | - - | 8188 | 22762 |
| 10000000 | - - | 25730797 | 88165288 | - - | 4169 | 14852 |

Table 2.1: *US Joint returns, tax revenue and number of returns distribution for year 1998, 2002 and 2007. Source: IRS Statistics of Income*

| | **1996** | **1998** | **2002** | **2007** |
|---|---|---|---|---|
| CPI-U | 100 | 103.89 | 114.66 | 132.15 |
| Deflator factor | 1 | 0.9626 | 0.8722 | 0.7567 |

Table 2.2: *Consumer prince index and corresponding deflator factor for 1996,1998, 2002, 2007. Source: US bureau of labor statistics.*

| AGI Range | Income Tax After Credit | | | Number of Returns | | |
|---|---|---|---|---|---|---|
| | 1998 | 2002 | 2007 | 1998 | 2002 | 2007 |
| 10000 | 66227 | 6750 | 481 | 436181 | 88780 | 752 |
| 15000 | 919633 | 385111 | 33859 | 1577940 | 1222199 | 253851 |
| 20000 | 2119803 | 973335 | 406118 | 1836516 | 1287690 | 902130 |
| 25000 | 3577258 | 1665153 | 891364 | 2196305 | 1413266 | 906688 |
| 30000 | 13369392 | 7813912 | 4451586 | 5349047 | 3933032 | 2490405 |
| 40000 | 20963037 | 14486214 | 9043270 | 5595343 | 4728996 | 3132765 |
| 50000 | 74453664 | 64700492 | 53544695 | 11716060 | 11494860 | 10052491 |
| 75000 | 71483043 | 80840237 | 80617873 | 6130908 | 7536872 | 8752815 |
| 100000 | 123900652 | 165382561 | 239678850 | 5367156 | 7187168 | 11161429 |
| 200000 | 97848161 | 121594698 | 216828745 | 1383004 | 1634831 | 2946286 |
| 500000 | 50962725 | 62528401 | 115644145 | 258898 | 289023 | 546882 |
| 1000000 | 124699776 | 26512957 | 53866089 | 143371 | 65254 | 138388 |
| 1500000 | - - | 15463205 | 32527923 | - - | 26373 | 58464 |
| 2000000 | - - | 37133033 | 85606143 | - - | 36735 | 89523 |
| 5000000 | - - | 18852705 | 47635839 | - - | 8188 | 22762 |
| 10000000 | - - | 29502679 | 116509669 | - - | 4169 | 14852 |

Table 2.3: *Example of inflation adjusted US Joint returns, tax revenue and number of returns distribution for year 1998, 2002 and 2007. Numbers are rounded to one dollar. Income intervals below $10.000 are consolidated for purposes of comparison.*

## 2.3 Extended Cubic Splines

A cubic spline is a function made up of piecewise cubic polynomials. It is widely used to approximate functions that are only known at a finite set of data points, Plato (2003); Wang (2010). A cubic polynomial is fitted between two consecutive points. In addition, the first and second derivative of any two consecutive cubic polynomials must have the same value at these data points. Hence, the cubic spline not only fits the data but it is continuous and has continuous first and second derivatives over its domain.

Fitting a distribution from the histogram data is not a straightforward application of the cubic spline formulas. The difference here is that we are given the histogram data instead of the probability density function value at each point. At each income level in the data points, instead of the corresponding revenue values, we have the cumulated revenues collected from taxpayers between two levels of income. In other words, instead of the tax revenue density function values, we have the integrals between two successive points. The cubic spline has to be extended to account for this change in the structure of the data.

Between two points of an income interval, we construct a cubic polynomial so as to preserve the amount of revenue collected within that interval. That is, the area below the polynomial is the same as the corresponding amount of revenue from the data. In addition, we add the constraint that the density function as well as its first and second derivative is continuous over the entire income distribution. That is, two consecutive piecewise cubic polynomials have the same function value, derivative and second derivative at their intersection point.

We follow Plato (2003) to derive the cubic spline formulation for the tax revenue distribution. Deriving the density for the number of returns is obtained the same way by using the distribution of number of returns histogram data given in Table A.15 in appendix A.5. The following technique is then well-known.

Let the different income thresholds (income intervals from the histogram data points of Table 2.3) be given by the sequence $z_k,\ k = 1, \ldots, N$. Let $y_k$ and $x_k, \quad k = 1, \ldots, N$ be the corresponding density function value and its second derivative at $z_k$, respectively. $x_k$ define curvatures at the $k$-th data point. Let $S_k(.)$, $h_k$ and $R_k, \quad k = 1, \ldots, N$ respectively be the cubic polynomial to be fitted, the step size and amount of revenue between $z_{k+1}$ and $z_k$, for $k = 1, \ldots, N-1$. That is, $h_k = z_{k+1} - z_k$ and $R_k = \int_{z_k}^{z_{k+1}} S_k(z)\, dz, \quad k = 1, \ldots, N-1$. Recall that $R_k$ and $z_k$ are given data as taken from Table 2.1. These notations are illustrated in Figure 2.1.

Since we impose continuity of the second derivative of the spline, the second derivative passes

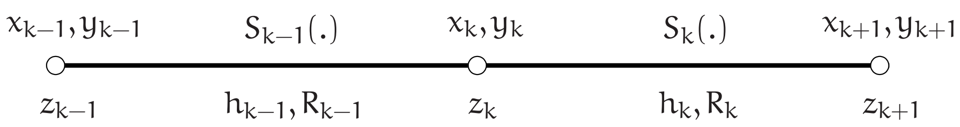



Figure 2.1: *Configuration of notations, on two consecutive intervals, on the income axis.*

through the points $(z_k, x_k)$ and $(z_{k+1}, x_{k+1})$. That is,

$$S_k''(z_k) = x_k \quad \text{and} \quad S_k''(z_{k+1}) = x_{k+1}. \tag{2.1}$$

Because $S_k(.)$ is a cubic polynomial, its second derivative is a polynomial of degree one. From equation (2.1), we have,

$$S_k''(z) = \frac{x_{k+1}}{h_k}(z - z_k) + \frac{x_k}{h_k}(z_{k+1} - z). \tag{2.2}$$

Taking the anti-derivative on both sides yields:

$$S_k'(z) = \frac{x_{k+1}}{2h_k}(z - z_k)^2 - \frac{x_k}{2h_k}(z_{k+1} - z)^2 + C_k, \tag{2.3}$$

where $C_k$ is a constant of integration.

Similarly, from equation (2.3),

$$S_k(z) = \frac{1}{3!}\frac{x_{k+1}}{h_k}(z - z_k)^3 + \frac{1}{3!}\frac{x_k}{h_k}(z_{k+1} - z)^3 + C_k z + D_k, \tag{2.4}$$

where $D_k$ is a constant of integration. Thus, $S_k$ is fully determined by the sequence of second derivatives $x_k$ and the constants $C_k$ and $D_k$.

To preserve the area under $S_k$ between $z_k$ and $z_{k+1}$ to be the amount of revenue given by the data, we have,

$$\begin{aligned} R_k &= \int_{z_k}^{z_{k+1}} S_k(z)\,dz \\ &= \left[\frac{1}{4!}\frac{x_{k+1}}{h_k}(z - z_k)^4 - \frac{1}{4!}\frac{x_k}{h_k}(z_{k+1} - z)^4 + \frac{C_k}{2}z^2 + D_k z\right]_{z=z_k}^{z=z_{k+1}}, \end{aligned} \tag{2.5}$$

that is,

$$R_k = \frac{1}{4!}h_k^3 x_{k+1} + \frac{1}{4!}h_k^3 x_k + \frac{C_k}{2}(z_{k+1}^2 - z_k^2) + D_k h_k. \tag{2.6}$$

This is one equation with four unknowns $x_k, x_{k+1}, C_k, D_k$. To decrease the degrees of freedom, we look at the other constraints we impose on the overall spline.

The second derivative is continuous at each point $z_k$ by construction. However, continuity of the first derivative at $z_k$ implies that $S_{k-1}'(z_k) = S_k'(z_k)$. From equation (2.3), we have,

$$\begin{cases} S_k'(z_k) = & -\frac{1}{2}x_k h_k + C_k \\ S_{k-1}'(z_k) = & \frac{1}{2}x_k h_{k-1} + C_{k-1} \end{cases} \tag{2.7}$$

Therefore, we have the recursion,

$$C_k = \frac{1}{2} x_k (h_k + h_{k-1}) + C_{k-1}. \tag{2.8}$$

Similarly, continuity of $S_k$ at $z_k$ implies that $S_k(z_k) = S_{k-1}(z_k)$. From equation (2.4), we have,

$$\begin{cases} S_k(z_k) = & \frac{1}{3!} x_k h_k^2 + C_k z_k + D_k \\ S_{k-1}(z_k) = & \frac{1}{3!} x_k h_{k-1}^2 + C_{k-1} z_k + D_{k-1}. \end{cases} \tag{2.9}$$

Using equation (2.8), we have the recursion for the sequence $D_k$,

$$D_k = \frac{1}{3!} x_k (h_{k-1}^2 - h_k^2) - \frac{1}{2} z_k x_k (h_k + h_{k-1}) + D_{k-1}. \tag{2.10}$$

Now, from equation (2.6), we have,

$$\begin{cases} \frac{R_k}{h_k} & = \frac{1}{4!} h_k^2 x_{k+1} + \frac{1}{4!} h_k^2 x_k + \frac{C_k}{2h_k} (z_{k+1}^2 - z_k^2) + D_k \\ \frac{R_{k-1}}{h_{k-1}} & = \frac{1}{4!} h_{k-1}^2 x_k + \frac{1}{4!} h_{k-1}^2 x_{k-1} + \frac{C_{k-1}}{2h_{k-1}} (z_k^2 - z_{k-1}^2) + D_{k-1}. \end{cases} \tag{2.11}$$

Replacing $D_k$, using equation (2.10), we eliminate $D_{k-1}$ by subtracting the first and second equation in equation (2.11). That is,

$$\frac{R_k}{h_k} - \frac{R_{k-1}}{h_{k-1}} = \frac{1}{4!} h_k^2 x_{k+1} + t_k x_k - \frac{1}{4!} h_{k-1}^2 x_{k-1} + C_k \frac{u_k}{h_k} - C_{k-1} \frac{u_{k-1}}{h_{k-1}}, \tag{2.12}$$

where $t_k = (\frac{1}{4!} - \frac{1}{3!}) (h_k^2 - h_{k-1}^2) - \frac{z_k}{2} (h_k + h_{k-1})$ and $u_k = \frac{1}{2} (z_{k+1}^2 - z_k^2)$.

Similarly, replacing $C_k$ and successively $C_{k-1}$ in equation (2.12), we eliminate $C_{k-2}$ by subtracting it with its equivalent first recursive form. That is, equation (2.12) evaluated at $k-1$. Finally, we end up with a set of linear equations in the $x_{k+1}, x_k, x_{k-1}, x_{k-2}$ unknowns given by equation (2.12). The overall formulation is depicted in Table 2.4.

The second difference equation in Table 2.4 results in a system of linear equations in the curvatures $x_k$. Since the equation involves four consecutive terms, $x_{k+1}, x_k, x_{k-1}$ and $x_{k-2}$, three additional equations are required to obtain a unique solution. These equations are provided by the continuity constraints of the Pareto tail of the distribution. As argued earlier, income distribution shows long Pareto-like tails. Continuity of the density function as well as the first and second derivative implies exactly three more equations have to be satisfied.

Let us assume that the distribution follows the Pareto distribution from a certain income level $z_p$. The derivation of the Pareto parameter will be explained in the next section (see equation (2.23) ) but let us assume that the Pareto distribution is well-defined. Let the function value at $z_p$ be $f_p$, its first derivative $f'_p$ and second derivative $f''_p$. We have,

**Notations:**

$h_k = z_{k+1} - z_k$

$u_k = \frac{1}{2}(z_{k+1}^2 - z_k^2)$

$a_k = \frac{u_k}{h_k} - \frac{u_{k-1}}{h_{k-1}}$

$t_k = (\frac{1}{3!} - \frac{1}{4!})\,(h_{k-1}^2 - h_k^2) - \frac{1}{2} z_k\,(h_{k-1} + h_k)$

$b_k = t_k + \frac{1}{2}\,(h_{k-1} + h_k)\,\frac{u_k}{h_k}$

$m_k = \frac{-b_{k-1}}{a_{k-1}} - \frac{1}{4!}\,\frac{h_{k-1}^2}{a_k} + \frac{1}{2}\,(h_{k-1} + h_k)$

$n_k = \frac{-b_k}{a_k} - \frac{1}{4!}\,\frac{h_{k-1}^2}{a_{k-1}}.$

**Revenue Constraint:**

$\frac{R_k}{h_k} = \frac{1}{4!}\,h_k^2\,x_{k+1} + \frac{1}{4!}\,h_k^2\,x_k + C_k\,\frac{u_k}{h_k} + D_k.$

**First Difference:**

$G_k := \frac{R_k}{h_k} - \frac{R_{k-1}}{h_{k-1}} = \frac{1}{4!}\,h_k^2\,x_{k+1} + b_k\,x_k - \frac{1}{4!}\,h_{k-1}^2\,x_{k-1} + C_{k-1}\,a_k.$

**Second Difference:**

$O_k := \frac{G_k}{a_k} - \frac{G_{k-1}}{a_{k-1}} = \frac{1}{4!}\,\frac{h_k^2}{a_k}\,x_{k+1} + n_k\,x_k + m_k\,x_{k-1} + \frac{1}{4!}\,\frac{h_{k-2}^2}{a_{k-1}}\,x_{k-2}.$

$C_k$ **and** $D_k$ **recurrence:**

$C_{k-1} = C_k - \frac{1}{2}\,(h_k + h_{k-1})\,x_k$

$D_{k-1} = D_k - (\frac{1}{3!}\,(h_{k-1}^2 - h_k^2) - \frac{1}{2} z_k\,(h_k + h_{k-1}))\,x_k.$

**Initial Conditions:**

$C_{p-1} = f_p' - \frac{1}{2}\,x_p\,h_{p-1}$

$D_{p-1} = f_p - f_p' z_p + (-\frac{1}{3!}\,h_{p-1}^2 + \frac{1}{2}\,h_{p-1}\,z_p)\,x_p$

Table 2.4: *Extended Cubic Spline fundamental equations.*

$$\begin{cases} f_p = S_{p-1}(z_p) = \dfrac{1}{3!} h_{p-1}^2 \, x_p + C_{p-1} z_p + D_{p-1} \\ f'_p = S'_{p-1}(z_p) = \dfrac{1}{2} x_p \, h_{p-1} + C_{p-1} \\ f''_p = S''_{p-1}(z_p) = x_p. \end{cases} \tag{2.13}$$

These yield,

$$\begin{cases} C_{p-1} = & f'_p - \frac{1}{2} x_p \, h_{p-1} \\ D_{p-1} = & f_p - \frac{1}{3!} x_p \, h_{p-1}^2 - C_{p-1} z_p = f_p - f'_p z_p + (-\frac{1}{3!} h_{p-1}^2 + \frac{1}{2} h_{p-1} z_p) \, x_p. \end{cases} \tag{2.14}$$

Substituting $C_{p-1}$ and $D_{p-1}$ in the revenue matching constraint in Table 2.4, yields the following equation:

$$\frac{R_{p-1}}{h_{p-1}} - f_p + f'_p \, (z_p - \frac{u_{p-1}}{h_{p-1}}) = [(\frac{1}{4!} - \frac{1}{3!}) \; h_{p-1}^2 - \frac{u_{p-1}}{2} + \frac{1}{2} h_{p-1} z_p] \, x_p + \frac{1}{4!} \, h_{p-1}^2 x_{p-1}. \tag{2.15}$$

Similarly, equation (2.14), (2.8) and (2.10) yield,

$$\begin{cases} D_{p-2} = & f_p - f'_p z_p + (-\frac{1}{3!} h_{p-1}^2 + \frac{1}{2} h_{p-1} z_p) \, x_p \\ & + (-\frac{1}{3!} (h_{p-2}^2 - h_{p-1}^2) + \frac{1}{2} z_{p-1} (h_{p-1} + h_{p-2})) \, x_{p-1} \\ C_{p-2} = & f'_p - \frac{1}{2} h_{p-1} x_p - \frac{1}{2} (h_{p-1} + h_{p-2}) \, x_{p-1}. \end{cases} \tag{2.16}$$

It follows from equation (2.6) that

$$\begin{aligned} \frac{R_{p-2}}{h_{p-2}} - f_p + f'_p \, (z_p - \frac{u_{p-2}}{h_{p-2}}) = & [-\frac{1}{2} h_{p-1} \frac{u_{p-2}}{h_{p-2}} - \frac{1}{3!} h_{p-1}^2 + \frac{1}{2} h_{p-1} z_p] \, x_p \\ & + [\frac{1}{4!} \, h_{p-2}^2 - \frac{1}{2} (h_{p-1} + h_{p-2}) \, \frac{u_{p-2}}{h_{p-2}} \\ & \quad - \frac{1}{3!} (h_{p-2}^2 - h_{p-1}^2) + \frac{1}{2} z_{p-1} (h_{p-1} + h_{p-2})] \, x_{p-1} \\ & + \frac{1}{4!} h_{p-2}^2 x_{p-2}. \end{aligned} \tag{2.17}$$

The last equation (2.13), that constrains the curvature $x_p$ at $z_p$ to be the Pareto tail curvature, that is $x_p = f''_p$ should in theory provide the third equation. Unfortunately, in practice, the US data makes it numerically very unstable (the conditioning number of the corresponding matrix very large). Instead, the third equation is obtained by imposing a more comprehensive constraint that the revenue must be zero at zero taxable income, or equivalently at $z_1$ (the standard deduction) income. From equation (2.4), we have

$$S_1(z_1) = \frac{1}{3!} h_1^2 \, x_1 + C_1 z_1 + D_1 = 0. \tag{2.18}$$

Hence, equation (2.6) yields

$$\frac{R_1}{h_1} = \frac{1}{4!}\,h_1^2 x_2 - \frac{3}{4!}\,h_1^2 x_1 + C_1\,[\frac{u_1}{h_1} - z_1]. \tag{2.19}$$

The first difference equation in Table 2.4 implies,

$$\frac{G_2}{a_2} = \frac{1}{4!}\,\frac{h_2^2}{a_2} x_3 + \frac{b_2}{a_2} x_2 - \frac{1}{4!}\,\frac{h_1^2}{a_2} x_1 + C_1. \tag{2.20}$$

Now, equation (2.19) and (2.20) yields,

$$\frac{G_2}{a_2} - \frac{R_1}{d_1\,h_1} = \frac{1}{4!}\,\frac{h_2^2}{a_2} z_3 + [\frac{b_2}{a_2} - \frac{1}{d_1}\,\frac{h_1^2}{4!}]\,z_2 + [\frac{3}{d_1} - \frac{1}{a_2}]\,\frac{h_1^2}{4!} z_1, \tag{2.21}$$

with $d_1 = \frac{u_1}{h_1} - z_1$.

Thus, the overall system is given by,

$$\left(\begin{array}{ccccccc}
A_1 & A_2 & \frac{1}{4!}\frac{h_2^2}{a_2} & 0 & \dots & & 0 \\
\frac{1}{4!}\frac{h_1^2}{a_2} & m_3 & n_3 & \frac{1}{4!}\frac{h_3^2}{a_3} & 0 & \dots & \\
0 & \frac{1}{4!}\frac{h_2^2}{a_3} & m_4 & n_4 & \frac{1}{4!}\frac{h_4^2}{a_4} & 0 & \\
\vdots & \ddots & \ddots & \ddots & \ddots & \ddots & \vdots \\
. & . & . & . & . & . & 0 \\
 & \dots & 0 & \frac{1}{4!}\frac{h_{p-3}^2}{a_{p-2}} & m_{p-1} & n_{p-1} & \frac{1}{4!}\frac{h_{p-1}^2}{a_{p-1}} \\
. & & \dots & 0 & \frac{1}{4!}h_{p-2}^2 & A_3 & A_4 \\
0 & & . & \dots & 0 & \frac{1}{4!}h_{p-1}^2 & A_5
\end{array}\right)
\left(\begin{array}{c} x_1 \\ x_2 \\ \vdots \\ x_{p-1} \\ x_p \end{array}\right)
= \left(\begin{array}{c} B_1 \\ O_3 \\ O_4 \\ \vdots \\ O_{p-1} \\ B_2 \\ B_3 \end{array}\right) \tag{2.22}$$

where

$$
\begin{aligned}
A_1 &= [\frac{3}{d_1} - \frac{1}{a_2}]\,\frac{h_1^2}{4!},\\
A_2 &= \frac{b_2}{a_2} - \frac{1}{d_1}\,\frac{h_1^2}{4!},\\
A_3 &= \frac{1}{4!}\,h_{p-2}^2 - \frac{1}{2}\,(h_{p-1}+h_{p-2})\,\frac{u_{p-2}}{h_{p-2}} - \frac{1}{3!}\,(h_{p-2}^2 - h_{p-1}^2) + \frac{1}{2}\,z_{p-1}\,(h_{p-1}+h_{p-2}),\\
A_4 &= -\frac{1}{2}\,h_{p-1}\frac{u_{p-2}}{h_{p-2}} - \frac{1}{3!}\,h_{p-1}^2 + \frac{1}{2}\,h_{p-1}\,z_p,\\
A_5 &= -\frac{3}{4!}\,h_{p-1}^2 - \frac{u_{p-1}}{2} + \frac{1}{2}\,h_{p-1}\,z_p \quad \text{and}\\
B_1 &= \frac{G_2}{a_2} - \frac{R_1}{d_1\,h_1},\\
B_2 &= \frac{R_{p-2}}{h_{p-2}} - f_p + f_p'\,(z_p - \frac{u_{p-2}}{h_{p-2}}),\\
B_3 &= \frac{R_{p-1}}{h_{p-1}} - f_p + f_p'\,(z_p - \frac{u_{p-1}}{h_{p-1}}).
\end{aligned}
$$

This system of linear equations yields the curvatures $x_k$. The coefficient sequences $C_k$ and $D_k$ are computed iteratively backward from (2.8) and (2.10) respectively, with initial values given by (2.14). As a result, the distribution density function is completely defined by equation (2.4). Evidently, the spline is defined on condition that the Pareto distribution on the tail is specified. This will be explained in the next section.

## 2.4 Pareto tail

Ideally, one would want to use the extended cubic spline over all income sizes, to let the data show the distribution on its own. The main drawback of that method is that it cannot accommodate infinite income. It is quite unrealistic to assume infinite income but as mentioned earlier, Saez (2001) showed empirical evidence that the income distribution is indeed a Pareto distribution at the tail.

Rahman and Pearson (2003) have done an extensive study on statistical methods to estimate Pareto distributions. We present a slightly different derivation here. The structure of the data given in the form of histogram allows us to estimate the Pareto parameters of the distribution. The difference is that we have the constraint of revenue matching for incomes above $z_p$.

Let us assume that the tail of the distribution is Pareto with parameters $a$ and $\alpha$. To facilitate the estimation of those parameters, we choose the following form for the density function,

$$f(z) = \alpha \frac{1}{z^{a+1}} \quad \forall z \geq z_p. \tag{2.23}$$

The revenue constraint for incomes between $z_p$ and $\infty$ is given by,

$$\begin{aligned} R_p + \ldots + R_N &= \int_{z_p}^{\infty} f(z)\,dz = \int_{z_p}^{\infty} \alpha \frac{a}{z^{a+1}}\,dz \\ &= \frac{\alpha}{a} [-\frac{1}{z^a}]_{z=z_p}^{z=\infty} \\ &= \frac{\alpha}{a} [\frac{1}{z_p^a}]. \end{aligned} \tag{2.24}$$

The last equality is true only for $a > 0$. This is verified from the results we obtain after estimation. This equation implies that $\alpha / a = z_p^a \sum_{i=p}^{N} R_i$.

The overall distribution has a Pareto tail above certain income $z_p$. Saez (2001) showed that above the top bracket the US income distribution is surely a Pareto distribution. However, following the underlying philosophy of this thesis to let the data reveal the distribution, we decide to only fit a Pareto distribution for income level above the last income threshold $z_N$ given in the data. This is \$1 million from 1996 to 1999 and \$10 millions from 2000 to 2008, whereas the top bracket is less than \$360.000 in 2008. We have, $z_p = z_N$. Hence,

$$\frac{\alpha}{a} = z_N^a R_N. \tag{2.25}$$

Nevertheless, assuming the data should show a Pareto-like distribution above the top bracket threshold[1], the parameters $\alpha$ and $a$ satisfy,

$$\begin{aligned} R_k + \ldots + R_{N-1} &= \sum_{i=k}^{N-1} R_i = \int_{z_k}^{z_N} f(z)\,dz \\ &= \frac{\alpha}{a} [-\frac{1}{z^a}]_{z=z_p}^{z=z_N} = \frac{\alpha}{a} [\frac{1}{z_k^a} - \frac{1}{z_N^a}] \\ &= R_N [(\frac{z_N}{z_k})^a - 1]. \end{aligned} \tag{2.26}$$

The last equality is obtained by substituting (2.25). Here $k$ goes from $b_5$ to $N-1$, where $b_5$ is the corresponding index to the income threshold from the histogram bin that is above the top bracket threshold. Therefore (2.26) provides many estimates of the Pareto parameter. For each

[1] We use Saez (2001) results to derive the parameters but we still use the extended cubic spline for $z < z_N$.

k, we have,

$$\hat{a}_k = \frac{\log[1 + \frac{\sum_{i=k}^{N-1} R_i}{R_N}]}{\log(z_N) - \log(z_k)},$$
$$\hat{\alpha}_k = \hat{a}_k\, z_N^{\hat{a}_k}\, R_N. \tag{2.27}$$

The second equation is from (2.25).

Quandt (1966) presents a variety of methods to estimate parameters of the Pareto distribution. Castillo and Haldi (1995) describe methods to choose a single estimate given a number of parameters to choose from quantile estimates. We will choose the parameter that has the least error. That is the combination $(\alpha_k, a_k)$ that minimizes the sum of the square of discrepancies in the revenues from the resulting Pareto distribution and the data.

## 2.5 Low Income Density

The US data provided by the IRS statistics is based on a sample of the actual collected revenue data. The sample gives a good representation of the actual distribution. The only flaw is that it has undersampled low income taxpayers. Very few taxpayers fell on the lowest bracket from the sample. The IRS itself acknowledges the lack of sample for low income taxpayers. For example, in 1998, there is no taxpayer that has below \$10.000 income. For this reason, special care should be taken when deriving the revenue distribution in this region.

The first approach is to combine all the low income revenue up to an income level which is no longer undersampled. In subsequent chapters we will combine the data on the first income bracket (which is under a single marginal tax rate). The sample size is one of the reasons that motivates that.

Second, we seek a simple polynomial fit. Using trial and error, we heuristically tune the degree of polynomial to be fitted on the first income interval (where taxpayers are undersampled), while ensuring that the revenue is zero at no income as well as making sure that the revenue density distribution is non-negative. We found that a cubic polynomial never worked on the US data. As a result, low income revenue distribution cannot be fitted with a cubic spline. However, we still impose smoothness constraints as in the theory of cubic splines.

We assume[2], $f(z) = a_0 z^n + a_1 z^{n+1} + a_2 z^{n+2} + a_3 z^{n+3}$ for $z \in [z_0, z_2]$ where $z_2$ is the income level from the data where we combined all revenue (denoted $R_1$ as before) collected from taxpayers whose income falls below that level. The constraints are given by,

[2] Of course, this is not the only formulation but it is one that is easy for imposing smoothness constraints.

$$\begin{cases} R_1 = a_0 \dfrac{z_2^{n+1}}{n+1} + a_1 \dfrac{z_2^{n+2}}{n+2} + a_2 \dfrac{z_2^{n+3}}{n+3} + a_3 \dfrac{z_2^{n+3}}{n+3} \\ f(z_2) = a_0\, z_2^{n} + a_1\, z_2^{n+1} + a_2\, z_2^{n+2} + a_3\, z_2^{n+3} \\ f'(z_2) = n\ a_0\, z_2^{n-1} + (n+1)\ a_1\, z_2^{n} + (n+2)\ a_2\, z_2^{n+1} + (n+3) a_3 z_2^{n+2} \\ f''(z_2) = n(n-1) a_0 z_2^{n-2} + (n+1) n a_1 z_2^{n-1} + (n+2)(n+1) a_2 z_2^{n} + (n+3)(n+2) a_3 z_2^{n+1}. \end{cases} \tag{2.28}$$

The first equation is the revenue constraint. The area under the curve $f$ must match the corresponding revenue from the data. The remaining equations are the smoothness constraints. The density function $f$ has to be continuous, differentiable and second differentiable at $z_2$ where the low income curve intersects the extended cubic spline curve. Solving equation (2.28) for the coefficients $a_0, \ldots, a_3$ completely determines the density function $f$ on $[z_0, z_2]$.

The main idea here is to provide a smooth curve for an irregular data set while avoiding negative density as well as multiple oscillations of the curve. At first, this seems to be a drawback of the use of the extended cubic spline. It cannot accommodate undersampling in the data. Statistical methods would not have to undergo special treatments even when the data presents some irregularities. However, undersampling might substantially bias the distribution's parameter estimates from their true values. In other words, the statistical estimation procedure works for any type of data but the resulting distribution can be seriously misleading. On the other hand, the extended cubic spline is independent of any misrepresentations that happen anywhere else on the income axis. Undersampling on some parts of the domain changes the shape of the distribution but does not affect the remaining parts. Finally, we note here that this low-income region is a small part of total income and will not play a significant role in the scope of our analysis in this thesis.

## 2.6 Links to Clamped Cubic Splines

A clamped cubic spline is a simple cubic spline, that fits a series of cubic polynomials given a set of data points, such that its derivative at the initial and final points are *"clamped"* to specific values as opposed to constraints on the curvature imposed at those points for the natural cubic spline, Plato (2003); Wang (2010). Since the tax revenue per income interval is given by the data, one could alternatively fit a clamped cubic spline to the cumulative distribution function of the revenue. The density follows by taking the first derivative of the resulting fitted-function. We show here that this does not always work well in practice and a higher degree clamped spline is required. The degree four clamped spline corresponds to the extended cubic spline derived in this chapter.

Let $F$ denote the cumulative PIT revenue distribution function. We have, $F(z_k) = \sum_{i=1}^{k-1} R_i$ for $k = 2, \ldots, N$. Naturally, $F(z_1) = F(0) = 0$. The Pareto distribution at the tail provides additional constraints on $f$ as well as $f'$ (and higher derivatives if desired). The Pareto parameters of the distribution are estimated as explained in the previous section. We have, for $z \geq z_N$,

$$\begin{aligned} F(z) &= \sum_{i=1}^{N-1} R_i + \int_{z_N}^{z} \frac{\hat{\alpha}}{t^{\hat{a}+1}} \, dt \\ &= \frac{\hat{\alpha}}{\hat{a}} \left[\frac{1}{z_N^{\hat{a}}} - \frac{1}{z^{\hat{a}}}\right] + \sum_{i=1}^{N-1} R_i. \end{aligned} \tag{2.29}$$

By construction $F$ is continuous. Continuity of $f$ and $f'$ are given by the following constraints:

$$\begin{cases} F'(z_N) = \dfrac{\hat{\alpha}}{z_N^{\hat{a}+1}}, \\ F''(z_N) = -\dfrac{\hat{\alpha}}{\hat{a}} \dfrac{1}{z_N^{\hat{a}}}. \end{cases} \tag{2.30}$$

A clamped cubic spline on $F$ is fitted using constraints given by equation (2.30) as well as $F(0) = 0$. As in the extended cubic spline, undersampled low income tax revenue has to be combined into a single revenue.

The theory of cubic spline is a well developed theory and there are readily available numerical procedures for implementation. However, it is not recommended to fit the spline directly on the cumulative distribution function as this has to be monotonically increasing from 0 to 1. When income thresholds step sizes are small (very likely after nondimensionalisation), the spline tends to oscillate to enforce continuity and smoothness over a short range interval. This in turn results in the density function being negative on the decreasing part of the oscillations. Measurement and sampling errors could also results in the same effect. This is the reason we advocate fitting the spline on the density function itself. Oscillations may arise but the corresponding cumulative distribution function will still be increasing from 0 to 1. One could in theory cumulate more income intervals to mitigate possible oscillations but this has the adverse effect of underestimating the true curve.

Furthermore, enforcing both constraints (2.30) leads to a numerically unstable linear system. In fact, linear iterative equations of the cubic spline only require two initial conditions. There are three possibilities but the first two lead to either a completely smooth function that does not start at zero that is not a cumulative distribution function, or a function that starts at zero but is discontinuous at $z_N$. The third choice impose $F(0) = 0$ and discards the second equation in (2.30). However, the second equation on curvature ensures smoothness of the corresponding density function. Increasing the degree of the spline to four increases the degrees of freedom by one and allows us to enforce the curvature constraint in (2.30). This is precisely the extended cubic spline we advocate as smoothness is one of the requirements we impose on the revenue and number of returns density curve in chapter 4.

## 2.7 Normalization and nondimensionalization

Normalization of a distribution requires that the area under the curve is set equal to one. This is done by dividing all entries in the revenue data for a specific year by the total revenue for that year. Furthermore, having the area equal to one for any distribution allows us to perform intertemporal distributional analysis. Indeed, when the distribution is normalized, any change from year to year at an income level is a change in the tax burden regardless of the total amount of revenue raised.

The extended cubic spline, in theory, does not require normalization. In practice, however, large numbers can render the process either numerically unstable or inaccurate. Typical revenues are in the range of billions of dollars. Normalizing the revenue is a necessary step for comparison. However, normalizing revenue alone will not be enough, we have to nondimensionalize all other quantities involved in the tax revenue formulation.

Nondimensionalization is a standard technique used in mathematical modeling. Howison (2005) describes it as the process whereby variables are scaled with typical values to transform mathematical models (algebraic nonlinear equations, ODEs, PDEs) into models with dimensionless quantities. By this procedure, variables with different orders of magnitude are reduced to the same scale, in theory. The number of parameters in the theory may also be reduced. Indeed, since the dimensions in a model are always balanced, the typical values used to scale the variables have to balance as well on the nondimensionalized model. The size of these nondimensionalized parameters shows the relative importance of various terms and variables in the model. Terms involved with large parameters drive the model while those with small parameter do not affect much the dynamics.

For the PIT tax revenue distribution, we have a simple relationship where the revenue collected at taxable income $z$ is the tax liability for people whose taxable income is $z$ multiplied by their numbers. Let $f, g$ denote the density of the tax revenue and number of returns distributions respectively. We have by definition, Creedy and Gemmell (2006),

$$f(z) = \bar{r}(z)\, z\, g(z) \tag{2.31}$$

where

$$\bar{r}(z) = \frac{T(z)}{z}$$

is the average tax rate for a taxpayer with taxable income $z$ and tax liability $T(z)$.

Howison (2005) suggests as a general rule that the obvious candidate be chosen, and changed only if it does not lead to the desired form. Let us perform a dimension analysis of our variables. We have, for the year $t$,

$$\begin{aligned} R_k^t &= \int_{z_k^t}^{z_{k+1}^t} f_t(z)\, dz, \\ [Mo] &= \sum [f_t]\,[Mo], \end{aligned} \tag{2.32}$$

where $[Mo]$ denotes the money dimension. It follows that $f_t$ is dimensionless,

$$[f_t] = 1. \tag{2.33}$$

Similarly,

$$\begin{aligned} N_k^t &= \int_{z_k^t}^{z_{k+1}^t} g_t(z)\, dz, \\ 1 &= \sum [g_t]\,[Mo]. \end{aligned} \tag{2.34}$$

Hence,

$$[g_t] = \frac{1}{[Mo]}. \tag{2.35}$$

Naturally,

$$[z] = [Mo] \text{ and } [\bar{r}] = [\frac{T(z)}{z}] = \frac{[Mo]}{[Mo]} = 1, \tag{2.36}$$

where $T(z)$ is the tax liability of a taxpayer with taxable income $z$, that is $[T(z)] = [Mo]$ .

Let

$$f_t = f_0\, f_t^*,\ g_t = g_0\, g_t^*,\ z = z_0\, z^* \text{ and } \bar{r} = r_0\, \bar{r}^*$$

be the factorization of $f, g, z$ and $\bar{r}$ at time $t$ respectively. For typical dimensions/scales $f_0, g_0, z_0, r_0$ of our problem and nondimensionalized variables $f^*, g^*, z^*, \bar{r}^*$. The scale factors $f_0$ have the dimensions

$$[f_0] = [f_t] = 1,\ [g_0] = [g_t] = \frac{1}{[Mo]},\ [z_0] = [z] = [Mo] \text{ and } [r_0] = [\bar{r}] = 1.$$

The total revenue and number of returns in year t satisfy,

$$R_0^t = \int_0^\infty f_t(z)\,dz \quad \text{and} \quad N_0^t = \int_0^\infty g_t(z)\,dz. \tag{2.37}$$

Normalizing,

$$1 = \int_0^\infty \frac{f_t(z)}{R_0^t}\,dz \quad \text{and} \quad 1 = \int_0^\infty \frac{g_t(z)}{N_0^t}\,dz, \tag{2.38}$$

where $[R_0^t] = [Mo]$ and $[N_0^t] = 1$, in a given year t.

Changing the integration variable $z$ to the dimensionless variable $z^* = z/z_0$ on both integrals implies,

$$1 = \int_0^\infty z_0\,\frac{f_t(z^*)}{R_0^t}\,dz^* \quad \text{and} \quad 1 = \int_0^\infty z_0\,\frac{g_t(z^*)}{N_0^t}\,dz^*. \tag{2.39}$$

As $dz^*$ is dimensionless, the dimensionless revenue and number of returns densities are respectively,

$$f_t^*(z^*) = z_0\,\frac{f_t(z^*)}{R_0^t} \quad \text{and} \quad g_t^*(z^*) = z_0\,\frac{g_t(z^*)}{N_0^t}. \tag{2.40}$$

Equivalently,

$$f_0 = \frac{R_0^t}{z_0} \text{ and } g_0 = \frac{N_0^t}{z_0}.$$

Clearly,

$$[f_0] = \frac{[R_0^t]}{[z_0]} = \frac{[Mo]}{[Mo]} = 1 \quad \text{and} \quad [g_0] = \frac{[N_0^t]}{[z_0]} = \frac{[1]}{[Mo]} \tag{2.41}$$

as desired.

The tax revenue density formula given by equation (2.31) implies,

$$\begin{aligned} f_t(z) &= \frac{R_0^t}{z_0}\,f_t^*(z^*) \\ &= \bar{r}(z)\,z\,g(z) \\ &= r_0\,\bar{r}^*(z^*)\,z_0\,z^*\,\frac{N_0^t}{z_0}\,g_t^*(z^*), \end{aligned} \tag{2.42}$$

where $\bar{r}^*(z^*) = r(z_0 z^*)/r_0$. Equation (2.42), in turn, implies,

$$f_t^*(z^*) = r_0\,z_0\,\frac{N_0^t}{R_0^t}\,\bar{r}^*(z^*)\,z^*\,g_t^*(z^*). \tag{2.43}$$

Since (2.31) has no parameter, we have to choose either $r_0$ or $z_0$ to reduce the parameter in (2.43) to one. The obvious (or 'typical') choice for $r_0$ would be $r_0 = 1$, 100% is the maximum tax rate that is physically applicable but economically meaningless. In this case the nondimensionalized averaged tax rate is the same as the average tax rate with no dimension.

Other typical values for $r_0$ are the marginal rates per income bracket on a specific year or the arithmetic mean of the marginal rates.

When $r_0$ is fixed, we have ,

$$z_0 = \frac{R_0^t}{r_0 \, N_0^t}. \tag{2.44}$$

In this case $z_0$ is the amount of income per taxpayer that would be required if everyone is taxed at fixed rate $r_0$, to obtain the total revenue $R_0^t$. Naturally,

$$[z_0] = \frac{[R_0^t]}{[r_0]\,[N_0^t]} = \frac{[\mathrm{Mo}]}{[1]\,[1]} = [\mathrm{Mo}],$$

as expected.

For the case $z_0$ is fixed, typical values would be the average income or the different income brackets thresholds. In the next chapter, we conveniently use the smallest income threshold $z_k$ that is greater than the top bracket. We have,

$$r_0 = \frac{R_0^t}{z_0 \, N_0^t}. \tag{2.45}$$

In this case, $r_0$ is the tax rate that would be required to obtain the total revenue $R_0^t$ if everyone has the same income $z_0$. Again, $[r_0] = \frac{[R_0^t]}{[z_0]\,[N_0^t]} = \frac{[\mathrm{Mo}]}{[\mathrm{Mo}]\,[1]} = [1]$.

Equations (2.33) and (2.36) imply that inflation adjustment surprisingly has no effect on the revenue density curve and the average tax rate. However, it does play a role for the number of returns curve as evidenced by (2.35). This is to be expected as most taxpayers automatically move to higher income thresholds and/or brackets without inflation indexation. Furthemore, the nondimensionalizing factor of $g$, given in (2.39) by

$$g_0^t = \frac{N_0^t}{z_0}$$

implies that the effects of inflation indexation on $g$ goes like $1/z_0$. That is, it is inversely proportional to income. In other words, inflation indexation decreases the number of returns at $z$ inversely proportionally to the deflator index used.

While the main objective of nondimensionalization as explained above is to scale large numbers down to small ones. The $z_0$ factor in equation (2.40) seems to counter the effect of the normalization by $R_0^t$ ( or $N_0^t$ for $g_t^*$). However, it is a necessary term as the result of the nondimensionalization of the income $z$. Not only incomes are as large as the tax revenues themselves (as a percentage of the taxable income) but scaling only selected variables does not help improve numerical accuracy and stability. We will show now that the $z_0$ factor can actually be obtained by scaling the income thresholds on the data by $z_0$.

Let $z_1,\ldots,z_N$ be the income thresholds given by the data (first column of Table 2.1) and let $z_1^* = z_1/z_0,\ldots,z_N^*/z_0$. Let $f_Z^t(z)$ be the revenue density function derived from the data using the $z_1,\ldots,z_N$ income thresholds and let $f_{Z^*}^t(z)$ be the revenue density function derived from the data using the $z_1^*,\ldots,z_N^*$ income thresholds. We aim to prove that $f_{Z^*}^t(z) = z_0 f_Z^t(z)$.

First, let us note every variable from Table 2.4 corresponding to $z_k^*, k = 1,\ldots,N$ with a $^*$. For example, $h_k^* = z_{k+1}^* - z_k^*$. Let us find the relationship between the curvatures $x_k, k = 1,\ldots,N$ of the spline using the $z_1,\ldots,z_N$ thresholds and the curvatures $x_k^*, k = 1,\ldots,N$ corresponding to $z_1^*,\ldots,z_N^*$. We have,

$$\begin{cases} h_k = z_0\, h_k^* & b_k = z_0^2\, b_k^* \\ u_k = z_0^2\, u_k^* & m_k = z_0\, m_k^* \\ a_k = z_0\, a_k^* & n_k = z_0\, n_k^* \\ t_k = z_0^2\, t_k^* & \end{cases} \tag{2.46}$$

On the other hand, if in addition to factorizing $z = z_0 z^*$, we also factorize $R_i = R_0\, R_i^*$, that is we divide all revenue entries on year[3] t from the data, we have $G_k^* = R_0\, G_k/z_0$ and $O_k^* = R_0\, O_k/z_0^2$. It follows that in (2.22), $Ax = b$ is equivalent to $z_0\, A^* x = R_0\, b^*/z_0^2$, where $A$ is the matrix in (2.22), $x$ the vector of unknown curvatures and $b$ the right hand side vector. Hence, $x^* = \frac{z_0^3}{R_0}\, x$.

Assuming, $C_p^* = \frac{z_0^2}{R_0} C_p$ and $D_p^* = \frac{z_0}{R_0} D_p$ which we will prove below, we have $C_k^* = \frac{z_0^2}{R_0} C_k$ and $D_k^* = \frac{z_0}{R_0} D_k$ for $k = 1,\ldots,p-1$. The extended cubic spline function from (2.4) satisfies,

$$\begin{aligned} S_k(z) &= \frac{1}{3!}\,\frac{x_{k+1}}{h_k}\,(z - z_k)^3 + \frac{1}{3!}\,\frac{x_k}{h_k}\,(z_{k+1} - z)^3 + C_k\, z + D_k, \\ &= \frac{1}{3!}\,\frac{R_0\, x_{k+1}^*/z_0^3}{z_0\, h_k^*}\,[z_0(z^* - z_k^*)]^3 + \frac{1}{3!}\,\frac{R_0\, x_k^*/z_0^3}{z_0\, h_k^*}\,[z_0\,(z_{k+1}^* - z^*)]^3 + \\ &\quad + \frac{R_0}{z_0^2}\, C_k^*\,(z_0\, z^*) + \frac{R_0}{z_0}\, D_k^*, \\ &= \frac{R_0}{z_0}\left\{\frac{1}{3!}\,\frac{x_{k+1}^*}{h_k^*}\,(z^* - z_k^*)^3 + \frac{1}{3!}\,\frac{x_k^*}{h_k^*}\,(z_{k+1}^* - z^*)^3 + C_k^*\, z^* + D_k^*\right\}, \\ &= \frac{R_0}{z_0} S_k|_{R^*,Z^*}(z^*), \end{aligned} \tag{2.47}$$

[3] We drop all superscript t for now, to avoid heavy notations.

where $S_k|_{R^*,Z^*}(z^*)$ denotes the corresponding spline using the $z_1^*,\ldots,z_N^*$ income threshold and the revenue entries $R_1^*,\ldots,R_N^*$ as the observed data.

On the Pareto tail of the distribution, first we note from equation (2.27) that the Pareto parameter $\hat{a}_k$ is a dimensionless parameter which does not change when scaling the data, neither $z_k$ to $z_k^*$ nor $R_k$ to $R_k^*$. That is,

$$\hat{a}_k = \hat{a}_k^*.$$

Next, we have $\hat{\alpha}_k = \hat{a}_k\, z_N^{\hat{a}_k}\, R_N = \hat{a}_k^*\, [z_0\, z_N^*]^{\hat{a}_k^*}\, R_0\, R_N^* = R_0\, z_0^{\hat{a}_k^*}\, \hat{\alpha}_k^*$. The Pareto tail density function from equation (2.23), for $z \geq z_p$, satisfies[4]

$$\begin{aligned}
f(z) &= \alpha \frac{1}{z^{a+1}}, \\
&= R_0\, z_0^a\, \alpha^* \frac{1}{[z_0\, z^*]^{a+1}} = \frac{R_0}{z_0}\, \alpha^* \frac{1}{[z^*]^{a+1}}, \\
&= \frac{R_0}{z_0}\, f|_{R^*,Z^*}(z^*),
\end{aligned} \tag{2.48}$$

where $f|_{R^*,Z^*}(z^*)$ denotes the Pareto density derived using the $z_1^*,\ldots,z_N^*$ income threshold and the revenue entries $R_1^*,\ldots,R_N^*$ as the observed data.

Equation (2.49) implies that $f_p, f'_p$ in (2.13) satisfy,

$$\begin{aligned}
f_p &= f(z_p) = \frac{R_0}{z_0} f_p^*, \\
f'_p &= f(z_p) = -(a+1)\,\alpha \frac{1}{z_p^{a+2}} = -(a+1)\, R_0\, z_0^a\, \alpha^* \frac{1}{[z_0\, z_p^*]^{a+2}} \\
&= -(a+1)\, \frac{R_0}{z_0^2}\, \alpha^* \frac{1}{[z_p^*]^{a+2}}, = \frac{R_0}{z_0^2} f_p'^*.
\end{aligned} \tag{2.49}$$

This proves that $C_p^* = \frac{z_0^2}{R_0} C_p$ and $D_p^* = \frac{z_0}{R_0} D_p$, which proves (2.47). Equations (2.47) and (2.49) in turn prove that $f(z) = \frac{R_0}{z_0} f|_{R^*,Z^*}(z^*)$ for all $z$ on the domain.

We have showed that to compute the nondimensionalized and normalized density function $f^*(z^*) = \frac{z_0}{R_0} f(z)$ we have to compute $f|_{R^*,Z^*}(z^*)$. That is, we compute $f$ with the same algorithm but using the scaled quantities $R^*$ and $Z^*$ as new data on the scaled income $z^*$. Thus, we managed to nondimensionalize the density functions by just scaling the data appropriately. In other words, we are able to use small numbers, and thus improve numerical accuracy and stability, and at the same time we are able to go back and forth between nondimensionalized and real scale density functions. Of course, year-to-year comparison of $f(z), g(z), \bar{r}(z)$ is now sensible.

[4] We drop the subscript $k$ above as we assume we have chosen one of the $a_k$ (see earlier section) and since it is dimensionless, any choice will not change the previous equality, allowing us to drop the subscript $k$ on $\alpha$ too.

## 2.8 Testing the Extended Cubic Splines

The revenue requirement, $R_k = \int_{z_k}^{z_{k+1}} f_k(z)\,dz$, intuitively suggests that as income thresholds stepsize $z_{k+1} - z_k$ is decreased, the approximation to $f$ converges to the true curve. Indeed, $f_k(z) \sim \frac{R_k}{\Delta z_k}$ as $\Delta z_k = (z_{k+1} - z_k) \to 0$. However, revenue and number of returns data are given through a specific set of income thresholds. As a result, convergence is not guaranteed. Furthermore, income thresholds stepsizes in the data increase as income increases, suggesting that approximation could get worse as we move along the income axis. We investigate in this section the convergence performance of the extended cubic splines on the lognormal density function. Our goal is to investigate the precision to which the extended cubic splines approximate curves with similar properties to the actual data. We use functions with peaks, modes and skewness in the range of actual data as general test functions. We note that these test functions do not have to be probability density distributions but any function with the same properties.

Figure 2.2 shows histograms of the revenue and number of returns from 1996 to 2008. The revenue histograms are plotted on the left. Interestingly, there are only two modal classes. The first one, plotted in black, is \$50,000-75,000 between 1996 -1999. The second modal class, plotted in red, is \$75,000-100,000 from 2000-2008. The peaks range from $3 \times 10^{-6}$ to $6.5 \times 10^{-6}$ when normalized. Plotted on the right, the number of returns shows three lower modal classes. The first modal class, plotted in black is \$30,000-40,000 in 1996 and 1997. The second, plotted in red, is \$40,000-50,000 from 1998 to 2002 and the third, plotted in blue, is \$50,000-75,000 from 2003 to 2008. The peaks are much higher from $1.17 \times 10^{-5}$ to $1.46 \times 10^{-5}$ when normalized. Both the revenue and the number of returns are left-skewed.

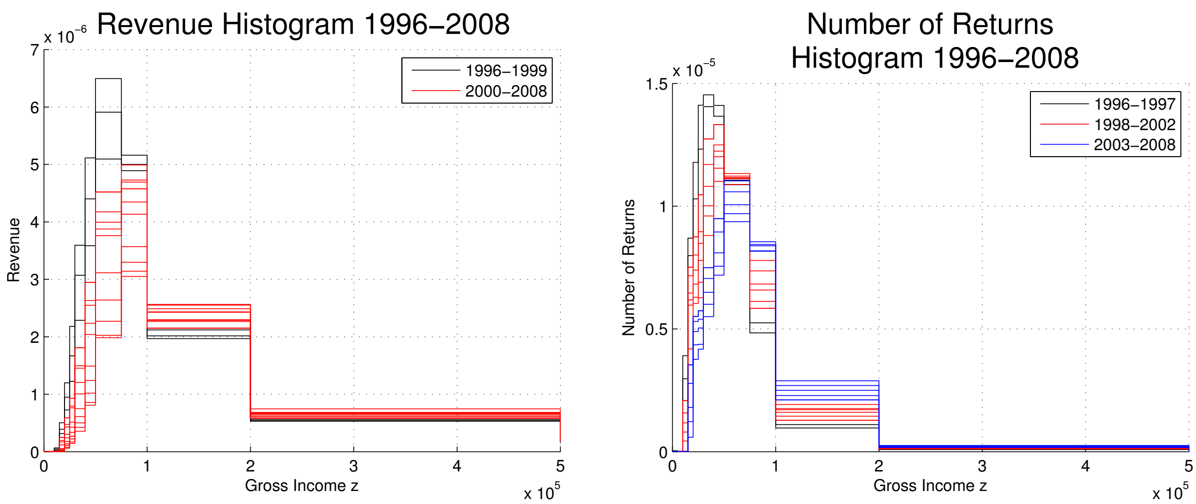



Figure 2.2: *Revenue and Number of Returns histograms from 1996 to 2008.*

We use the lognormal density curve as a general test function. It is left-skewed and is used in the literature to model income distribution. As a start, one could fit an *"eyeball"* lognormal distribution to the data. That is a lognormal distribution with the same peak and mode as the data given in Table 2.3 each year from 1996 to 2008. Figure 2.3 displays the four extreme cases for the eyeball revenue and number of returns. That is the lognormal distribution with lower and upper bound of the mode and peak different combinations. For example, the upper left distribution curve is the reconstruction of the lognormal with the lower bound of the mode and upper bound of the peak from data. The original lognormal curve is plotted in black whereas the extended cubic spline reconstruction is in red. We plot an additional histogram in blue to show the location of the bins. The height of the histogram is not relevant here, it is chosen to keep the figure neat and visible.

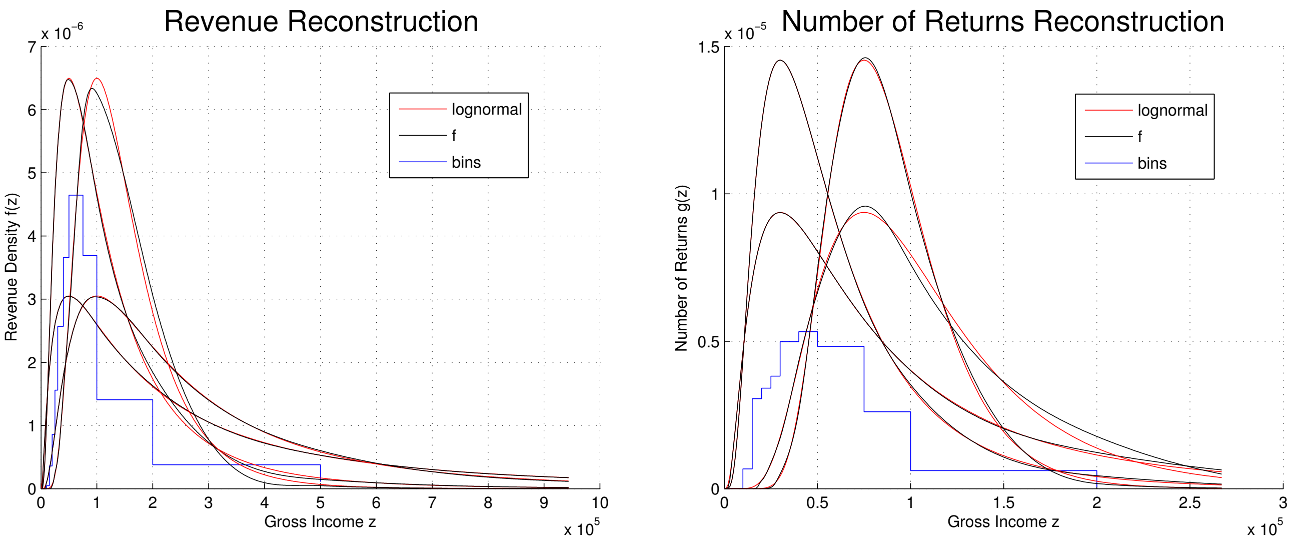



Figure 2.3: *Revenue and Number of Returns with various combination of the lower and upper bound of the mode and peak, reconstructed.*

Clearly, the larger the width of the distribution, that is the lower its peak, the more accurate is the extended cubic splines while using the data income thresholds. The peak and the mode are identified extremely precisely by the extended cubic spline reconstruction. Table 2.5 and 2.6 below shows in row 2 and 3 the relative error of the peak and the mode for respectively the revenue and number of returns combination of lower and upper bounds of the peak and mode from the data. We provide in the first row the root mean square (RMS) of the relative errors between the original lognormal density curve and the extended cubic spline approximation to give an overall estimate of its performance. We use the relative error as a measure of the distance between the two curves to account for the extremely low magnitude of the lognormal distribution. With income above \$10 millions to have non zero density, at the peak the lognormal distribution is roughly around $10^-6$. We use the following formula for the RMS:

$$\text{RMS} = \sqrt{\frac{1}{(z_N - z_1)} \int_{z_1}^{z_N} (\frac{f_k(z) - \hat{f}_k(z)}{f_k(z)})^2 \, dz},$$

where $f_k$ and $\hat{f}_k$ denote the actual lognormal and its extended cubic spline approximation respectively. $z_1$ and $z_N$ denote income thresholds within which the overall performance of the extended cubic spline is investigated. Similarly, the relative error of the mode is $\frac{|m-\hat{m}|}{m}$ where $m = \log(\mu - \sigma^2)$, the mode of the lognormal distribution with parameter $\mu$ and $\sigma$ and $\hat{m}$ the mode of the extended cubic spline. At the peak, the relative error is $\frac{|f(m)-\hat{f}(\hat{m})|}{f(m)}$.

Table 2.5 and 2.6 confirm that as the peak and/or the mode increases the extended cubic spline approximation of the peak and the mode is getting less precise. An upper bound of 8% and 0.8% relative error of the mode is given for the revenue and number of returns respectively. For the peak relative error, the upper bound are 2.46% and 2.3% for the revenue and number of returns respectively. With extremely small relative errors, one could conclude that the extended cubic spline is performing well given the income bins from the data used. However, two important remarks must be given. First, the peak relative error does not grow linearly as the mode relative error for the number of returns. The mode relative error is higher for the combination maximum mode and minimum peak (lower right curve in the right hand plot in Figure 2.3) than the combination maximum mode, maximum peak. Second, the RMS relative error reveals that the extended cubic splines, though doing extremely well on low income range, are not approximating the upper tail equally well. This is to be expected as the number of income bins is highly spaced in that range. The overall performance of the extended cubic spline as shown in Table 2.5 and 2.6 do not follow a specific trend. Hence, one cannot draw a strong conclusion on the overall performance and mode reconstruction trend.

| | **( min(m), min(p))** | **( min(m), max(p))** | **( max(m), min(p))** | **( max(m), max(p))** |
|---|---|---|---|---|
| **RMS** | 0.1696 | 0.0739 | 0.6305 | 0.2864 |
| **RPE** | 0.0015 | 0.0033 | 0.0045 | 0.0246 |
| **RME** | 0.0219 | 0.0246 | 0.0281 | 0.0869 |

Table 2.5: *Relative RMS error, peak and mode error from the combination of mode and peak at the upper and lower bound from the revenue data.*

Figure 2.4 shows respectively the relative peak error, relative mode error and relative RMS trend in the revenue and number of returns mode and peak overlapping range. All three plots show bumps and oscillations over the mode and peak range but the overall trend is that the extended

| | ( min(m), min(p)) | ( min(m), max(p)) | ( max(m), min(p)) | ( max(m), max(p)) |
|---|---|---|---|---|
| **RMS** | 0.1540 | 0.6299 | 1.9200 | 0.1138 |
| **RPE** | 0.0007 | 0.0001 | 0.0228 | 0.0056 |
| **RME** | 0.0042 | 0.0004 | 0.0074 | 0.0086 |

Table 2.6: *Relative RMS error, peak and mode error from the combination of mode and peak at the upper and lower bound from the number of returns data.*

cubic spline is more precise as the mode and/or peak decrease. Furthermore, the relative peak error upper bound is 4% and the relative mode upper bound is 8%. One can conclude that within the mode and peak of the data range, the extended cubic spline reconstructs the peak and the mode successfully. We note that these results apply for other left-skewed test functions such as $f(x,\alpha,\beta) = x^{\alpha}e^{-\beta x}$ but is not reported here for space convenience.

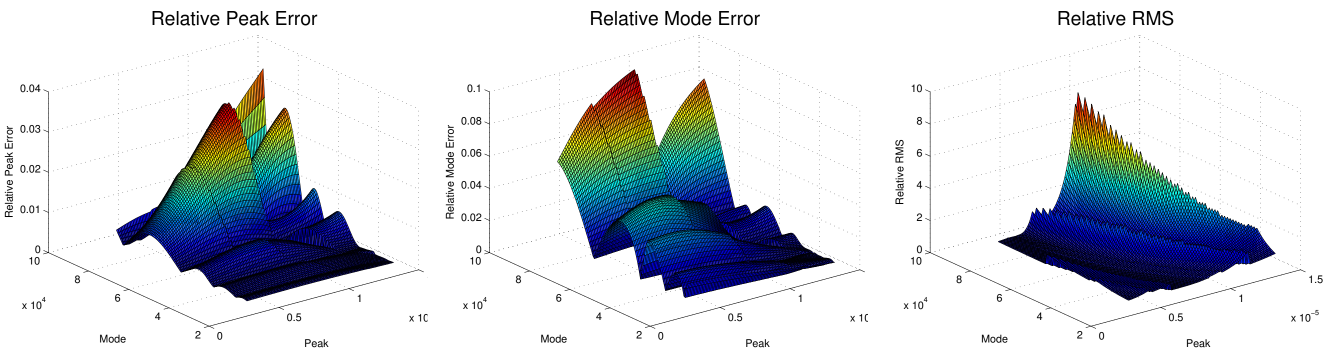



Figure 2.4: *Trend in the relative peak error, relative mode error and relative RMS in the number of returns and revenue combined range.*

## 2.9 Testing the Pareto tail

To test the dependency and sensitivity of the extended cubic splines with respect to its Pareto tail, we remove the tail of the lognormal fitted to the data, keeping the same peak and mode as the original data in Table 2.3, and we fit instead a Pareto distribution. Table **??** depicts the RMS relative error, relative peak error and the relative mode error from 1996 to 2008. The first three columns correspond to the case where we generate the Pareto tail using the area above the last income thresholds and its subsequent area that is between the second last and the last income thresholds. In other words, we fit the Pareto tail precisely according to the way they are estimated in our extended cubic spline algorithms. Surprisingly the relative peak and mode errors

| Years | Area 1 and Area 2 | | | Area 1 and Continuity | | | Continuity and Smoothness | | |
|---|---|---|---|---|---|---|---|---|---|
| | RMS | p | m | RMS | p | m | RMS | p | m |
| **1996** | 0.1470 | 0.0079 | 0.0509 | 0.1470 | 0.0079 | 0.0509 | 0.1373 | 0.0079 | 0.0509 |
| **1997** | 0.1324 | 0.0068 | 0.0494 | 0.1324 | 0.0068 | 0.0494 | 0.1167 | 0.0068 | 0.0495 |
| **1998** | 0.0649 | 0.0053 | 0.0465 | 0.0649 | 0.0053 | 0.0465 | 0.0485 | 0.0054 | 0.0466 |
| **1999** | 0.0887 | 0.0015 | 0.0021 | 0.0887 | 0.0015 | 0.0021 | 0.0731 | 0.0015 | 0.0020 |
| **2000** | 0.0827 | 0.0015 | 0.0039 | 0.0827 | 0.0015 | 0.0039 | 0.0658 | 0.0016 | 0.0038 |
| **2001** | 0.0928 | 0.0015 | 0.0008 | 0.0928 | 0.0015 | 0.0008 | 0.0781 | 0.0015 | 0.0007 |
| **2002** | 0.0993 | 0.0014 | 0.0013 | 0.0993 | 0.0014 | 0.0013 | 0.0859 | 0.0014 | 0.0014 |
| **2003** | 0.0919 | 0.0015 | 0.0011 | 0.0919 | 0.0015 | 0.0011 | 0.0770 | 0.0015 | 0.0010 |
| **2004** | 0.0777 | 0.0015 | 0.0057 | 0.0777 | 0.0015 | 0.0057 | 0.0594 | 0.0016 | 0.0055 |
| **2005** | 0.0685 | 0.0013 | 0.0102 | 0.0685 | 0.0013 | 0.0102 | 0.0458 | 0.0014 | 0.0096 |
| **2006** | 0.0661 | 0.0012 | 0.0123 | 0.0661 | 0.0012 | 0.0123 | 0.0410 | 0.0013 | 0.0115 |
| **2007** | 0.0651 | 0.0010 | 0.0144 | 0.0651 | 0.0010 | 0.0144 | 0.0374 | 0.0012 | 0.0131 |
| **2008** | 0.0654 | 0.0010 | 0.0136 | 0.0654 | 0.0010 | 0.0136 | 0.0386 | 0.0012 | 0.0125 |

Table 2.7: *Relative RMS, peak and mode error for different Pareto tails of the lognormal distribution with peak and mode from the original data, from 1996 to 2008.*

have tremendously decreased to be as low as half their original value. However, the extended cubic splines did not recover the Pareto tail for year 1996, 1997 very badly and 2001, 2002, 2003 to a lesser extent. As a result, the overall RMS error has increased for those years. For the remaining years, it has successfully recovered the Pareto tail with decreased RMS. However, one deep flaw of this experiment is that by requiring the Pareto tail to have the same area in between three income thresholds, one has to forgo the continuity/differentiability of the overall original distribution. As those are key components of the extended cubic splines, one has to investigate if they actually worsen the approximation.

Columns 4,5,6 of Table 2.9 show the RMS relative error, relative peak error and the relative mode error where we fit on the last income threshold, the Pareto distribution that preserves the area above the last income threshold, continuous with its lognormal piece. That is, it has the same function value as the lognormal piece. There is absolutely no change in the results. This is to be expected as the extended cubic spline algorithm only uses each area per income thresholds and does not depend on the continuity, smoothness or any other property of the original curve.

The last three columns of Table 2.9 give the RMS relative error, relative peak error and the relative mode error where we fit on the last income threshold, the Pareto distribution that has the same function value and derivative as its lognormal upper piece. In other words, we impose continuity and smoothness at the income threshold where the lognormal ends and the Pareto distribution begins. The overall RMS improves but the peak and mode error increase. This is due to the fact that by requiring continuity and smoothness of the original distribution, the area above the last income threshold is no longer the same as previously given by the lognormal distribution. Therefore, the extended cubic spline is changed but it does not improve the precision at the peak.

### 2.9.1 Sensitivity and error propagation from the Pareto tail

The Pareto tail of the extended cubic spline curve is completely determined by its parameters $\alpha$ and $a$. Those are dependent on the area above the last income thresholds and its subsequent areas, as estimated by (2.27). The series of cubic polynomials on the subsequent income thresholds (counting backwards) are estimated based on the function value and derivative at the last income thresholds $z_N$. Let us now determine the effect of small variations in the function value $f_N = f(z_N)$ and derivative $f'_N = f'(z_N)$.

Let, $f_N = f_N + \Delta f_N$ and $f'_N = f'_N + \Delta f'_N$. First, the matrix in (2.22) does not depend on $f_N$ or $f'_N$. The last two values of the right hand side depend linearly on $f_N$ and $f'_N$. As a result, the curvature will remain the same except at the very last two income thresholds. To see this, we solve the linear system of the form $Ax = b + \Delta b$. It follows that, $x = x + A^{-1}\Delta b = x + \Delta x$. In our case, $x_N = x_N + \Delta x_N$, $x_{N-1} = x_{N-1} + \Delta x_{N-1}$ and $x_k = x_k$ for $k = 1, \ldots, N-2$. As $B2, B3$ in (2.22) depend linearly on $f_N$ and $f'_N$, we have,

$$x_N = x_N + \alpha_N \Delta f_N + \beta_N \Delta f'_N \quad \text{and} \quad x_{N-1} = x_{N-1} + \alpha_{N-1}\Delta f_N + \beta_{N-1}\Delta f'_N.$$

Therefore,

$$\begin{cases} C_N = f'_N - \frac{1}{2} x_N h_{N-1} = C_N + \Delta f'_N - \frac{1}{2} h_{N-1} (\alpha_N \Delta f_N + \beta_N \Delta f'_N) = C_N + C^1_N \Delta f_N + C^2_N \Delta f'_N \\ D_N = f_N - z_N f'_N + (-\frac{1}{3!} h^2_{N-1} + \frac{1}{2} h_{N-1} z_N) x_N = D_N + \Delta f_N - z_N \Delta f' N + \\ \quad (-\frac{1}{3!} h^2_{N-1} + \frac{1}{2} h_{N-1} z_N)(\alpha_N \Delta f_N + \beta_N \Delta f'_N) \\ \quad = D_N + D^1_N \Delta f_N + D^2_N \Delta f'_N \end{cases}$$

Second, from the recursion in the coefficient $C_k$ given by (2.8) and recursion in the coefficient $D_k$ given by (2.10), we have,

$$\begin{cases} C_{N-1} = C_N - \frac{1}{2}(h_N + h_{N-1})x_N = C_{N-1} + C_N^1 \Delta f_N + C_N^2 \Delta f'_N - \frac{1}{2}(h_N + h_{N-1})(\alpha_N \Delta f_N + \beta_N \Delta f'_N) \\ \qquad = C_{N-1} + C_{N-1}^1 \Delta f_N + C_{N-1}^2 \Delta f'_N \\ D_{N-1} = D_N - (-\frac{1}{3!}(h_{N-1}^2 - h_N^2) - \frac{1}{2}(h_{N-1} + h_N)z_N)\, x_N = D_{N-1} + D_N^1 \Delta f_N + D_N^2 \Delta f'_N + \\ \qquad (-\frac{1}{3!}h_{N-1}^2 + \frac{1}{2}h_{N-1}z_N)(\alpha_N \Delta f_N + \beta_N \Delta f'_N) \\ \qquad = D_{N-1} + D_{N-1}^1 \Delta f_N + D_{N-1}^2 \Delta f'_N. \end{cases}$$

Similarly,

$$\begin{cases} C_{N-2} = C_{N-2} + C_{N-2}^1 \Delta f_N + C_{N-2}^2 \Delta f'_N \\ D_{N-2} = D_{N-2} + D_{N-2}^1 \Delta f_N + D_{N-2}^2 \Delta f'_N \end{cases}$$

and for $k = N-3, \ldots, 1$,

$$\begin{cases} C_k = C_k + C_{N-2}^1 \Delta f_N + C_{N-2}^2 \Delta f'_N \\ D_k = D_k + D_{N-2}^1 \Delta f_N + D_{N-2}^2 \Delta f'_N. \end{cases}$$

Thus, we have proven that the curvatures $x_k$, the coefficients $C_k, D_k$ for all $k$, depend linearly on the variations $\Delta f_N$ and $\Delta f'_N$. Now, for $z \in [z_k, z_{k+1}]$,

$$f(z) = S_k(z) = \frac{1}{3!}\,\frac{x_{k+1}}{h_k}\,(z - z_k)^3 + \frac{1}{3!}\,\frac{x_k}{h_k}\,(z_{k+1} - z)^3 + C_k z + D_k.$$

Hence, $f(z) = f(z) + \alpha \Delta f_N + \beta \Delta f'_N$ for some finite coefficients $\alpha$ and $\beta$. For $\Delta f_N, \Delta f'_N \longrightarrow 0$, $f \longrightarrow f$. In other words, for small variations in the function value and first derivative at the last income threshold $z_N$, either as a result of a Pareto distribution or any other distributions, the resulting extended cubic spline function is not significantly altered.

The same analysis can be carried if the estimation was done where the start of the Pareto tail is changed. Two curves with different start of the Pareto tail at $z_{N1}$ and $z_{N2}$, with $z_{N1} \le z_{N2}$, will have an error of the form $f_{N1} = f_{N2} + \Delta f$ and $f'_{N1} = f'_{N2} + \Delta f'$. The rest of the proof is the same. Hence, the error propagation with respect to changes in the start of the Pareto tail is also linear. Note however that this is not valid in between the two different start of the Pareto tail, $z_{N1} \le z \le z_{N2}$. One has to be mindful of the implication of the wide bins at high income range.

## 2.10 Conclusion

The extended cubic spline described in this chapter approximates distributions given data constraints. This method is particularly designed to estimate the revenue and number of returns distributions where data is given per income bin threshold. By fitting a smooth curve and matching each revenue/number of returns per income bin, the procedure results in a simple system of linear equations. Unlike conventional statistical distribution fitting, we do not assume the data stem from a known distribution with unknown parameters for the middle and high income range. Instead, we use the powerful numerical technique of cubic splines and let the data reveal the true shape of the distribution itself. Careful nondimensionlization and normalization simplify the formulation and show that only the distribution of number of returns requires inflation indexation for year to year comparisons. A Pareto distribution is fitted at the upper tail on the very high income range, based on previous results in the literature.

Simulations on lognormal curves show that within the peak and mode from the revenue and number of return data range, the extended cubic spline successfully recovers the original curve. Its performance depends on the number of income bins and its frequency over the income axis. Using income bins of the data, the extended cubic spline performs extremely well on the low income range. The highest relative error in the mode and peak data range is in the order of 8% and 4% respectively. With income bins from data, as the mode and/or the peak increase, that is they are located further away from the numerous bins of the low income range, the procedure decreases in performance. Good approximation on low income range can be offset by bad performance at the upper tail. This is not a linear decrease and local behavior can show adverse trends but in general higher peak and mode lead to worse overall approximations. The Pareto fit at the tail of the distribution does not influence the overall shape of the cubic splines on the middle and lower income ranges. Changes to the tail result in limited changes to the slope and function value of the splines.

# Chapter 3

# Comparative Analysis of Intertemporal Changes to the US Joint Returns Income Tax Revenue Distribution from 1996 to 2008

## 3.1 Introduction

The individual income tax revenue distribution shows the true distribution of the income tax burden. Surprisingly, the income tax literature has not paid much attention to its shape or how it changes over the years. To the best of our knowledge, none has so far estimated and analyzed the overall distribution of income tax revenue, particularly the joint return income tax revenue distribution in the US. Government and congressional agencies such as the Congressional Budget Office (CBO), Joint committee on Taxation (JCT) and the Office of Tax Analysis (OTA) in the Department of Treasury do extensive distributional analyses on taxation, CBO (2007a); OTA (1999, 2008); Bradford (1995). However, they work on specific groupings, quantiles and percentiles of taxpayers population and have focused primarily on the distribution of effective tax rates rather than directly estimating and characterizing the overall tax revenue distribution itself.

Estimating the associated distribution of number of returns, the average tax rate and taxable income are also important. These are the main factors that explain changes in the distribution of revenue. Thus, various economic theories of taxpayers responses to changes in the tax parameters and how to estimate them, can be identified by individually estimating those factors. In particular, changes in the distribution of number of returns measures the true indirect response at each income to changes in the tax system and economic conditions overall. Administrative costs such as tax fraud and avoidance are automatically included as it is the observed distribution that is obtained.

Analysis by filing status is necessary as different filing status is associated with different tax preferences prior to the enactment of the Economic Growth and Tax Relief Reconciliation Act of 2001 (EGTRRA). For example, deductions and tax brackets amounts were not double that of the single taxpayers amounts prior to EGTRRA. Effects on taxpayers behavior as observed on the distribution of revenue could not be expected to be the same for single and joint taxpayers.

In this chapter, we compare the American individual income tax revenue distributions of married taxpayers who filed jointly from 1996 to 2008. We use the extended cubic spline method explained in the previous chapter on the IRS data. Estimating the overall distribution of revenue gives a clear picture of the dynamics of the tax burden and changes to taxpayers' behavior over the years. It not only shows features at specific points such as the widely used quantiles and percentiles but also reveals local behaviors surrounding them. By using a particular nondimensionalization factor, year to year comparison reveals a collapse in the middle income range of the distribution. Government intervention such as the gradual marginal tax rate reductions brought by EGTRRA restores temporarily this revenue trend. Normalization implies that the revenue follows the opposite trend on higher income ranges. That is, higher income ranges corresponding revenue increases as a counterpart to the systematic fast decrease on the middle income range. This is not a dramatic increase as it is on a much wider income range. In addition, there are specific clouds of points, group of taxpayers, that are invariant and do not move in the distribution over the years.

This chapter is organized as follows. Section 2 specifies the nondimensionalization and normalization factors to be used in order to observe year to year revenue properties changes. In section 3, we describe the estimation of the taxable income from the data, that we use to compute the average tax rate. In section 4, we show the year to year changes in the overall tax revenue distribution. Section 5 gives the interpretations. We give a detailed analysis per income range and compare different income range responses. We conclude in section 6.

## 3.2 Choosing the non-dimensionalization

We use the extended cubic spline detailed in the previous chapter to estimate the distribution of income tax revenue for married taxpayers filing jointly in the US, from 1996 to 2008. Revenue data for the corresponding years are given in Appendix 1. We only investigate regular income tax as the AMT only covers 3% of the taxpayers population and is subject to a different set of brackets and marginal rates.

Our primary objective is to observe the combined effects of EGTRRA and JGTRRA[1] on the distribution of revenue. We index for inflation to have the money dimension in constant dollar as explained in the previous chapter. We normalize the revenue data, so as to compare changes in the distribution of the burden relative to the distribution of taxpayers and/or distribution of AGI. We nondimensionalize to ease numerical computations as well as reveal and emphasize particular properties of the distribution that would have otherwise been unnoticed. The normalization and non-dimensionalization formulae in the previous chapter are general, we now select the nondimensionalizing factors $z_0, r_0, N_0$ and $R_0$. It is important that these factors may vary from year to year, but to simplify notation we emphasize the nondimensionalization property by the suffix zero.

* Define $z = z_0\, z^*$ where the nondimensionalization factor $z_O$ is the first income threshold from the data above the top marginal rate bracket threshold. That is the first income on the first column of Table A.10, above the last row of the corresponding year in the bracket Table A.9. A complete list of $z_0$ from 1996 to 2008 is given in Table A.12 in Appendix A.
* The revenue data $R|_Z$ with the income threshold $Z$ is transformed to $R|_Z = R_0\, R^*|_{Z/z_0}$, where $R_0 = \sum R|_Z$ is the total revenue and $R^*|_{Z/z_0}$ is the nondimensionalized revenue with its income thresholds $Z$ divided by $z_0$. In other words, Table A.10 is transformed every year in such a way that the income threshold column is divided by the corresponding $z_0$ at that year and the corresponding revenue column is divided by the corresponding $R_0$. The complete transformed set of income thresholds is given in Table A.13, and the corresponding revenue is given in Table A.14 in Appendix A.
* Similarly, the number of returns $N|_Z$ with the income threshold $Z$ is transformed to $N|_Z = N_0\, N^*|_{Z/z_0}$, where $N_0 = \sum N|_Z$ is the total revenue and $N^*|_{Z/z_0}$ is the nondimensionalized revenue with its income thresholds $Z$ divided by $z_0$. That is, every year the corresponding number of returns column in Table A.11 is divided by the corresponding total number of returns $N_0$. The complete set of transformed number of returns is given in Table A.15 in Appendix A.
* The Pareto tail of the distribution is started at $z_0$ for both the revenue and number of returns distributions. As shown in the previous chapter, having more income bins in the high income

[1] For the remainder of this chapter we will only use EGTRRA but it should be understood as the combination of EGTRRA and JGTRRA.

ranges does not help identify the distribution accurately as income bins in those ranges are wider and wider in range. Using the Pareto tail as early as at $z_0$ insures a better approximation for the overall distribution.

It follows from (2.15) in the previous chapter that the revenue density $f_t$ in year $t$ satisfies $f_t(z) = R_0^t f_t^*(z^*)/z_0$. Similarly the number of returns density $g_t$ satisfies $g_t(z) = N_0^t g_t^*(z^*)/z_0$ while the average tax rate $\bar{r}_t$ satisfies $\bar{r}_t(z) = z_0 N_0^t/R_0^t \bar{r}_t^*(z^*)$. Of course, $\int_0^\infty f_t^*(x)\,dx = \int_0^\infty g_t^*(x)\,dx = 1$. Using these transformations, we can easily go back and forth from nondimensionalized to nominal curves. We will show using this particular nondimensionalization, that the revenue distribution displays an instability and tax parameters changes brought by EGTRRA temporarily restored the original distribution.

## 3.3 Taxable Income and Average Tax Rate

At income $z$, changes in the revenue distribution $f_t(z)$ account for changes in the number of returns $g_t(z)$ and/or changes in the individual tax liabilities $T_t(z)$ as outlined by the formula

$$f_t(z) = T_t(z)\, g_t(z).$$

These two effects are not necessarily independent of each other but as the distribution of number of returns is directly observable from the data, we estimate the changes in the individual tax liabilities separately to explain changes in the revenue distribution. We use the average tax rate formulation. For ease of notation, we drop the index $t$ on all variables used in this section but it should be understood that the analysis is done for a particular year $t$.

By definition, Creedy and Gemmell (2006), the average tax rate is given by

$$\bar{r}(z) = \frac{T(I(z))}{I(z)},$$

where $I(z)$ denotes the individual taxable income for taxpayers with gross income $z$. It represents the total tax liability of taxpayers with gross income $z$ as a percentage of his taxable income. It follows that, $T(z) = T(I(z)) = I(z)\,\bar{r}(z)$.

Meanwhile, the tax liability of a taxpayer whose taxable income falls in the $i$-th tax bracket is given by

$$T(z) = r_i(I(z) - b_i) + \sum_{k=1}^{i-1} r_k(b_{k+1} - b_k),$$

with $I(z) \in [b_i, b_{i+1}]$ where $b_i$, $i = 1, \cdots, 5$ denote the tax brackets. Clearly, this simple formula works as long as the taxable income $I(z)$ is given. The IRS data however do not give explicitly the corresponding taxable income, given the gross income. Instead, the taxable income histogram corresponding to gross income threshold bins is given. Hence, we can estimate the distribution of taxable income using the extended cubic spline following the same procedure as with the distribution of revenue and number of returns. If $TI(z)$ denotes the density of taxable income distribution, the taxable income per individual distribution is $I(z) = TI(z)/g(z)$.

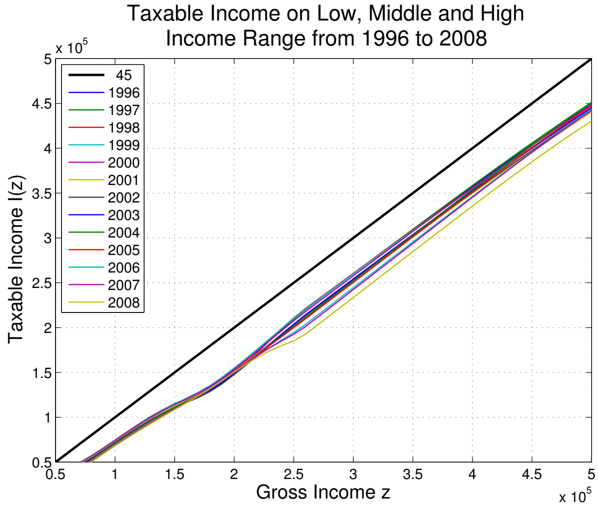
Alternatively, one can subtract the total deductions from the gross income to recover the individual taxable income. That is,

$$I(z) = z\,(1 - \frac{d(z)}{z\,g(z)})\,, \tag{3.1}$$

where, $d(z)$ denotes the distribution of total deductions at $z$. The ratio $d(z)/g(z)$ represents the average deduction per taxpayer with gross income $z$. On the last income bin of the data, this ratio is approximated by the total deduction divided by the total number of returns of the corresponding income bin. This formulation proves to be numerically more stable and preserves the increasing nature of taxable income as gross income increases. We shall use the (3.1) formulation of the taxable income in the remainder of this chapter.

Figure 3.1 displays the taxable income per return at the gross income $z$. The low, middle and high income range (all below \$500,000) is depicted on the left for the years 1996 to 2008. The taxable income is very regular and is constant over the years for gross income up to about \$250,000 at which point the spread of the yearly variation in the taxable income is widened. Taxable income in the most recent years, 2006, 2007, 2008 is lower than those of the previous years. This variation is in the order of 10%. This is explained by the fact that prior to 2006, the exemption amount could be zero whereas after 2006, the minimum exemption is \$1,100 per taxpayer. Moreover, the standard deduction for married taxpayers filing jointly was increased to double that of single taxpayers. More importantly, starting in 2006, the itemized deduction for taxpayers above threshold is now only 2/3 of the original reduction. This becomes 1/3 in 2008. In part, this explains the wider gap in taxable income in 2008. The financial crisis in 2008 perhaps explains the remaining part, whereby high income taxpayers had less income as a result of the crisis.

Similarly, for the very high income range (above \$500,000) corresponding taxable income is depicted on the right. We note that the taxable income in this range is given by the gross income subtracted by the ratio of the total deduction to the total number of returns in the highest income bin of the data. Here the year to year difference is not apparent. Total deduction is relatively small compared to gross income in this range. Hence, the taxable income is close to the 45 degree line. Furthermore, year to year variation in taxable income is much smaller than the taxable income itself and thus no apparent changes in taxable income are observed in the figure. This will be treated in more detail in section 5 where we specifically look at changes in taxable income per income ranges as part of the analysis.



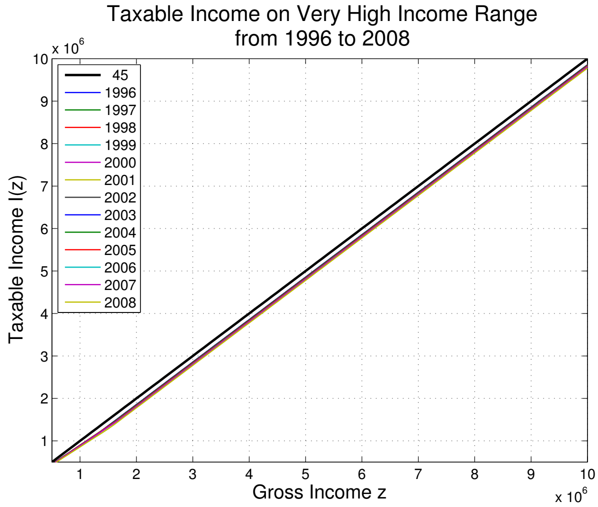




Figure 3.1: *Taxable income per return as a function of the gross income from 1996 to 2008. The thick black curve is the 45 degree line representing the identity function of the gross income.*

## 3.4 Year to Year Comparison of the Overall Joint Returns Income Tax Revenue Distribution

### 3.4.1 Pre-EGTRRA Era

Figure 3.2 displays in the left column normalized distributions and in the right column nominal distributions for the years before EGTRRA was enacted. The first row is the distribution of income tax revenue. We observe a consistent drop in the peak of the normalized distribution that cannot be explained by numerical errors. The peak of the nominal distribution shifts erratically every year and does not reveal the trend of normalized $f_t^*$. We do note an overall consistent weak shift to the right but redistribution of the tax burden cannot clearly be discussed from $f_t$. Recall that $f_t^*$ and $f_t$ are independent of inflation as shown in the previous chapter and these trends are owing purely to tax policy.

In the second row of Figure 3.2, we plot the distribution of number of returns. The normalized distribution on the left and nominal distribution on the right show no significant differences in pattern. There is no significant trend of the peak. Up to random small changes in $g_t^*$ or $g_t$, the distribution of taxpayers is virtually constant over the years. The peak of the number of returns distribution drops by about 7% in the pre-EGTRRA era compared to about 30% in the case of normalized tax revenues. It is located on the left of the peak of the revenue distribution. Nominal returns decrease by about 16% at the peak, but again, we cannot claim redistribution of returns across gross income.

In the third row, we plot the average tax rates for the same years. We can see that the nondimensionalized average rate $\bar{r}_t^*(z^*)$ on the left displays, at fixed taxable income, a consistent drop as well. In contrast, the nominal average rate $\bar{r}_t(z)$ displayed on the right, shows an insignificant shift to the right. Although during the pre-EGTRRA era, marginal tax rates did not change, brackets were inflation adjusted. This highlights the use of nondimensionalization since the nondimensionalized curve clearly reveals significant changes in average rates.

In the fourth row of Figure 3.2, we show in the left the nondimensionalized taxable income and in the right the nominal taxable income. Both nondimensionalized and nominal taxable income around the peak of the normalized revenue distribution did not change over the years prior to EGTRRA.

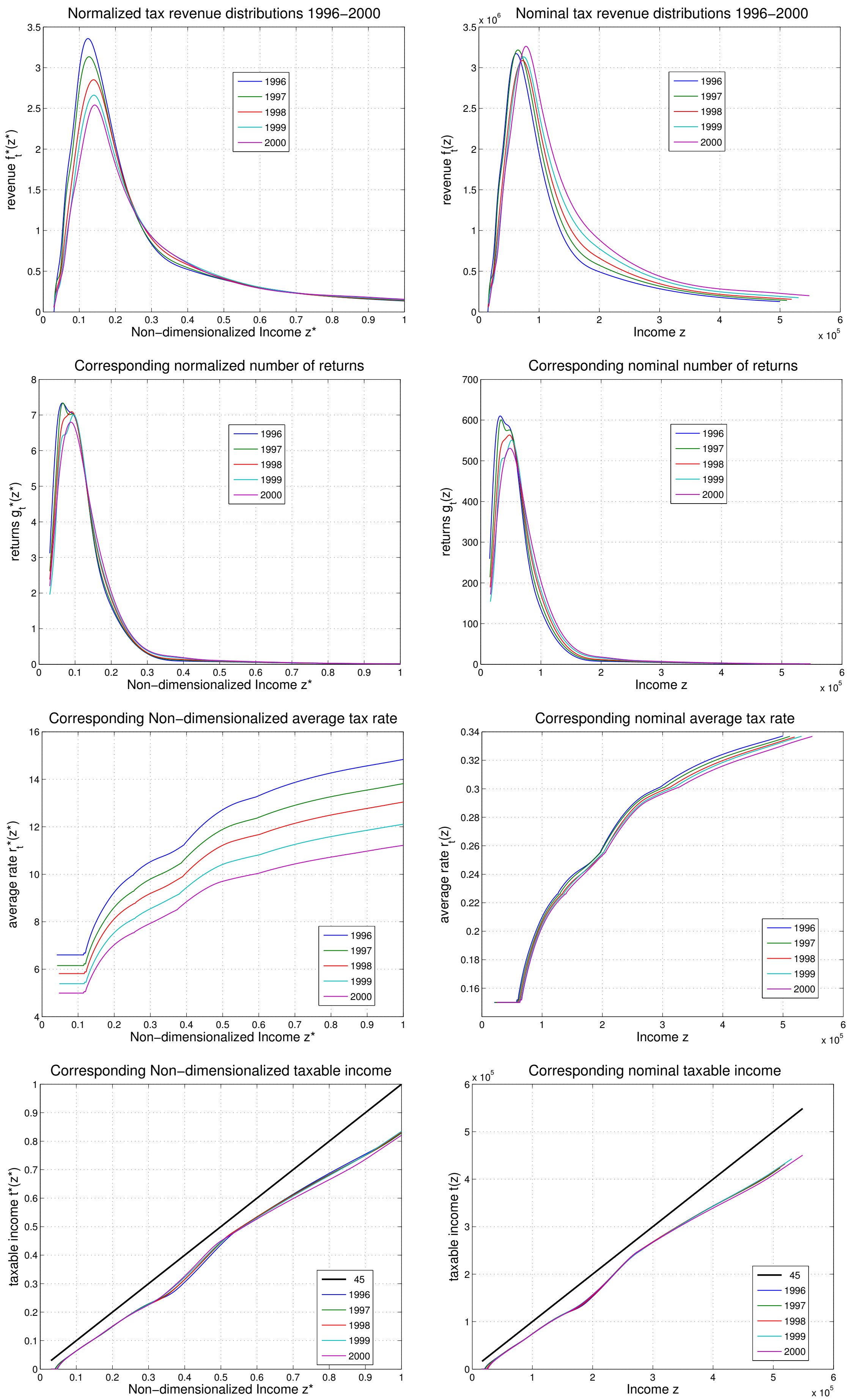



Figure 3.2: *Year to year changes prior to EGTRRA 2001.*

### 3.4.2 EGTRRA Phase In

We plot in Figure 3.3 the same distributions as in Figure 3.2 during the EGTRRA phase of marginal tax rate reductions. We also include year 2000 and 2004 to give a clear picture of the overall changes and a connection to the previous plots. As before, nondimensionalized distributions are displayed on the left and corresponding nominal distributions from the raw data on the right side. The first row is the distribution of income tax revenue. The normalized revenue distribution now follows the opposite trends of the pre-EGTRRA era. The peak of the distribution increases every year from 2000 to 2003. The marginal rate changes were fully implemented by 2003 and there were no rate changes in 2004 and the peak of the normalized distribution drops back close to the 2000 level. In contrast, the nominal distribution shows a consistent drop every year during this period.

In the second row, we plot the distribution of number of returns. The normalized distribution on the left shows no pattern around its peak and we see no change in the distribution of number of returns. The nominal distribution on the right is systematically decreasing so that the absolute number of returns in that range decreases. However, because $g_t^*$ is constant, we conclude that $g_t$ decreases everywhere in the same approximate proportion. Again, the peak of the number of returns distribution is located on the left of the peak of the revenue distribution.

The average tax rates for the same years are displayed on the third row of Figure 3.3. The nondimensionalized average rate on the left as well as the normalized revenue distribution show significant change. It increases from 2000 to 2002 and in 2003 it starts to decrease again. The nominal average rate, displayed on the right, decreases yearly as the reduction in the marginal tax rate brought by EGTRRA is reflected on the average tax rate. This explains the consistent drop in the nominal revenue distribution. However, this is the only visible part of the overall picture. As the total revenue increased every year (see Table in Appendix A), some parts of the income range must show an increase in the revenue distribution. Yet, the average rate decreased over the whole income range. Thus, a decrease in marginal rate does not automatically imply a decrease in revenue. This will be discussed in detail in a later section.

As before, the taxable income overall and in particular, at the peak of the revenue distribution, did not change during the EGTRRA phase-in period. Deductions and exemptions cannot be the cause of the reverse trend in the peak of the revenue. This implies that changes in marginal rates effectively disrupt and reverse the decreasing trend of the revenue.

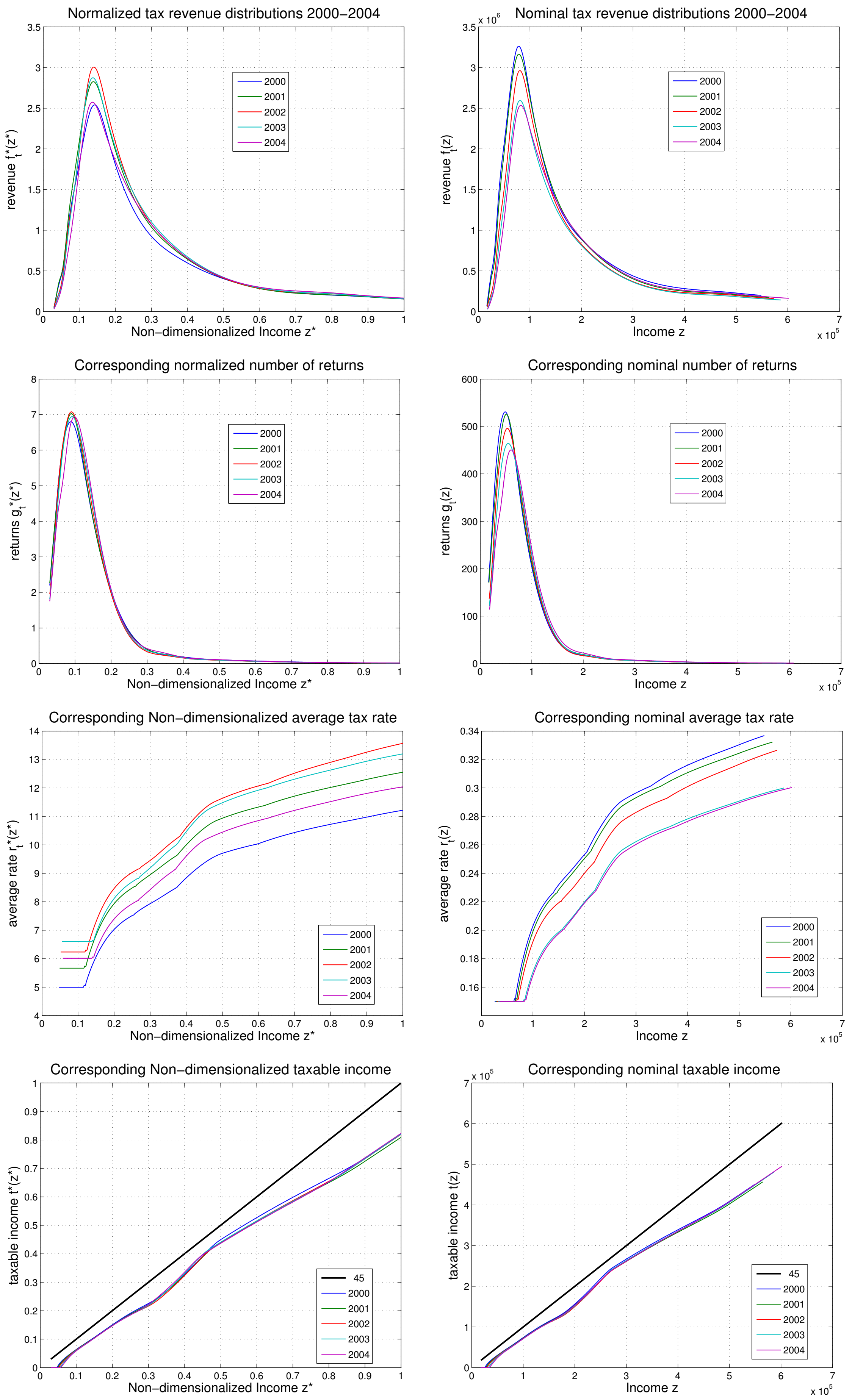



Figure 3.3: *Year to year changes during the EGTRRA 2001 phase in period.*

### 3.4.3 Post-EGTRRA Era

Figure 3.4 displays in the left side the nondimensionalized distributions and in the right side the nominal distributions from the raw data for the years after the EGTRRA marginal rate reductions are fully phased in. That is, the statutory marginal rates are the same as the last change in EGTRRA. As before, we include the year 2003 in the plot to provide a link to the picture in the previous section. The first row is the distribution of income tax revenue. We observe a similar consistent drop between 5 and 10% in the peak of the normalized distribution as in the pre-EGTRRA phase. In both eras, marginal rates were constant over the era. However in the year 2008, it jumps back while even tax rates were constant. The nominal distribution drops by about $10^{-4}$% and clearly shifts progressively to the right every year, as for the pre-EGTRRA era.

The distributions of number of returns are displayed in the second row. The peak of the normalized distribution on the left increases only slightly (2%) every year where the nominal distribution decreases by about 6%. Again, the normalization readjusts the curves to show that the comparable year to year changes do not significantly affect the distribution of taxpayers. The peak is located on the left of the peak of the revenue distribution. More detailed pictures for different income ranges are given and discussed in the next section. The same conclusion as in the pre-EGTRRA period can be made. When there is no change in the marginal tax rates, the revenue distribution $f_t^*$ and not $g_t^*$ consistently decreases year after year.

In the third row, we plot the average tax rates for the same years. As in the pre-EGTRRA period, the nondimensionalized average rate on the left displays a consistent drop. The nominal average rate on the right shows a slight shift to the right. As before, changes in the brackets did not follow the same inflation indexation, a slight yearly change is to be expected. Once more the nondimensionalization gives a clearer picture that links changes in average rates and changes in the distribution of revenue.

The taxable income plotted in the last row still shows virtually no change at the peak of the revenue distribution. There is however a gradual decrease at high income range as the changes in deductions and exemptions brought by EGTRRA is enacted. They were specifically set to apply after the changes in marginal rates were fully phased in. Hence, changes in nondimensionalized revenue at high income are the combination of the drop in average rate further emphasized by the drop in taxable income.

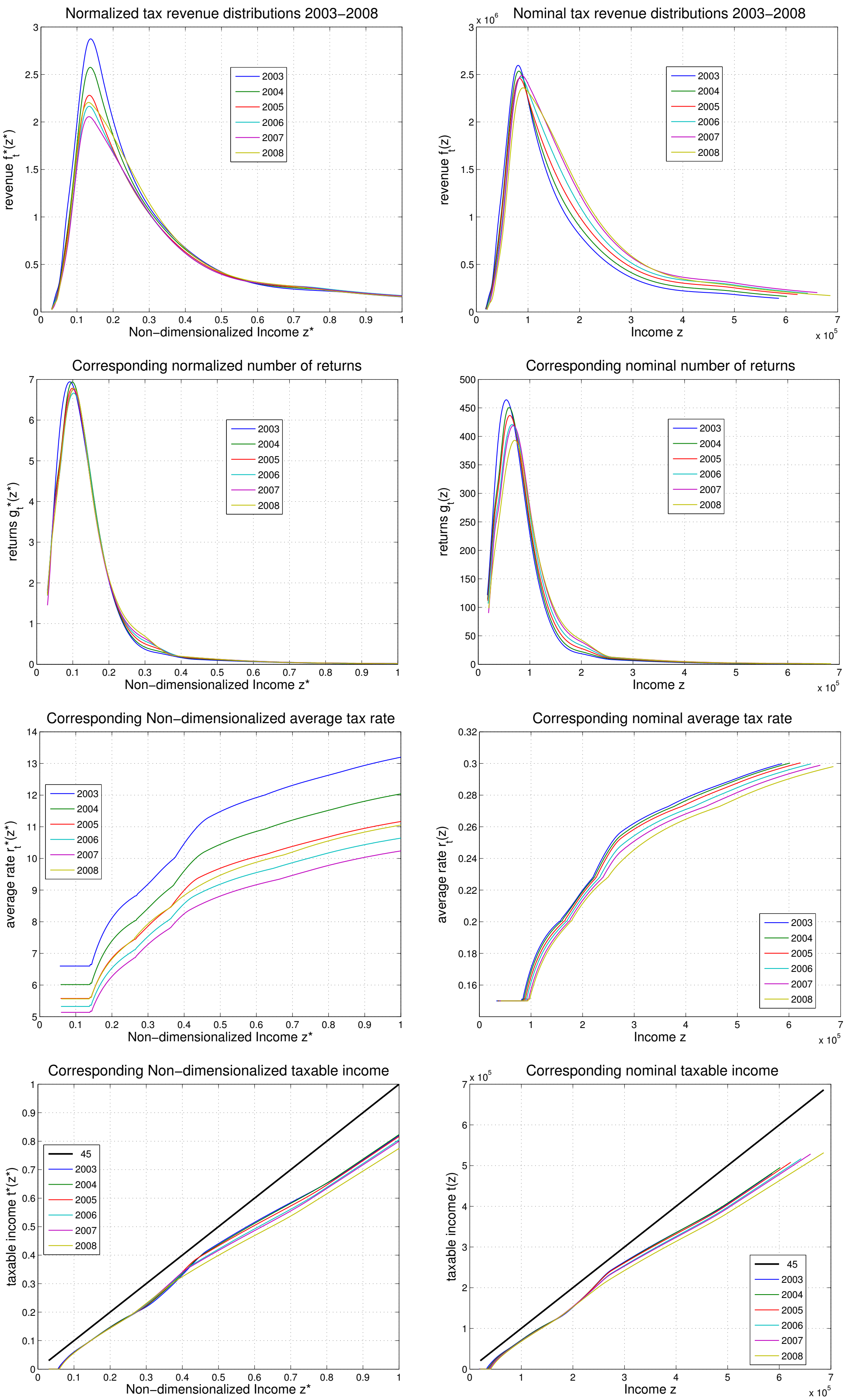



Figure 3.4: *Year to year changes after EGTRRA 2001 is fully phased in.*

## 3.5 Interpretations

The figures in the previous section give the overall picture of the revenue distribution for married taxpayers filing jointly over the years from 1996 to 2008. In this section, we will look closely and in more details in some income ranges and particular points of interest in the distribution by following their time series.

We divide the income axis into four income ranges. First, the low income families are taxpayers who fall in the first tax bracket. They are not discussed nor displayed in the figures ahead. Because of the low sampling in this range, we cannot make reliable conclusions. Second, there is the middle income range, those between the first bracket and the last bracket before the top bracket. We will plot the peak of the distribution but it is understood that the middle income range follows the same pattern. There will be some invariant points. Those are the points where the middle income range stop. This is why sometimes the upper bound of the middle income comes before the last bracket. Third, there is the high income range. This is from the invariant point to the last income threshold from the data. As we start the Pareto section of the curve at $z_N = z_0$, we will plot densities at $z_0$ or equivalently at the nondimensionalized income $1$, to represent the trend at high income range. Fourthly, there is the very high income range for taxpayers above the last income threshold $z_N$ from the data where we fit the Pareto distribution. We will plot exactly at $10\, z_0$. As the Pareto distribution is completely identified by its shape and height parameters, plotting the Pareto at a point is enough to observe the year to year trends in the very high income region. As before, plots on the left are for the nondimensionalized distributions while plots on the right are for the nominal distributions.

### 3.5.1 Middle Income Range

Figure 3.5 displays the time series of the revenue, number of returns, average rate and average taxable income at the peak of the revenue distribution from 1996 to 2008. At a glance, the revenue distribution and the corresponding average tax rate shows similar trends for both the normalized and nominal distribution. This confirms our observation in the previous section that the average tax rate is the main factor that drives the revenue. More precisely, changes in brackets alone, albeit small, involve persistent decrease in the revenue peak. Number of returns and taxable income are variable during this time.

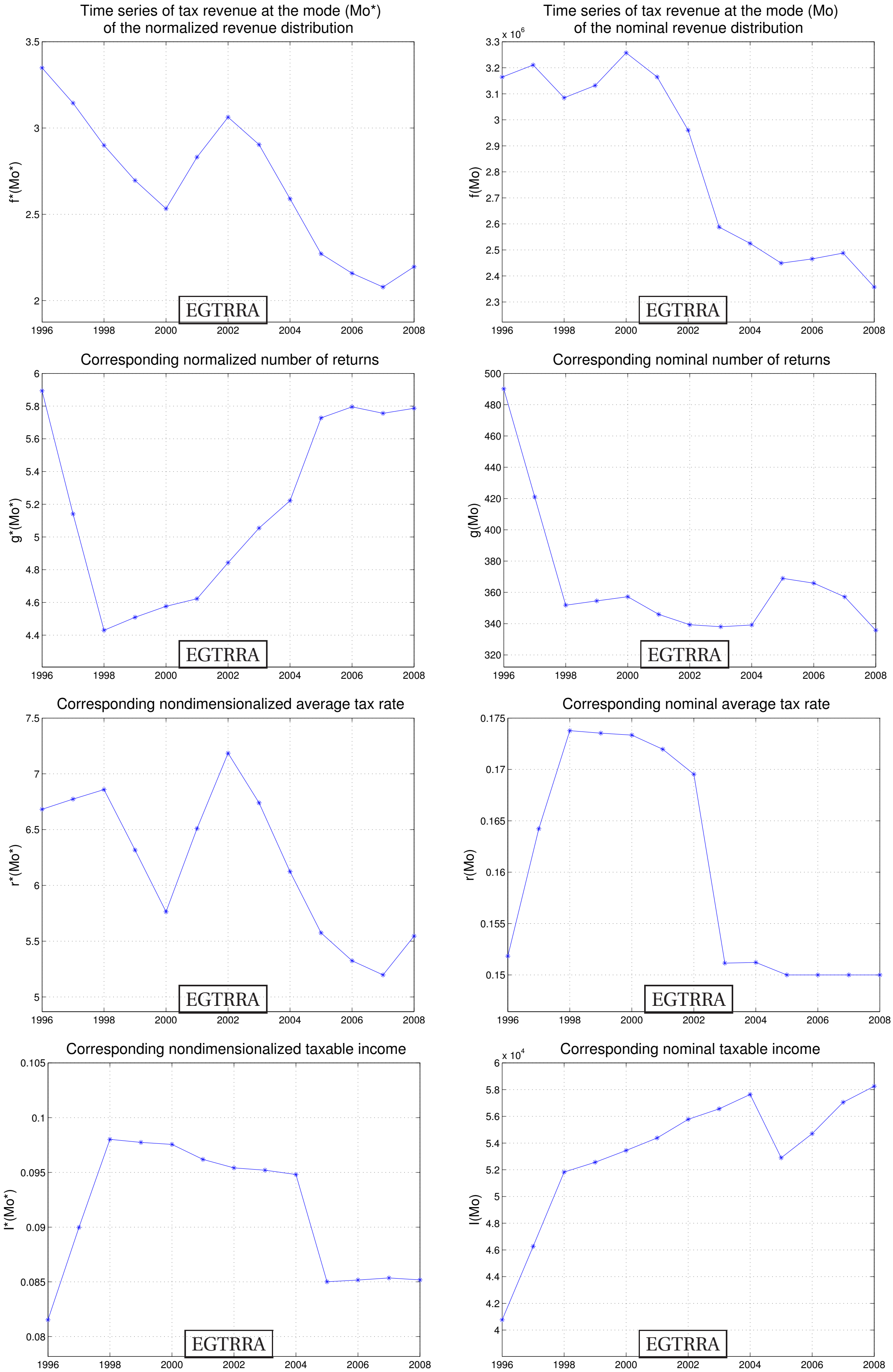



Figure 3.5: *Time series of the peak of the revenue distribution together with the corresponding number of returns, average rate and taxable income.*

On a closer look, one observes that from 1996 to 1998, the nondimensionalized average rate actually increased whereas the average taxable income increased at the peak of the revenue. It is the number of returns that decreased consistently over those three years that implies decreasing revenue. From 1998 to 2000, as the nondimensionalized number of returns as well as average nondimensionalized taxable income became constant, the average nondimensionalized rate decreased to cause the revenue to decrease as well. The link between peak revenue and nondimensionalized revenue is then emphasized.

During the EGTRRA phase-in period, while the nondimensionalized average taxable income slightly decreased, both the average rate and the number of returns increased and the revenue had to increase in response. The marginal tax rate decrease brought by EGTRRA is reflected in an increase in nondimensionalized average rate that induced an increase in revenue as a positive taxpayers' response. As the average taxable income decreased slightly, few taxpayers moved downward in the taxable income scale but the overall effect is an increase in the number of returns at the peak. This implies that more taxpayers were willing to pay taxes as a result of the marginal rate decrease than taxpayers who shifted on the income axis.

After the EGTRRA rate changes are fully phased in, the average rate again decreased and the number of returns stabilized. It follows that the revenue density also decreased on the middle income range. In 2008, because of the financial crisis, the US economy went into recession. This is reflected as an increase in the middle income tax revenue density. As the economy entered recession, taxpayers saw their taxable incomes decrease and thus the number of returns increases in the lower ranges of income. It will be shown in a later section that in 2008, the high and very high income range all drew fewer taxpayers.

Figure 3.6 shows the year to year difference in the normalized revenue and number of returns as well as the nondimensionalized average rate and taxable income at the mode of the revenue distribution. A decrease in year to year change corresponds to a negative value in the year to year difference plots. As mentioned earlier, the revenue at a particular income is the total tax payment multiplied by the number of returns at that income. It follows that, year to year changes in revenue results either from changes in tax payment or changes in number of returns. More precisely, from the identity $f_t^*(z) = \bar{r}_t^*(z)\, I_t^*(z)\, g_t^*(z)$, we have,

$$df_t^*(z) = \bar{r}_t^*(z)\, I_t^*(z)\, dg_t^*(z) + I_t^*(z)\, g_t^*(z)\, d\bar{r}_t^*(z) + \bar{r}_t^*(z)\, g_t^*(z)\, dI_t^*(z).$$

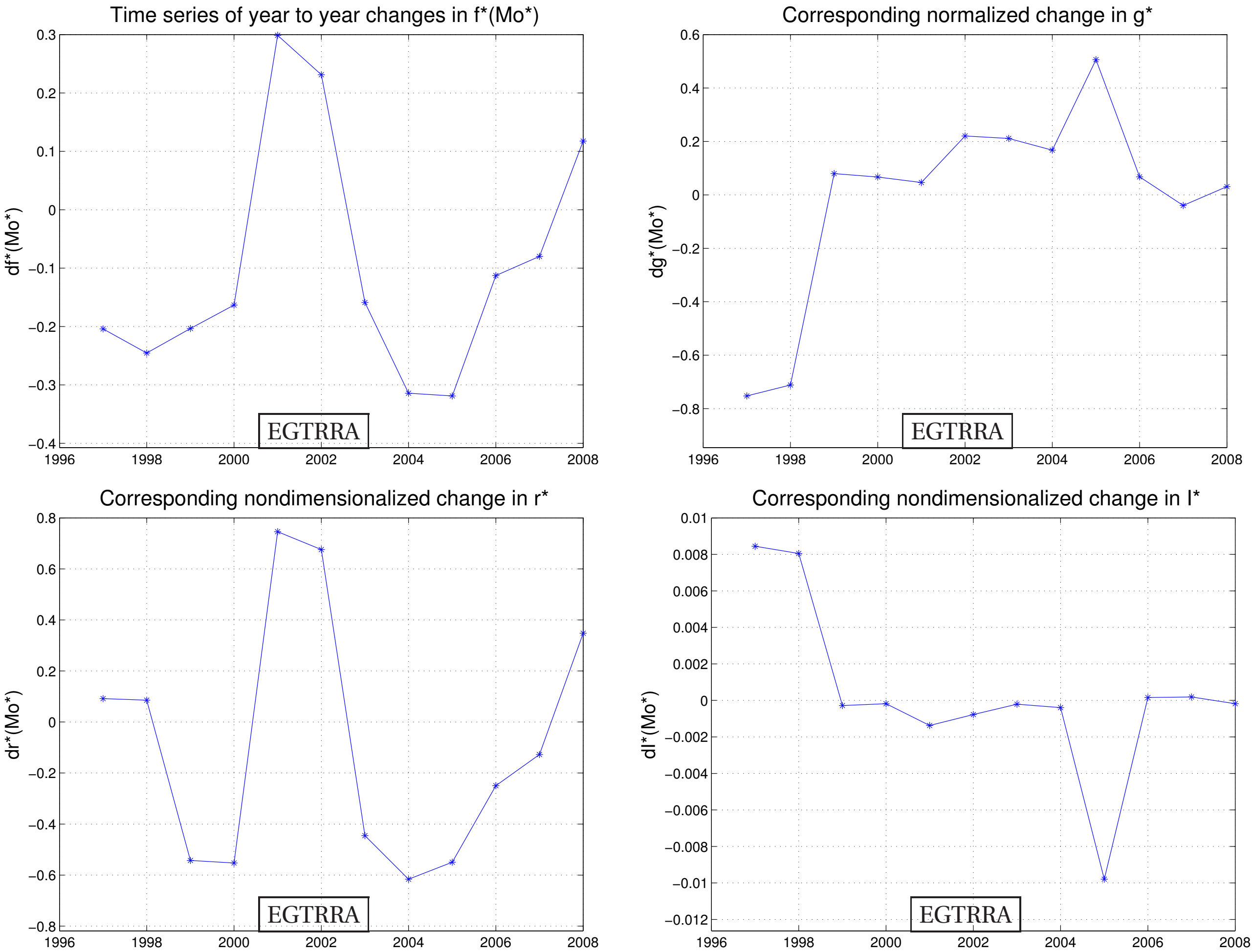



Figure 3.6: *Year to year changes in* $f^*(Mo^*)$, $g^*(Mo^*)$, $\bar{r}^*(Mo^*)$, $I^*(Mo^*)$ *at the peak* $Mo^*$.

The decreasing trend in revenue prior to EGTRRA is decomposed in two periods with different effects on revenue at the peak of the distribution. From 1996 to 1998, although both the average rate and taxable income increased, the decrease in the number of returns is large enough to imply a decrease in revenue. In other words, as the relative number of returns increases, that is more taxpayers move to higher income ranges, the relative revenue contribution decreases. From 1998 to 2000, we have the opposite effect. The number of returns increases to roughly 10% change per year but that is not large enough to counter the decrease in the average tax rate. There is virtually no change in the taxable income. Hence, the revenue decrease in the peak is now an average tax rate effect. Changes in the average rate is due to either changes in the marginal rates or tax brackets. Figure 3.7, shows that tax brackets have indeed been increased by inflation indexation only, every year from 1996 to 2008 and thus, the nondimensionalized tax brackets are virtually constant from year to year. Firstly, systematic change in the peak of $f_t^*$ is an unintended, hidden consequence of this simple tax policy. Secondly, the

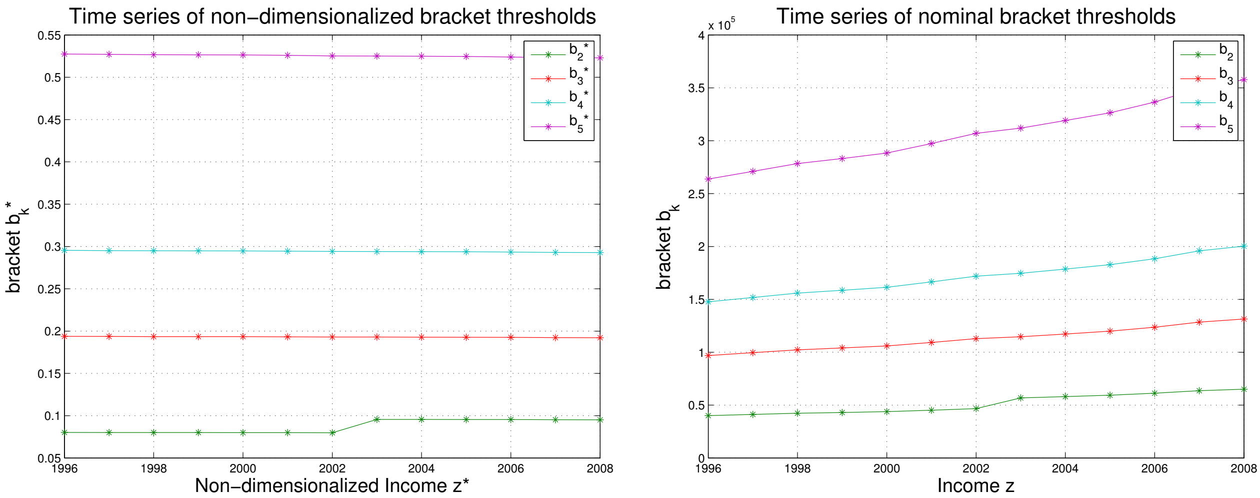



Figure 3.7: *Joint return tax brackets time series.*

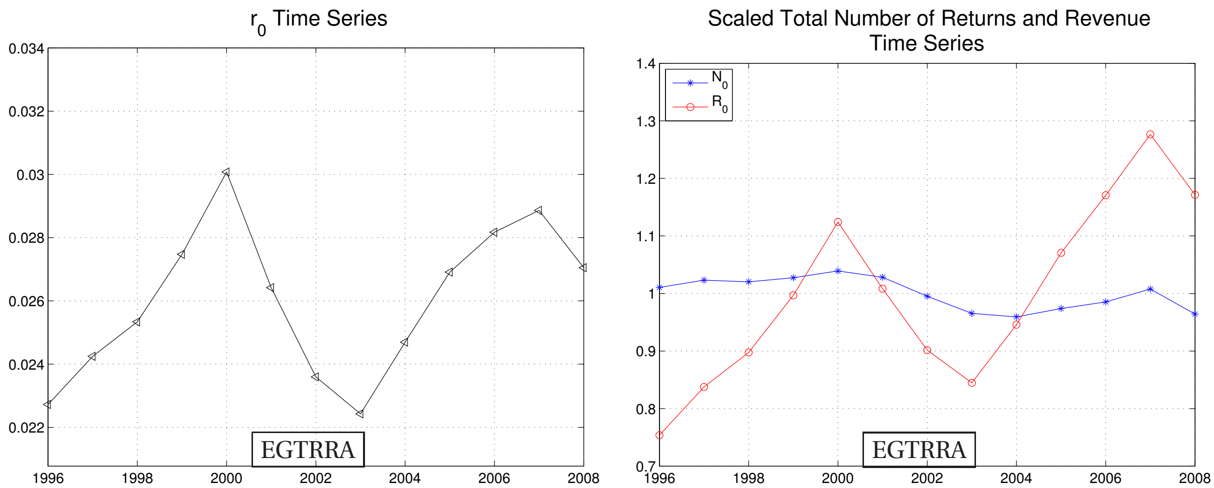



Figure 3.8: $r_0$ *and scaled* $R_0, N_0$ *time series.*

changes in peak revenue are due to changes in the nondimensionalized average tax rates. This sounds paradoxical at first as prior to EGTRRA, neither the marginal rates nor brackets nor the taxable income changed prior to EGTRRA. The nondimensionalized average tax rate has a time-dependent nondimensionalizing factor. We have,

$$\bar{r}(z) = r_0\,\bar{r}^*(z) \quad \text{where } r_0 = \frac{R_0}{z_0\,N_0}.$$

Although the average rate $\bar{r}(z)$ is virtually constant, the nondimensionalizing factor $r_0$ changes with time as $z_0, R_0$ and $N_0$ change with time. We plot on the left side of Figure 3.8 $r_0$ time series and on the right side $N_0, R_0$ divided by their respective means. The nondimensionalized marginal rates changed as nondimensionalization factors in the revenue and number of returns changed from year to year. More importantly, changes in $N_0$ are small and insignificant compared to changes in $R_0$. In addition, the inflation indexation that causes changes in $z_0$ is canceled with the same inflation indexation performed on $R_0$. In other words, changes in the nondimensionalized average rate is owing to changes in total revenue collections. Put simply, changes to the normalized revenue from 1998 to 2000 is a pure average rate effect, in turn owing to relative change in total revenue collection. The average rate effect is important here as the total revenue collection is a constant but the revenue response or redistribution of tax burden, varies across income ranges.

During the EGTRRA phase-in, the taxable income slightly decreased but was not large enough to counter the increase in both the average tax rate and number of returns. The increase in the revenue at the peak is primarily an average tax rate effect as the number of returns effect is three times smaller. Reductions in marginal tax rates resulted in year to year decreases in the total revenue collection. Consequently, the nondimensionalizing factor $r_0$ decreased. As a result, the nondimensionalized average tax rate increased which in turn implied an increase of the normalized revenue. In other words, the marginal tax rate cuts provided by EGTRRA implied an increased contribution to the relative revenue distribution at the peak. That is, larger marginal rate cuts in higher income ranges decreased their relative contribution compared to the middle income range. However, these gross effects cannot explain local changes to $f_t$ and again we note the hidden consequence of tax policy that it redistributes tax burden.

After the EGTRRA was fully phased in, the marginal rates were constant and once again the nondimensionalized average rates decreased as a result of the increase in $r_0$. The taxable income suddenly decreased substantially in 2005 as the exemptions and deductions provisions in EGTRRA come into effect. From 2006, the taxable income saw virtually no change prior to 2005. In contrast, the number of returns increases in 2005 but reverts to smaller positive changes in the remaining years. Increases provided in the number of returns are not large enough to counter the decrease brought by average tax rates. Hence, the decrease in the revenue at the peak after EGTRRA is an average tax rate effect. When the marginal rates remain unchanged, higher income ranges relative contribution increases as a result of the simple inflation indexed income tax system.

### 3.5.2 Invariant Points

One of the discoveries obtained by using the normalized revenue distribution is that it shows the responses of different income groups in the population to change in tax policy. As we have seen in previous sections, middle and high income taxpayers have different responses and sensitivities to taxation. As a result, there is a point of income where people are not responsive to changes in the tax structure in that year. This is the point where the change in behavior from the middle and high income ranges meet. We investigate the changes in these points here.

#### Pre-EGTRRA Era

Prior to the enactment of EGTRRA 2001, there were two groups of invariant points as shown in Figure 3.9. The first one, displayed in the first row, is between 0.2 and 0.3 on the nondimensionalized income scale. That is roughly between \$100,000 and \$150,000 (US 1996 dollar) on the nominal income scale. This is located on the start of wider income bins in the data but is still very close to the narrower lower income bins and as proven in the previous chapter, it is identified very precisely by the extended cubic spline estimation algorithm. The changes in behavior represented by the crossing of year to year distribution does not happen exactly at the same income value every year but is located in the same bin. Thus, they form the clouds of invariant points whereby taxpayers are classified according to their behavior as represented by their location on the income axis. The implication is clear, at $z^* = 0.4$ on the non-dimensionalized income axis, the revenue density follows the reverse trend of that of the peak of the distribution. In contrast, the nominal revenue density distribution shows parallel curves shifting to the right from year to year. In other words, taxpayers in this neighborhood have the same response to the current tax structure, that is an increase in the revenue density from year to year. This shows once more the inappropriateness of using the nominal distribution to compare year to year changes, as the invariant points are completely missed if nominal distribution is used.

A second and possibly a third group of invariant points is located between 0.5 and 0.8 on the nondimensionalized income scale, around \$250,000 and \$400,000 on the nominal scale, as depicted in the second row of Figure 3.9. These are located on the widest income bins from the data and, as shown in the previous chapter, are not very reliable. First, their existence is questionable as they could just be the result of numerical errors or smoothness requirement. Second, if they are present, their location cannot be pinned down as accurately as the first group of invariant points owing to wide income bins here. However, if they are present, intermediate value theorem from calculus tells us that there must be at least two of them in this region. Indeed, the normalization of the revenue distribution implies that the consistent drop around its peak must be balanced by a proportional increase in revenue at higher income ranges. As the Pareto tails of the annual distributions become parallel to each other over the income axis, it follows that the revenue density must have the reverse trend of the peak before the start of the Pareto tail of the distribution. Since the first group of invariant point results in this reverse trend of the peak already, any invariant point between the first one and the start of the Pareto tail will reverse that trend and therefore must be followed by another invariant point to switch it back to the opposite trend of the peak.

With or without the second and third group of invariant points, a clear conclusion is that taxpayers behavior changes before the start of the Pareto tail and that those invariant points represent taxpayers that are indifferent to the underlying changes to the income tax system and indirectly to the economy as a whole. As before, the nominal revenue distribution shifts systematically to the right and does not allow us to draw a conclusion.

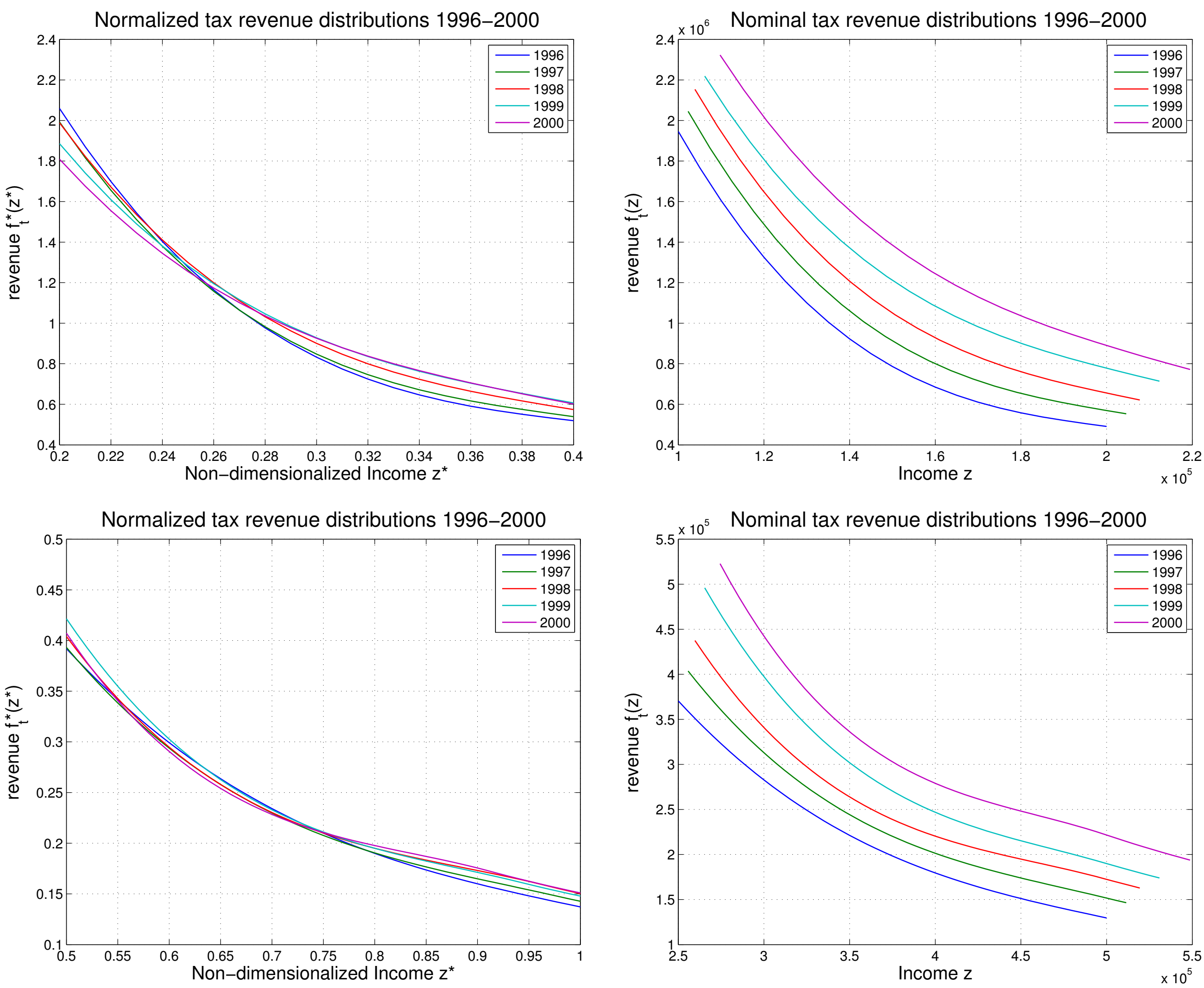



Figure 3.9: *Clouds of invariant points prior to EGTRRA 2001 shows a tendency to cluster at* $0.24 \leq z^* \leq 0.26$.

**EGTRRA Phase In**

Figure 3.10 displays the revenue distribution in the same income range during the EGTRRA phase in period. On the first row, between $z^* = 0.2$ and $z^* = 0.3$ nondimensionalized income, the 2001 normalized revenue curve crosses that of 2004 whereas the 2003 curve crosses that of 2002. The 2002 and 2004 curve join each other from 0.3 to 0.4. As we observe year to year changes, the only true invariant point here is at $z^* \simeq 0.25$ where the 2002 and 2003 curves intersect. One can conclude that taxpayer's behaviors were opposite above and below that point from 2002 to 2003 but only for those two years. The rest are two to three years invariant points. The nominal revenue distribution on the left shows the opposite trend of the pre-EGTRRA period. That is a shift to the left implying a decrease in revenue during the years where the marginal tax rates were reduced. This implies a decrease in revenue over that range but again, relative to the whole distribution the normalized distribution reveals that no such systematic year to year decrease occurs.

On the second row, we plot again the revenue distribution between $z^* = 0.5$ and $z^* = 1$ on the nondimensionalized income scale. Clear and specific invariant points are inconclusive. All the curves from 2000 to 2004 are close at $z^* = 0.5$. They separate in a non-uniform way over $z^* = 0.65$ to $z^* = 0.9$ where they are stuck together to form one thick curve until 1 except for the 2004 curve which lies above the rest from $z^* = 0.55$. As before, these curves are not reliable estimates of the true distribution and no firm conclusion can be drawn. Nonetheless, as the trend of the peak were not reversed in the [0.2 0.3] range, if the invariant points do not occur in the [0.5 1] range it is likely to happen around 1 where the Pareto tail of the distribution will start. In this case, only the general conclusion as to the existence of at least one invariant point can be drawn as the point 1 is again arbitrarily chosen, albeit some arguments in its favor as the start the Pareto tail. Again, the nominal distribution shows a shift to the left implying a reduction in revenue over that range.

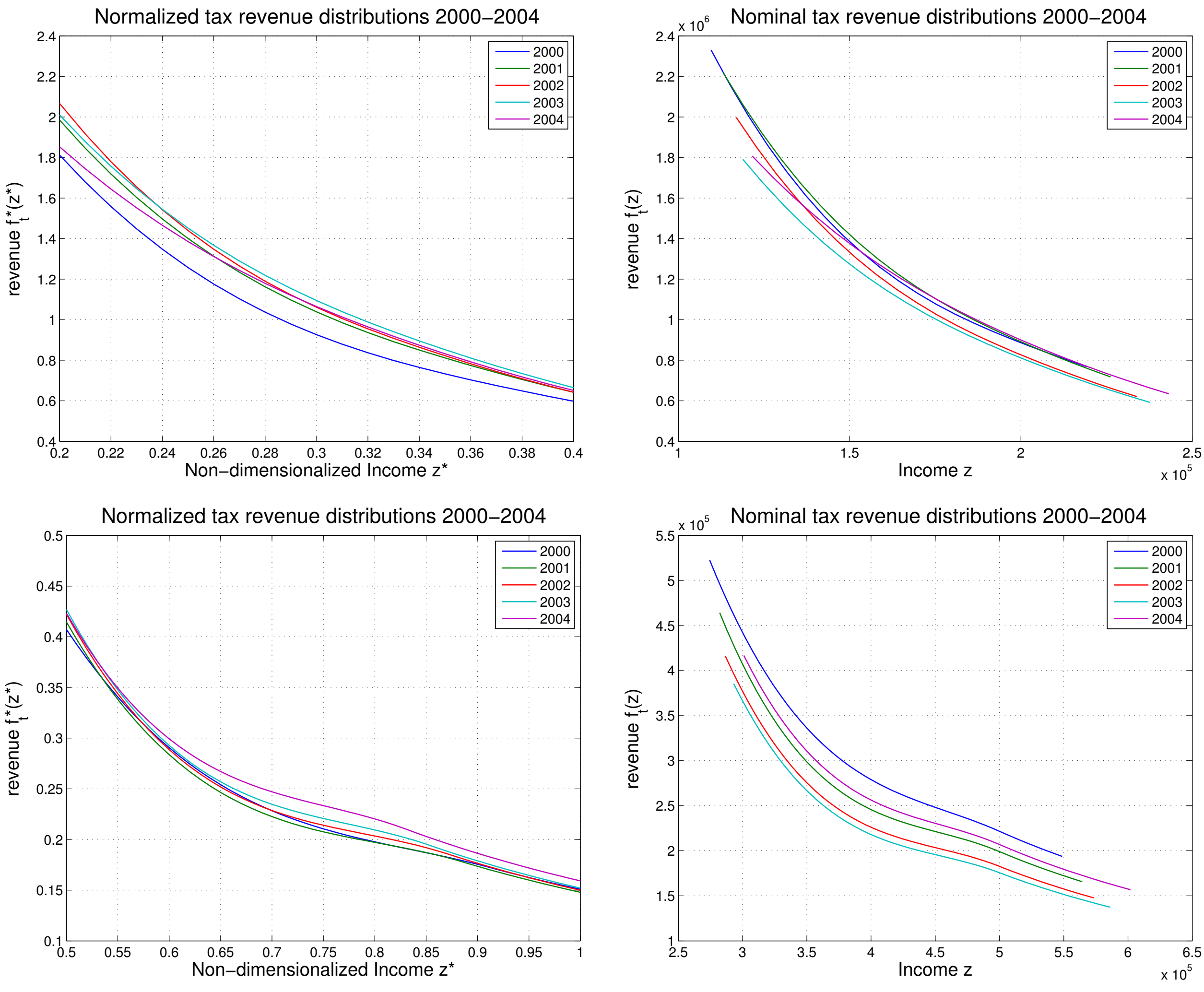



Figure 3.10: *Invariant points during EGTRRA phase-in period show no tendency to cluster.*

**post-EGTRRA Era**

Figure 3.11 depicts the revenue distribution from 2003 to 2008. The progression in higher income ranges from 2003 to 2007 shows clearly the counter-effect from the decrease in the middle income revenue trend. From 2005 to 2007 the normalized revenue curve is very stable and there is not much difference from year to year. The most striking result here is that there are two invariant points. There is no cluster here. Both points are identified precisely at 0.22 and 0.31 for all years between 2005 and 2007. Again, these points are in the narrow income bins of the data and are therefore estimated accurately by the extended cubic spline approximation. Invariant points indicate constant taxpayer behavior but as the curves were narrowly held to each other, this change is not significant. The 2008 revenue density is increased from its 2007 level. However it does not intersect the 2007 curve and thus does not imply any invariant point. As prior to EGTRRA, the steady decrease in the normalized revenue corresponds to a consistent shift to the right in the nominal curve. This is an increase in the year to year revenue over this range but relative to the overall distribution no such increase is observed as evidenced by the normalized curve.

The revenue density on [0.5, 1] nondimensionalized income range is shown on the second row of Figure 3.11. As there were two invariant points in the [0.2, 0.4] range, one cloud of invariant point has to be in the [0.5, 1] range. It is located at 0.576 on the nondimensionalized income scale. Again, this is located on the widest income bins of the data and therefore is not precise. This invariant point indicates the switch in taxpayer's behavior marking the end of the middle income range. The year to year decrease in the peak of the distribution is balanced with a corresponding increase above the last invariant point. As before, the nominal revenue distribution systematically shifts to the right.

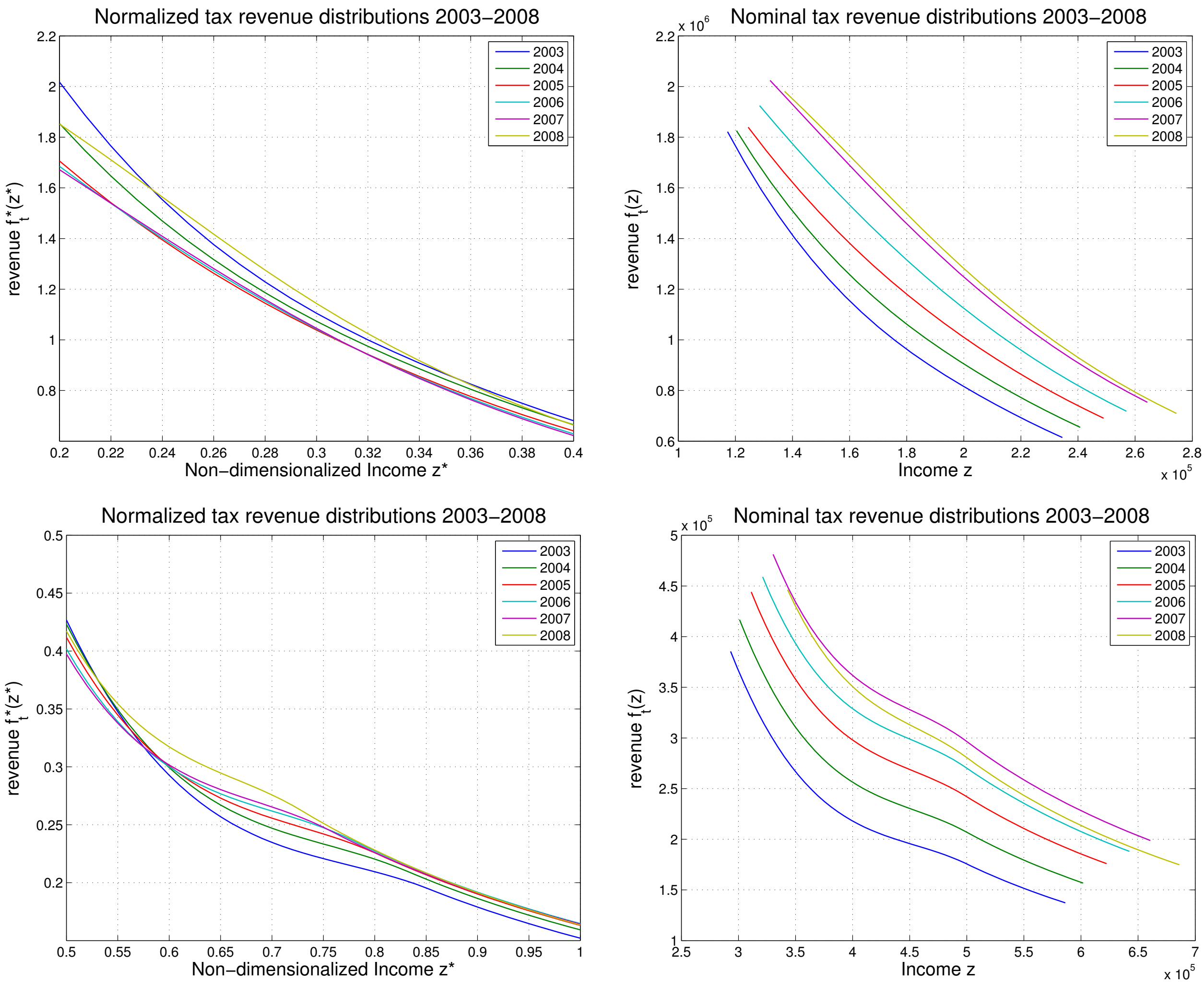



Figure 3.11: *Clouds of invariant points after EGTRRA is fully phase-in.*

### 3.5.3 High Income Range

Figure 3.12 displays the time series of the revenue, number of returns distributions, average rate and taxable income at $z_0$, the income threshold from the data above the highest bracket threshold or equivalently at the nondimensionalized income $z^* = 1$, from 1996 to 2008. This is a particular point of interest but it should also be understood as a representative of all high income taxpayers. Behavior at this point is the same over the high income range as there is no invariant point in the high income range. Intuitively, behaviors are the same but the magnitude of any response or properties are not the same.

At high income level, the overall trend of the normalized revenue distribution is increasing with some fluctuations from 1998 to 2003 and a decreasing trend in 2008. The period of fluctuations in revenue corresponds roughly to the years where there is no clear invariant point before $z^* = 1$. Therefore, as the invariant point is around $z^* = 1$, changes in taxpayers behavior occur near that point. Hence the slight decrease or increase according to where $z^* = 1$ is right before or right after the true invariant point. In contrast, the nominal revenue distribution increases prior and after EGTRRA and decreases during its phase-in period. This is the absolute change in revenue. At high income level, the average rate starts to converge to the marginal rate. Hence, reduction in marginal rates implies reduction in revenue.

The normalized and nominal distribution of number of returns follow similar patterns, with different magnitude as expected. During the pre-EGTRRA era, the number of returns increased. This decreased at a slower rate while the tax rate reductions where phasing in. It increased again after the tax rate reductions where fully phased in. At first, high income range taxpayers follow an inverse Laffer effect during the rate reduction phase-in period but combining the overall observation with the pre-EGTRRA and after EGTRRA periods, we deduce that rate reduction causes taxpayers to move to lower income range whereas constant rate pushes them to higher income ranges. Thus, the number of returns increases or decreases with constant rates or with rate reductions. Significantly worse economic conditions in 2008 decreased taxpayer's income and therefore decreased the number of returns at $z^* = 1$ as a general trend in the high income range.

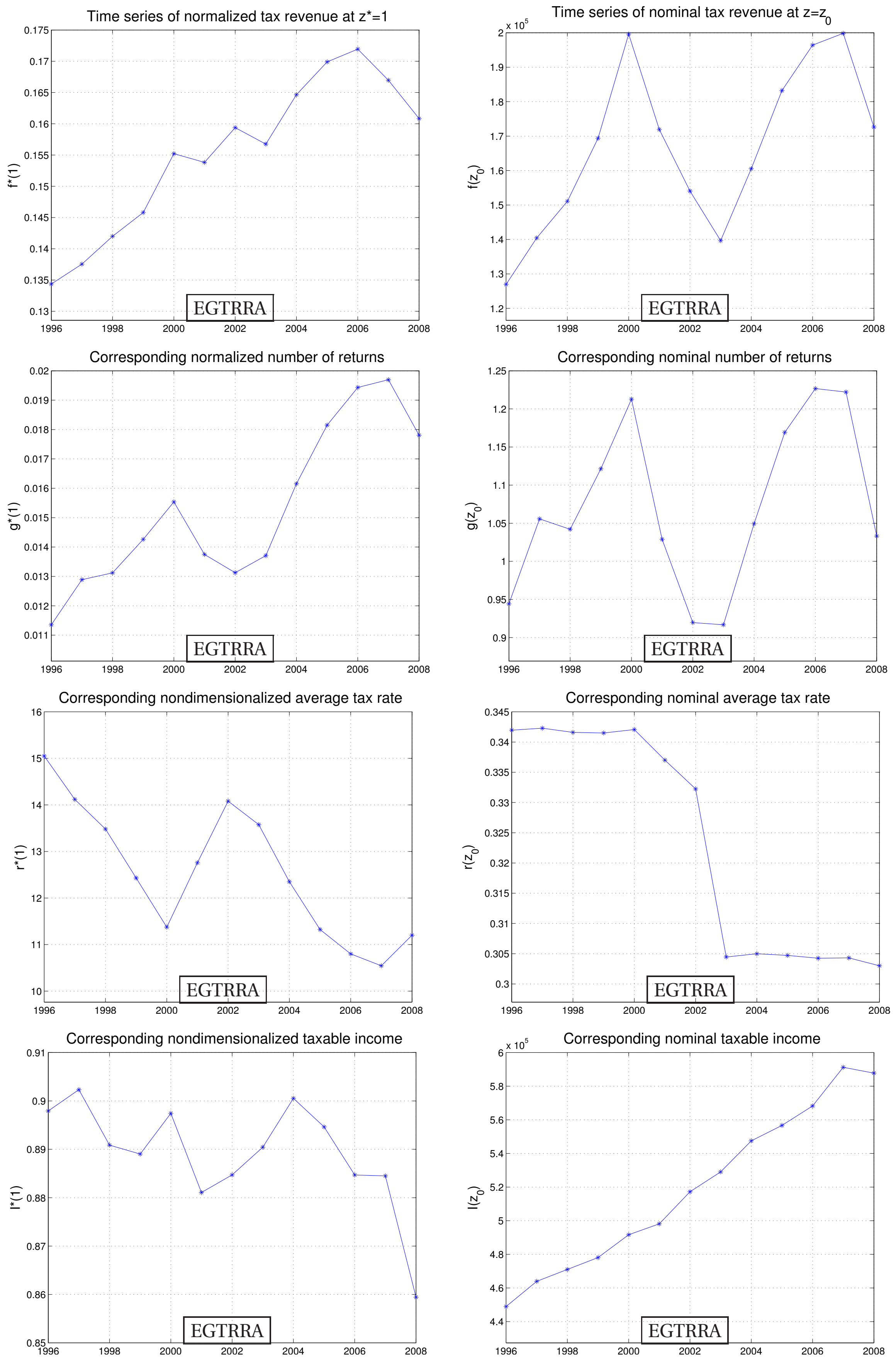



Figure 3.12: *Time series at high income levels, any* $z^* \geq 1$ *or equivalently any* $z \geq z_0$*. Note* $z_0$ *changes every year due to the inflation indexation. See Table A.12 in Appendix A for a complete list.*

The nondimensionalized average rate follows the opposite trajectory to the nominal revenue distribution. It decreases while there is no change in rate. While the tax rate is reduced, the corresponding nondimensionalized average tax rate actually increased. We note this is a similar trend to the middle income range, especially at the peak of the revenue distribution but the magnitude and ranges are not the same. The nominal average rate follows the same trend as the marginal rates at high income range. They are not the same as the average rate is still lower than the marginal rates but reduction in marginal rates corresponds to reduction in the nominal average rate.

The nominal taxable income increased every year except for 2008 where it decreased slightly. In contrast, the nondimensionalized taxable income decreased overall with some increases between 1996 to 1999 and 2001 to 2004. This is a direct result of simple use of inflation indexation above. The steady increase in nominal terms is not steep enough to overcome the inflation factor and thus result in a decrease when deflated.

Figure 3.13 shows the year to year difference in the normalized revenue and number of returns as well as the nondimensionalized average rate and taxable income at $z^* = 1$. As before, we decompose year to year changes in revenue into a number of returns effect, average tax rate effect and taxable income effect.

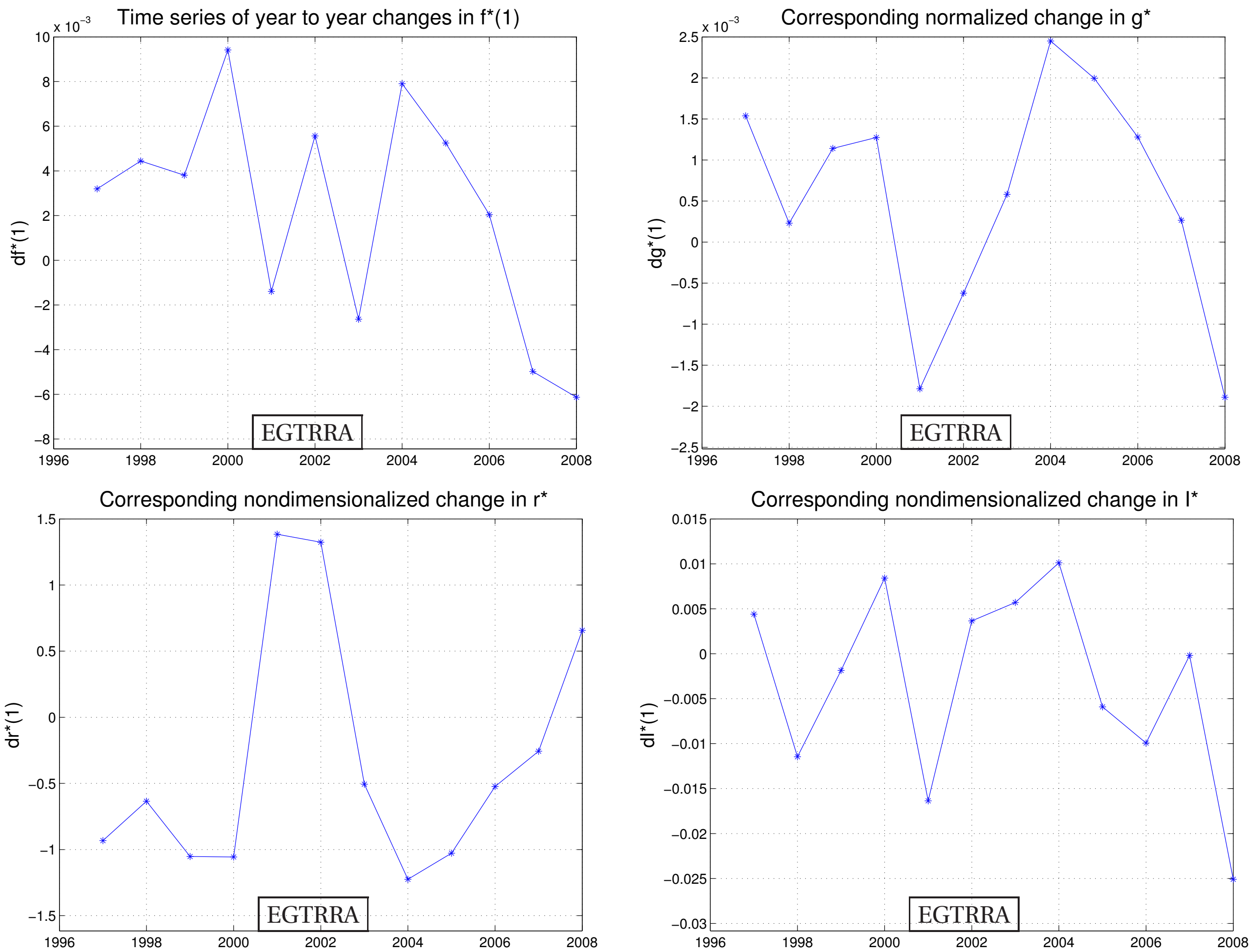



Figure 3.13: *Year to year changes in* $f^*(1)$, $g^*(1)$, $\bar{r}^*(1)$, $I^*(1)$.

From 1996 to 1998, the average rate decreased but this effect is not large enough to counter the positive effect from both the number of returns and taxable income and thus the revenue increased. This is primarily a number of returns effect as the taxable income effect is ten times smaller. Increases in number of returns pushed the revenue to increase. From 1998 to 2002, the revenue fluctuations are a sequence of the three effects whereby one effect is marginally larger than the rest. However, the magnitude of the fluctuations in revenue is very small, one cannot draw a solid conclusion. Again, this is due to the uncertainty about the location of the invariant point for those years.

From 2004 to 2007, both the taxable income and average tax rates were decreasing steadily but those effects are not large enough to balance a steeper increase in the number of returns. The revenue increase at high income range is a number of returns effect. After EGTRRA is fully phased in, nondimensionalized average rates decreased and in response the relative number of returns increased to increase the relative revenue contribution. In 2008, as the number of returns as well as the taxable income decreased, the increase in the average rate was not strong enough and the revenue stagnated.

### 3.5.4 Very High Income Range

For completeness of our study, we observe the changes in the distribution at income levels where we fit the Pareto distribution in the extended cubic spline procedure. We plot in Figure 3.12 the time series of the revenue, number of returns distributions, average rate and taxable income at $z = 10z_0$, from 1996 to 2008. Any point in the Pareto tail will have the same year to year relative behavior as two Pareto distributions can only cross each other at one point. The intersection point in our cases falls in the cubic spline part of the distribution, making the Pareto tails parallel to each other while asymptotically vanishing at different rates to infinity.

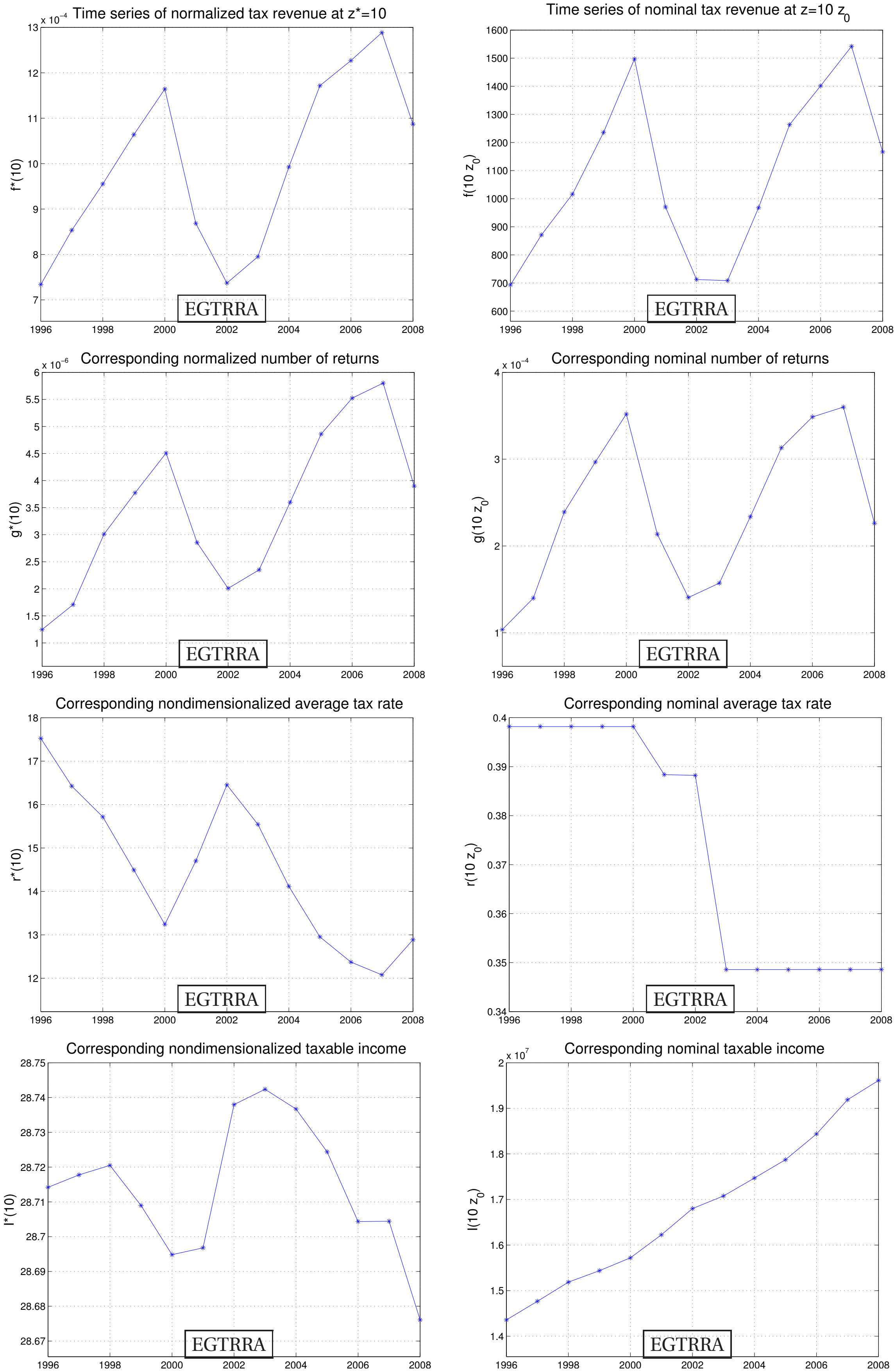



Figure 3.14: *Time series at very high income levels, at $z = 10\,z_0$.*

Nondimensionalization does not affect the shape of the revenue distribution. The normalized and nominal distributions have the same pattern. The revenue distribution increased during the years where there is no change in the tax rates. Unlike the high income taxpayers behavior, the revenue decreased rapidly to its 1996 level while the tax rate reductions of EGTRRA was phasing in. It increased faster above its 2000 peak in by 2007. Again, the financial crisis and the glooming recession in 2008 decreased the revenue below its 2005 level. Very high income taxpayers are very sensitive to tax parameters and economic conditions as they are the ones who invest major parts of their income.

The corresponding distribution of number of returns follows the same pattern. Slow increase precedes EGTRRA before it jumps in 2000, followed by a fast decrease while the tax rate changes where phased in increased afterwards to higher level above its 2000 peak by 2007. In 2008, the number of returns decreases below its 2005 level. Again, nondimensionalization does not affect the distribution. Normalized and nominal number of returns follow the same pattern and are proportional to each other. This is to be expected as a result of the Pareto tail of the number of returns distribution. However, the implication is a major shift from the supply-side economic theory where it is understood that decrease in tax rates such as those provided by EGTRRA encourage economic activities and stimulate growth. Here, the opposite is observed, EGTRRA implied temporarily a decrease in number of returns. Switching to nondimensionalized values corrects this contradiction as nondimensionalized average tax rates clearly increased during the EGTRRA phase in. All this is a hidden consequence of simple inflation-adjusted tax policy.

The time series of the average tax rate did not change much from the high income level. The nondimensionalized average rate decreased during the pre-EGTRRA era, increased while the statutory marginal rates were reduced and decreased again when the rate reduction where fully phased in. Again, these are as a result of the changes in the total revenue collection $R_0$ on the average tax rate nondimensionalizing factor $r_0$. The nominal average tax rate followed the same change as the marginal tax rate. In fact, at the very high income range, the nominal average rates converge to the top marginal rates. This one of the reason why they are particularly sensitive to changes in tax rate, at this income level.

The nominal taxable income of the very rich increases every year from 1996 to 2008. However, the year to year increase is not steep enough so that inflation indexation translates this into a corresponding increase in the nondimensionalized income. Instead, the nondimensionalized taxable income first increases from 1996 to 1998 then decreases until 2000. This is followed by year to year increases until 2003 after which it decreases until 2008 with constant taxable income between 2006 and 2007. In other words, changes to exemptions and deduction provisions saw a stable increase in the very high income taxpayers taxable income but in real terms their taxable income varied and decreased consistently after EGTRRA was fully phased in.

Figure 3.15 shows the year to year difference in the normalized revenue and number of returns as well as the nondimensionalized average rate and taxable income at $z^* = 10$. Again, we decompose the year to year change in revenue into a number of returns effect, average tax rate effect and taxable income effect.

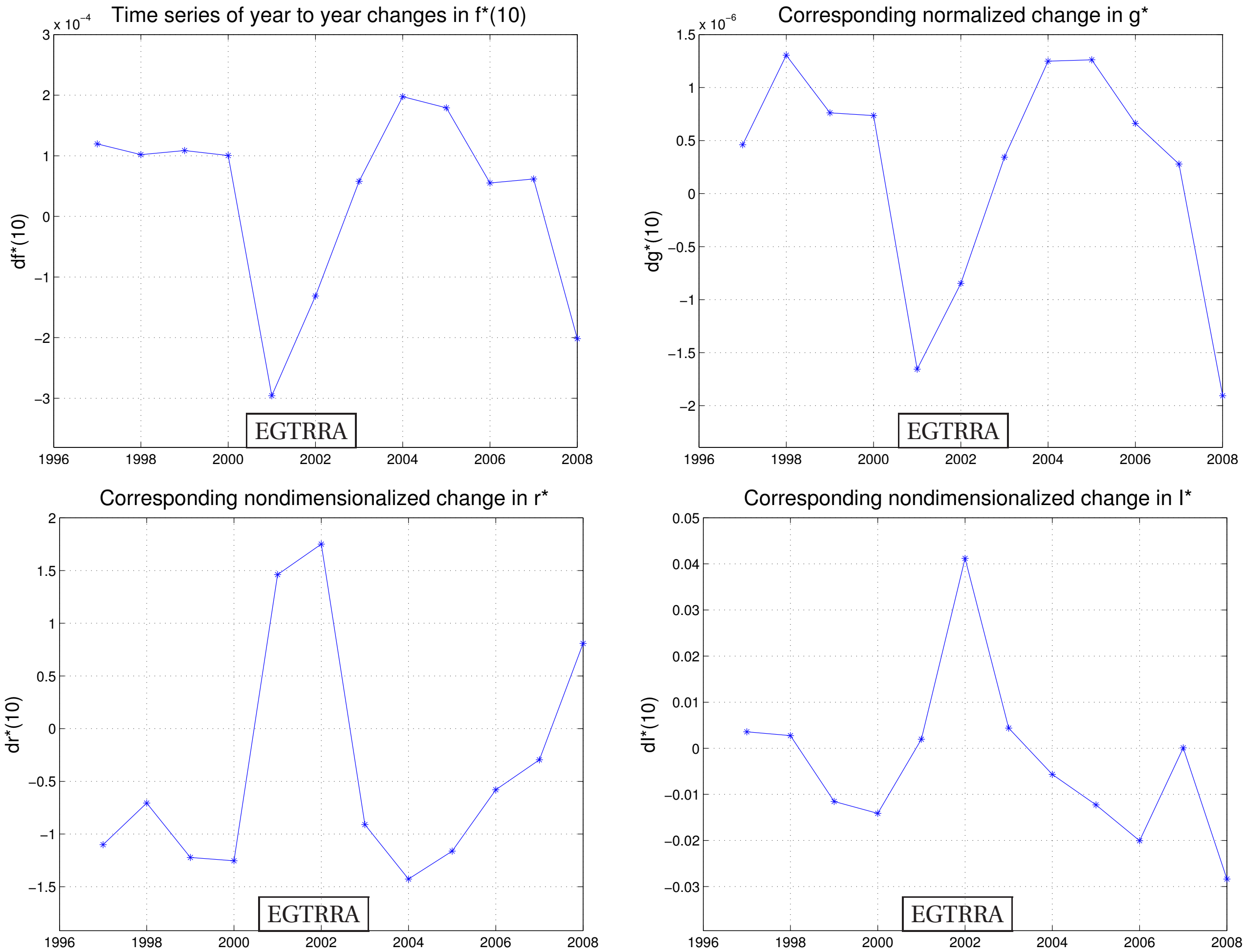



Figure 3.15: *Year to year changes in* $f^*(10)$, $g^*(10)$, $\bar{r}^*(10)$, $I^*(10)$.

Prior to EGTRRA, the change in average rate decreased and the taxable income increased from 1996 to 1998 but decreased from 1998 to 2000 whereas the number of returns increased consistently. When the taxable income increased from 1996 to 1998, this contributed half the number of returns effect. Hence, the revenue increase prior to EGTRRA is a number of returns effect. In other words, no change in marginal tax rates increases the number of returns in very high income taxpayers and that in turn implies higher relative revenue. As mentioned before, no change in marginal rates translates into a decrease in the nondimensionalized average rate that justifies an increase in the number of returns as a response. This is a pure indirect Laffer effect on revenue.

During the EGTRRA phase in period, we have the reverse trend. The number of returns decreased whereas both the average rate and the taxable income increased but the revenue still decreased. Clearly, this is yet another number of returns effect. As before, the decrease in number of returns is an indirect effect from the increase in the nondimensionalized average rate. Indeed, the marginal rate reduction brought by EGTRRA corresponds to nondimensionalized average rate increase that in turn implies a decrease in the relative number of returns, as taxpayers behavioral responses. This is also a pure Laffer effect on very high income taxpayers.

From 2003 to 2007, the original Laffer effect prior to EGTRRA is observed. Both the taxable income and average tax rates decreased, the number of returns increased as an indirect response and the revenue increased. In 2008, the average rate increased as a result of the financial crisis, the number of returns decreased and the relative revenue contribution decreased. This is also a Laffer effect on very high income taxpayers from the financial crisis and not EGTRRA as in the previous years.

#### Behavior of the Pareto parameters

The Pareto distribution has two parameters, the shape parameter and the associated height parameter. As both the revenue and number of returns distributions are not affected by normalization and nondimensionalization, we are interested in the link between year to year changes in the distribution and its respective shape parameter $a$ and/or height parameter $\alpha$ from (2.23). The shape parameter $a$ of the Pareto distribution is called the Pareto parameter in this section.

Taxpayers behavior at the very high income range is captured by the associated Pareto parameter. We plot in Figure 3.16 the time series of Pareto parameters for both the revenue and number of returns distribution. The changes in the Pareto parameter of the revenue distribution are similar to changes in nondimensionalized average tax rate of the very high income taxpayers. The reason is simple changes in Pareto parameter have the opposite effect to the change in revenue as the Pareto parameter is a negative power to the revenue distribution. That is, an increase in the Pareto parameter results in an decrease in revenue and inversely. Meanwhile, the revenue distribution responds negatively to nondimensionalized average tax rate for the very high income taxpayers as explained in the previous section. In other words, the overall very rich taxpayer's responses to changes in the tax system and the economic conditions at large are summarized by the Pareto parameters.

The same conclusion can be reached for the Pareto parameter of the distribution of number of returns. As an indirect Laffer effect from the nondimensionalized average tax rate at the very high income range, the number of returns increased while there is no change in tax rates and decreased when the marginal rates were cut. In contrast, the Pareto parameter decreased before and after EGTRRA and increased while the tax rates were reduced during the EGTRRA phase-in period. The Pareto parameter captures the relative changes to distribution of number of returns at the very high income range.

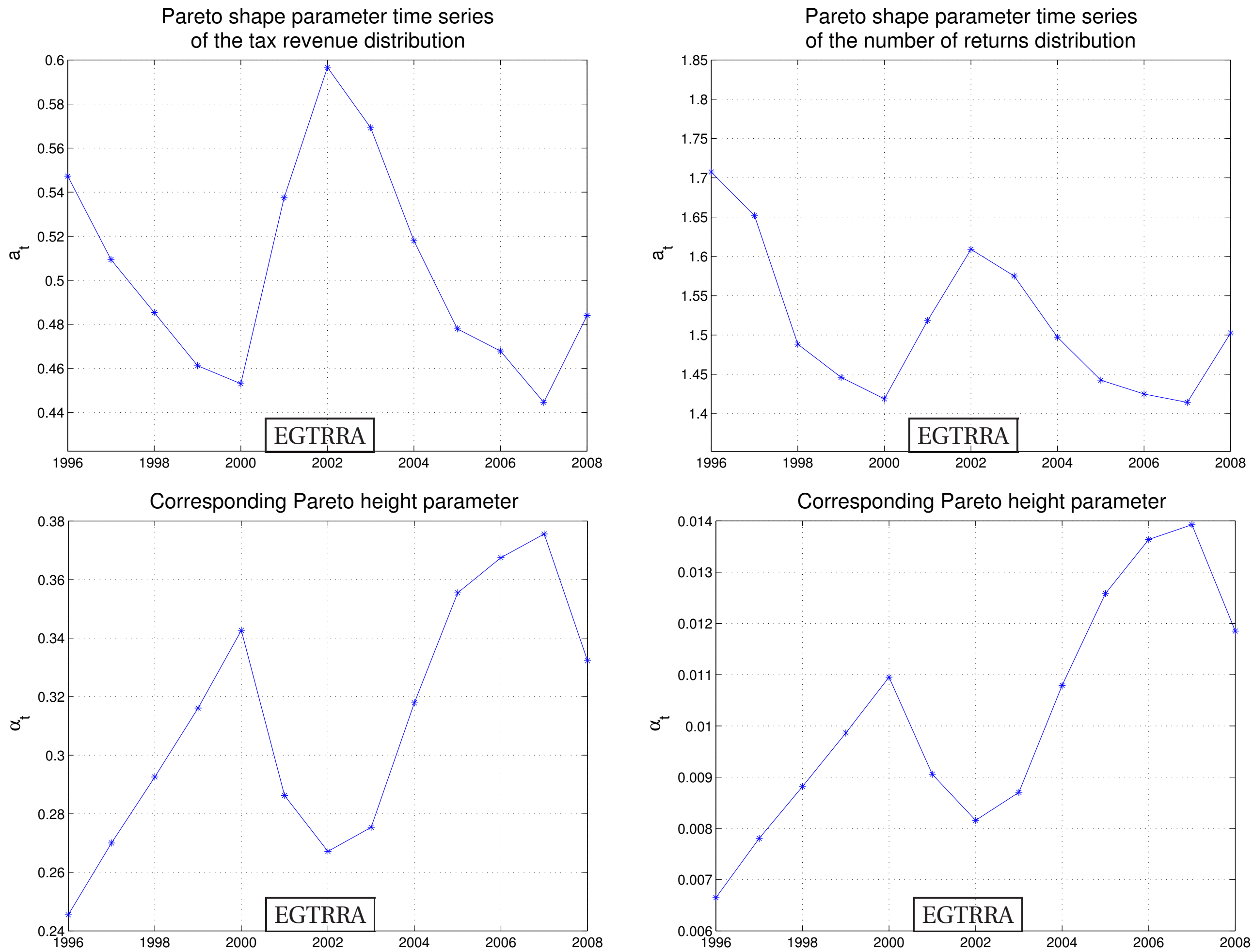



Figure 3.16: *Shape $a_t$ and height $\alpha_t$ Pareto parameter of the revenue and number of returns distributions time series.*

### 3.5.5 Income Range Comparison

To study interactions and relative behaviors between income classes, we plot in Figure 3.17 the revenue, number of returns densities, average tax rates and taxable income at its peak, at the nondimensionalized income $z^* = 1$ and $z^* = 10$. They are scaled by their average value from 1996 to 2008 for each item plotted. For example the time series of the revenue distribution at the very high income levels is divided by its average over time. As the revenue has different order of magnitude at different income level, scaling is necessary in order to fit them in the same figure. The same argument goes for the distribution of number of returns and average tax rates. Note, the scaling is lost here but the relative trend is preserved.

Plotting all income ranges on the same figure highlights the importance of normalization and nondimensionalization when it comes to comparing year to year distributions. Prior to EGTRRA, the middle income taxpayers showed their relative revenue contribution, to the normalized distribution, decreased consistently fast. It is increasing overall at high income range, transitioning to the reverse of the middle income range trend, reached at the very high income taxpayers, as compared to the nominal distribution where all three income ranges increased and decreased roughly together, during the same period. The normalization process reveals the relative contribution of taxpayers from different income ranges. Decreases in some ranges must be balanced by increases in other ranges.

In contrast, normalization does not affect much the relative trend between different income ranges to the number of returns. This is because the number of returns is inflation dependent as proved in the previous chapter and the total number of returns did not change much over the years. However, normalization still implies that increases and decreases in the middle income ranges must be balanced by decreases and increases respectively in high and very high income ranges.

There is no relative change between the high and very high income ranges nondimensionalized average tax rates. The middle income range is slightly higher but all three ranges of nondimensionalized average tax rates decreases when there is no change in the marginal rates and increases when the marginal rates decreased. The middle income range nondimensionalized average rate also increased when the nominal average rate increased as a result of the increase in the taxable income at the peak of the revenue distribution. There is irregular variation in the nondimensionalized taxable income as a result of the inflation indexation of the roughly uniform increasing nominal taxable income but it is constrained within 10% for the middle income range, 3% for the high income range, whereas the very high income range is virtually constant relative to the other two income ranges. In other words, changes to the deductions and exemptions rules do not affect much the high and very high income taxpayers taxable incomes relative to the middle income taxpayers as expected.

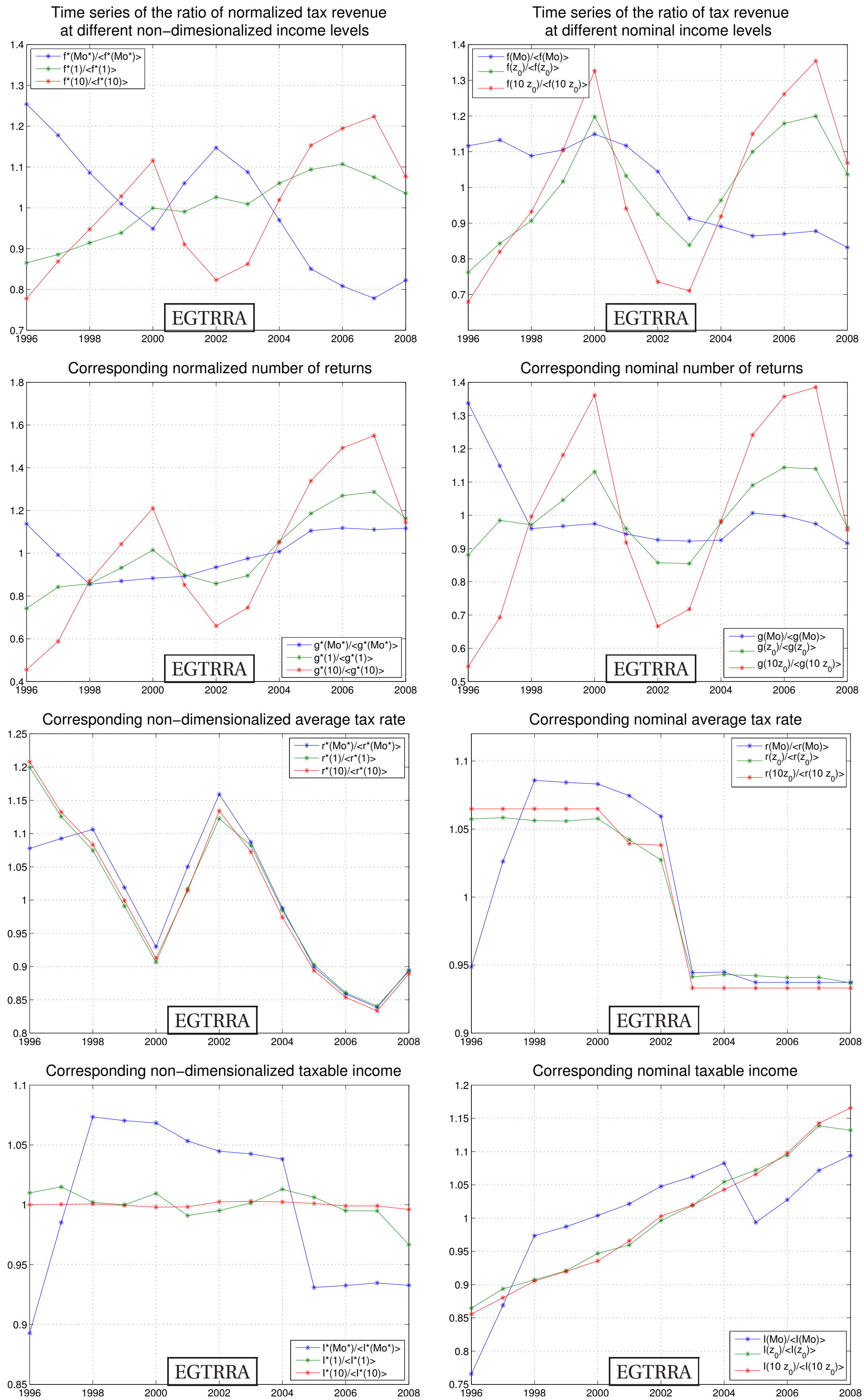



Figure 3.17: *Comparison of middle (blue), high (green) and very high income (red) sectors. The average rate* $r^*$ *is similar for all income groups. Both the revenue and the number of returns distributions have opposite effects at the mode and the Pareto income levels.*

## 3.6 Conclusion

Normalization and nondimensionalisation shed a new light in the dynamics of the US individual income tax revenue distribution. Normalization is necessary to evaluate the relative changes to the distribution from different income ranges. That is, to evaluate redistribution of the tax burden. Nondimensionalization as a means to facilitate numerical computations also ensures that the same income ranges are used in the year to year comparison of distributions. We have showed in this chapter multiple features of the US income tax distribution, undiscovered through the use of the simple nominal revenue distribution.

### 3.6.1 Redistribution of the tax burden

The US joint returns revenue distribution showed two opposite responses across income ranges. The middle income range, near the peak of the distribution, decreases every year when there is no change in the marginal tax rates. There is a cloud of invariant points on the income axis that are indifferent to changes in the tax system. In contrast, the high income range revenue distribution increases correspondingly year after year. This is a consistent shift in the distribution of the burden to the rich. In section 3.5, we have clearly shown that the falling revenue peak $f^*(Mode)$ in pre- and post-EGTRRA periods is associated with changes in average tax rates $\bar{r}^*(z^*)$, and not with changes in number of returns or deductions and exemptions. In turn, during these periods marginal tax rates are constant. Tax brackets, however, moved and constitute the single action of the IRS to ensure adequate revenues. However, Figure 3.7 proves that indeed even tax brackets are constant and cannot account for falling of the peak $f^*(Mode)$. It follows that, while real annual tax revenues are as desired, the policy of bracket adjustment with inflation leads to a systematic distortion of the tax burden from the middle income to the very rich.

it into a completely high income tax distribution only. Collapse of the peak will presumably occur if present taxpayers attitudes to tax policy do not change. Change in taxpayers attitudes will demand, at least, that $f_t^* \simeq constant$. Deeper critique of the optimal shape of the distribution will surely arise.

The tax rate reductions brought by EGTRRA reversed the decreasing trend of the revenue but only temporarily as the rate changes are occuring only a short period of time. The nondimensionalised average tax rate measures the applied average rate of every taxpayer. Government intervention such as marginal tax rate reductions or exogenous changes in economic conditions are captured in the nondimensionalised average rate. Changes in rates and/or changes in economic conditions and business cycles such as a recession is reflected in the total revenue collection and number of returns. Taxpayer's responses in the middle income range are an average rate effect primarily through the effect of the annual total revenue collection $R_0$ on average rate nondimensionalizing factor $r_0$.

Consider the normalized distribution of number of returns $g^*(z^*)$. The years preceding EGTRRA showed a constant shift of the middle income range taxpayers moving to higher income ranges. The rate of change is not as dramatic as for the revenue distribution. While the EGTRRA tax rate reductions were phased in, the shift of number of returns reversed back because taxpayers were progressively moving towards lower income ranges. This is again temporary as when the tax rate reductions were fully phased in, taxpayers shifted back to higher income ranges. We confirmed the sensitive Laffer effect responses of the very high income taxpayers. This is a reverse indirect effect of the average tax rates on the number of returns.

the instability responds to marginal tax rate changes, how should they be changed so as to preserve the shape of the distribution.

# Chapter 4

# Year to Year Control of Instability in the Individual Income Tax Revenue Distribution

## 4.1 Introduction

There is a vast amount of literature in the economics of income taxation following the seminal paper of Mirrlees (1971). His theory looks at individual utilities given their skills endowments and aggregates them on a welfare function that is to be maximized, given the government revenue constraints (see Tuomala (1990) for a comprehensive review). This formed the basic foundation of modern economics of income taxation. The main drawback of this theory as noted by Saez (2001), is that it is very general with few practical results and leads to complicated formulas, not understood in the applied taxation world.

The more applied alternative is the theory of optimal tax reforms (see Buddelmeyer et al. (2007)) where one looks to small tax changes to improve existing tax structures. Microsimulation models are used to estimate marginal welfare effects that show the direction of improvement (a technical review is given by Creedy and Herault (2009)). Cross-sectional tax data are used to simulate individual behaviors but the theory is again based on individual utilities and effects on the aggregate welfare function are obtained. The alternative to those two mainstream theories that we propose in this thesis is to discard individual behaviors and utilities altogether and look directly at the aggregate group level through the dynamics of the revenue and number of returns distributions.

Naturally, by-passing the individual utility formulation prevents us from using the same welfare function as is customarily done in the literature. As a result, we cannot compare directly our approach to the results in the field thus far. In fact, we start a new branch of taxation theory where optimality has to be redefined. The main emphasis in this theory is to obtain required revenue and keep the system in that regime. In a sense, this is the third best option compared to the Mirrlees framework that looks directly into the optimal rates and the microsimulation approach that points to the direction of optimal reforms. The motivation here is that in practice major tax reforms are rather rare, happening every 20 years or so in the US, and solutions from economic theory are subject to government intentions together with the political debate that accompanies them. We give a direct mathematical approach to sustain currently applied policy.

In the previous chapter, changes in the year to year revenue and number of returns distributions show changes in the relative contribution and how many contributed to the revenue precisely at each income level. By plotting the yearly revenue distribution, there is a consistent drop around the peak, matched by a corresponding increase in higher income range. Tax parameter changes brought by EGTRRA temporarily restored that trend. This motivates two points. Firstly, the shape of the distribution has to be controlled because the currently applied tax policy (of holding tax rates and adjusting brackets by the inflation rate) consistently changes the revenue distribution. Instead of attaining the optimal tax as in Mirrlees (1971)'s work or improving tax reforms using microsimulation models, one could look at preserving the current normalized revenue. That is, we keep constant the current relative revenue contribution. The consistent drop in the peak of the revenue cannot be going on forever because the distribution of revenue could migrate to a qualitatively different one. Secondly, one has to control every year as the restoration of tax distribution in 2000-2002 lasted only while EGTRRA policy adjusted marginal rates.

In this chapter, we will use the identity, tax revenue = tax liability × number of returns, to obtain a simple formulation that will keep the normalized revenue distribution unchanged from one year to the next. Controlling the overall revenue shape from year to year ensures that the consistent drop around the peak is vanishes. The end result of that year to year control should be the optimal tax parameters as those are understood by tax policy makers and the public in general. This chapter is organized as follows. In section 2, we define the year to year changes in the revenue distribution as an instability. In section 3, we derive the year to year control of the instability. In section 4, extensions and sensitivity to the number of returns distribution estimation errors are discussed. We conclude in section 5.

## 4.2 Instability in the normalized revenue distribution

In the theory of dynamical systems, an equilibrium solution of a system of differential equations, describing the trajectory of a physical process, is said to be stable if every other solution of the system that start close to the equilibrium solution remain close to it at all times, Lyapunov (1892). Instability refers to solutions whereby small perturbations to its initial conditions lead to solutions diverging from the equilibrium as time goes on.

As shown in the previous chapter, the current policy that consist of indexing tax brackets by the inflation rate associated with economic growth lead to a year by year consistent drop in the normalized revenue peak and a corresponding increase in the revenue for higher income ranges. This constitutes an instability in the normalized revenue distribution as year by year, the relative contribution shifts more and more to higher income ranges.

At fixed income $z$, the time trajectory of the normalized revenue distribution tends to zero at lower income and increases more and more at higher income. This calls for a systematic control, that is, changes in bracket and/or marginal rates policy, to contain the consistent drop in the normalized revenue distribution. If no action is taken, the normalized revenue distribution could become a middle and high income revenue distribution only. However, this could lead to a major shift in the dynamics of revenue as the middle and high income taxpayers are more sensitive to tax parameters changes.

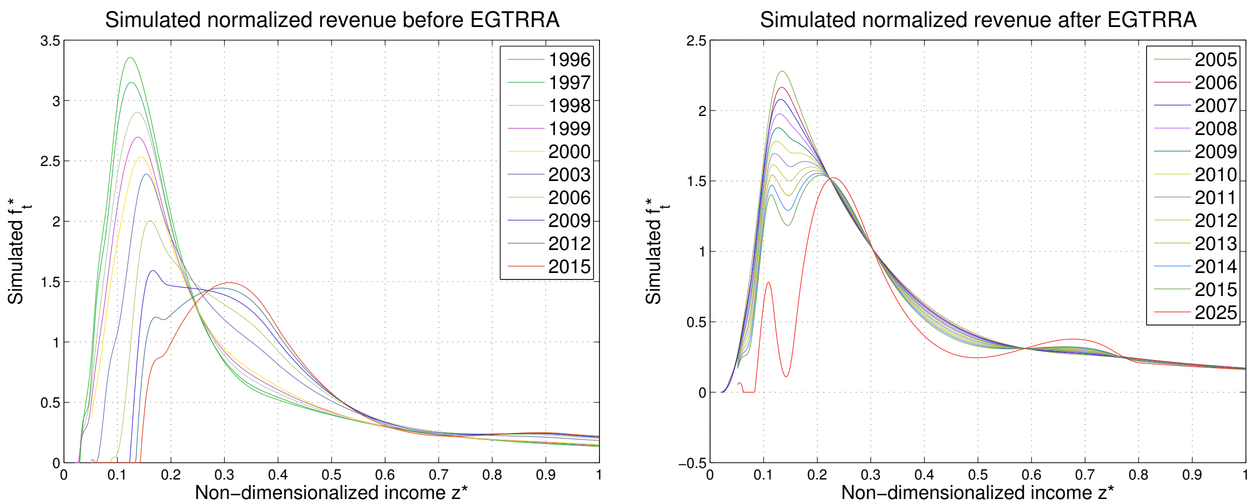



Figure 4.1: *Simulated revenue distribution assuming no changes to current policy, before and after EGTRRA.*

Figure 4.1 shows a simulated scenario where we progressively transform the distribution to a new one formed by middle and high income taxpayers only. We use a simple linear regression at each income $z$ based on a number of years of density distribution values. On the left, we use the years 1996 to 1999 to project future revenue distribution while carrying the same policy; that is indexing the brackets by the inflation rate and keeping the marginal rate the same as before EGTRRA, other things being equal. Similarly, on the right we use the years 2004 to 2007 to project future revenue distribution using the EGTRRA tax parameters changes. Both simulations show that the revenue distribution migrates to all-middle and high income taxpayers only. Changes brought by EGTRRA delayed and changed the shape of the distribution but the end result is the same: there is a range of taxpayers whose relative contribution is very little and the mode of the distribution is shifted toward the right in the high income taxpayers range.

Naturally, linear regressions are not very reliable for projections on a such a long time horizon and there are better sophisticated data mining and system identification techniques available. Moreover, it is possible that taxpayers behavioral responses change and automatically correct the drop in the peak of revenue. However, the point of the simulation is to show possible scenarios to emphasize the need for control of the yearly changes to the shape of the revenue distribution. Controlling for the instability has its own academic merit but more importantly, adding a rational control on the changes in the revenue distribution has its social and economic benefits such as avoiding a class conflict or social change. The responsibility for the shape of the controlled distribution rests squarely on the shoulders of policy makers.

## 4.3 Year to year control of Instability

We consider the simple case where the government controls the instability in the revenue distribution by containing the consistent drop in its peak. The objective is to find the optimal set of brackets and corresponding marginal tax rates so as to keep the normalized revenue density of the following year to be the same as that of the current year. That is $f_{t+1}(z) = f_t(z)$, for $z \geq 0$.

Let us assume that the distribution of number of returns $g_{t+1}$ as well as the taxable income per capita $I_{t+1}$ of the following year is known. The problem is simply $T_t\, g_t = T_{t+1}\, g_{t+1}$ which yields the optimal tax liability,

$$T_{t+1}(z) = T_t(z)\,\frac{g_t(z)}{g_{t+1}(z)}. \tag{4.1}$$

The optimal tax liability at gross income $z$, to keep the relative revenue contribution at the current year level, is the current tax liability multiplied by the inverse of the growth in the number of returns.

Figure 4.2 shows on the left the tax liability $T_t^*$ from data for 1996, the tax liability control $T_{t+1}^{*c}$

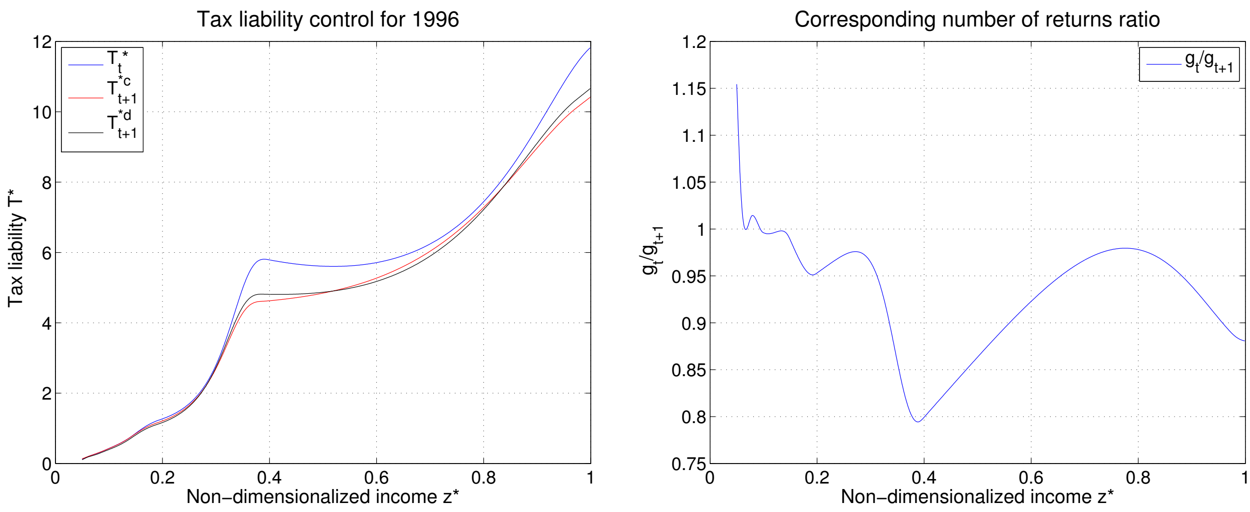



Figure 4.2: *Tax liability and number of returns ratio for 1996.*

from (4.1) and the tax liability $T^{*d}_{t+1}$ from data for 1997. All three curves are close enough to form one curve in the lower income range but as income increases the control and the tax liability data for 1997 become lower than that of 1996. This is supported by the ratio of number of returns displayed on the right. The more it is close to 1, the more the control and the 1996 tax liability are closer to each other. Surprisingly, the control and the 1997 tax liability curve from data are close to each other over all income ranges. This implies that the revenue curve is very sensitive to slight changes in the tax liability as clearly the 1997 curve from data causes the instability whereas the tax liability control (in red) corrects such an instability. The level of sensitivity is highest near the peak of the revenue distribution (around 0.1 in the non-dimensionalized income axis) where there is small difference in the control and the 1997 curve from data but still the peak is dropping year by year. Similar observations are made for the remaining years. The corresponding figures are put in Appendix B.

### 4.3.1 Effective Taxable Income

Equation (4.1) provides the optimal control of the instability observed in $f_t$ so long as the identity $f_t(z) = T_t(z)\, g_t(z) = r_t(z)\, I_t(z)\, g_t(z)$ holds. In principle, this identity is trivial as the relative contribution at $z$ is defined by the corresponding tax liability at $z$ multiplied by the corresponding number of taxpayers at that income. However, (4.2) shows that this is not always the case. In fact, the observed $f_t$ is always slightly less than $T_t(z)\, g_t(z)$. There are several reasons such as numerical errors, unclean data and measurement errors among other things but the most important one is the fact that not every return filed pays taxes and tax deductions are not uniformly distributed among taxpayers.

A standard technique used in the applied world in general is the use of an effective quantity when, the theory not being proved wrong, observations do not match its expected theoretical value. Here, we could restore the equality by either redefining the number of returns or the tax liability. We choose the latter as the former is stable and regular from the data whereas the tax liability is affected by errors in tax deductions identifications as previously argued. Thus, we define the effective taxable income, denoted $I^{eff}$, as the taxable income that would be required to obtain $f_t$ given the number of returns $g_t$. In other words, $I^{eff}$ is such that, $f_t(z) = r_t(I^{eff}(z))\, I_t^{eff}(z)\, g_t(z)$. It follows that,

$$I^{eff}(z) = b_i + \frac{1}{r_i}\,[\frac{f(z)}{g(z)} - \sum_{k=1}^{i} r_{k-1}(b_k - b_{k-1})], \tag{4.2}$$

for $b_i \le I^{eff}(z) \le b_{i+1}$.

The corresponding effective tax deductions and liability are respectively given by,

$$\begin{aligned} d^{eff}(z) &= (z - I^{eff}(z))\; g(z) \\ T^{eff}(z) &= r(I^{eff}(z))\; I^{eff}(z). \end{aligned} \tag{4.3}$$

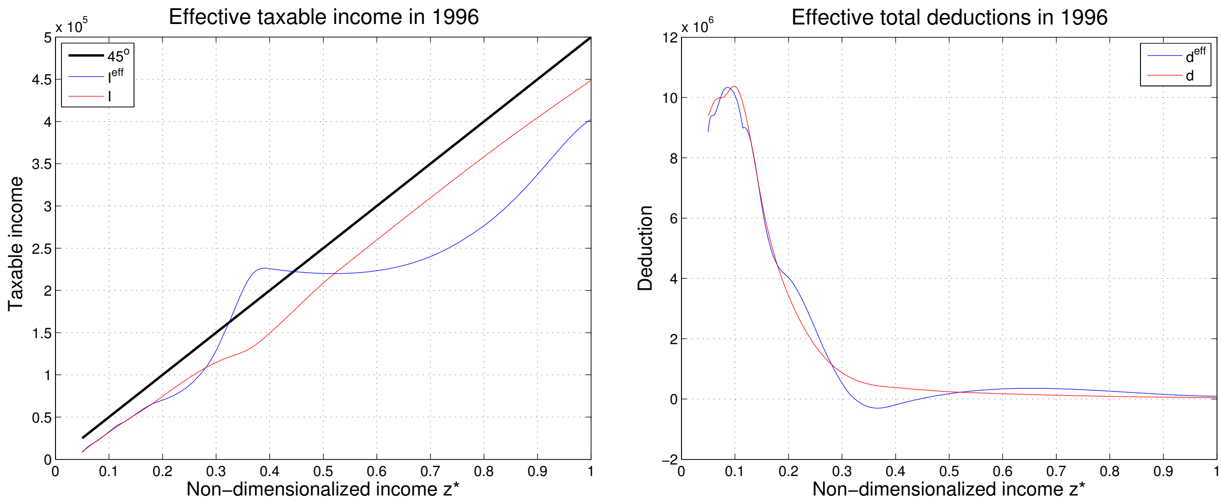



Figure 4.3: *Effective taxable income and deductions for 1996.*

Figure 4.3 shows the effective taxable income and deductions for 1996. On the left, the taxable income $I(z)$ as developed in the previous chapter is compared to the effective taxable income $I^{eff}(z)$. Clearly, the effective taxable income is not monotonically increasing and exceeds the 45 degree line for a range of income. As discussed above, the effective quantity is only used as there are observations unexplained by theory. On the right, the corresponding theoretical deductions $d(z)$, as developed in the previous chapter, are compared to the effective deductions $d^{eff}$. They are similar in shape overall but do not match for most income ranges. Thus, difference in deduction estimation contributes to the difference between the relative tax revenue effective taxable income anand the tax formula. Similar conclusions hold and the corresponding figures are reported in Figure B.1 in Appendix B.

For ease of notation, we will not use the upper script $^{eff}$ in the reminder of this chapter but $I_t$ and $T_t$ should be understood as their effective values.

### 4.3.2 Revenue Constraints

The revenue constraints in the original Mirrlees (1971) work states that total consumption, that is the overall individual consumption plus the government expenditure, must match the total production in the overall economy. In our case, as we are not looking at individual behaviour but rather the aggregate group behaviour directly through $g$ and $f$, it suffices to require that the total tax revenue matches the government expenditures. In other words, we control for the instability in the relative revenue contribution while requiring that total revenue is as desired, $\int_0^\infty f_{t+1}(z)\,dz = R_*$, for some $R_*$.

The tax liability $T_{t+1}$ as given by (4.1) ensures that $f^*_{t+1} = f^*_t$ but the required revenue $R_*$ is general and must not necessarily be equal to $R_t$, the total revenue at $t$. As we control for the instability on the normalized revenue, we have $f^*_{t+1} = f^*_t$. Hence, $\int_0^\infty f^*_{t+1}(z^*)\,dz^* = \int_0^\infty f^*_t(z^*)\,dz^* = 1$. However, this does not conflict with the revenue requirement that $\int_0^\infty f_{t+1}(z)\,dz = R_* \neq R_t = \int_0^\infty f_t(z)\,dz$. Using the correct nondimensionalization factor for each variable and function converts them into nominal terms which in turn insures the revenue constraints. We have,

$$\begin{aligned} f_{t+1}(z) &= \frac{R_*}{z_0^{t+1}} f^*_{t+1}(z^*) = \frac{R_*}{z_0^{t+1}} T^*_{t+1}(z^*)\, g^*_{t+1}(z^*) = \frac{R_*}{z_0^{t+1}} r^*_{t+1}(z^*)\, I^*_{t+1}(z^*)\, g^*_{t+1}(z^*), \\ &= \frac{R_*}{z_0^{t+1}} r^*_{t+1}(z^*)\, \frac{I_{t+1}(z)}{z_0^{t+1}}\, \frac{z_0^{t+1}}{N_{t+1}}\, g_{t+1}(z) = \frac{R_*}{z_0^{t+1} N_{t+1}} r^*_{t+1}(z^*)\, I_{t+1}\, g_{t+1}(z). \end{aligned} \tag{4.4}$$

As expected, $r_{t+1}(z) = \frac{R_*}{z_0^{t+1} N_{t+1}} r^*_{t+1}(z^*)$. In other words, the nominal average rate is directly proportional to the desired revenue $R_*$.

From (4.1),

$$r^*_{t+1}(z^*) = \frac{T^*_{t+1}(z^*)}{I^*_{t+1}(z^*)} = \frac{T^*_t(z^*)}{I^*_{t+1}(z^*)}\,\frac{g^*_t(z^*)}{g^*_{t+1}(z^*)} = r^*_t(z^*)\,\frac{I^*_t(z^*)}{I^*_{t+1}(z^*)}\,\frac{g^*_t(z^*)}{g^*_{t+1}(z^*)}. \tag{4.5}$$

Therefore,

$$\begin{aligned} r_{t+1}(z) &= \frac{R_*}{z_0^{t+1} N_{t+1}} r^*_{t+1}(z^*) = \frac{R_*}{z_0^{t+1} N_{t+1}} r^*_t(z^*) \frac{I^*_t(z^*)}{I^*_{t+1}(z^*)} \frac{g^*_t(z^*)}{g^*_{t+1}(z^*)}, \\ &= \frac{R_*}{z_0^{t+1} N_{t+1}} \frac{z_0^t N_t}{R_t} \frac{I^*_t(z^*)}{I^*_{t+1}(z^*)} \frac{g^*_t(z^*)}{g^*_{t+1}(z^*)} r_t(z), \\ &= \frac{z_0^t}{z_0^{t+1}} \frac{R_*}{R_t} \frac{I_t(z)}{I_{t+1}(z)} \frac{g_t(z)}{g_{t+1}(z)} r_t(z). \end{aligned} \tag{4.6}$$

Thus, the average rate that stabilizes $f^*$ is directly proportional to the ratio of the required revenue and the current year's total revenue. It is inversely proportional to the inflation rate and the growth rate of the taxable income and number of returns. Intuitively, the higher the revenue required compared to the current collection, the higher the average tax rate should be, other things being equal. The higher the inflation rate or the higher the taxable income or the higher the number of returns increased, the lower the average rate should be.

Similar derivation can be carried on for the effective tax liability. For space convenience, we only state the result. We have,

$$T_{t+1}(z) = \frac{R_*}{N_{t+1}} T^*_{t+1}(z^*) = \frac{z_0^t}{z_0^{t+1}} \frac{R_*}{R_t} \frac{g_t(z)}{g_{t+1}(z)} T_t(z). \tag{4.7}$$

### 4.3.3 Year to year optimal brackets

The tax liability formula as given by (4.1), (4.7) or that of the average tax rate given in (4.6) is not practicable as those are not physical control variables that the government can manipulate. Change in revenue has to be through changes in one or more tax parameters such as tax deductions, tax credits, tax brackets and marginal tax rates. In the case of tax brackets, there are only a number of studies of optimal tax brackets in the literature. To the best of our knowledge, there is no theoretical background to the setting of tax brackets except from inflation indexation to avoid bracket creep. Hence, the natural question one ought to ask is if this current practice is optimal in some sense? if not, how should the tax brackets be set, how many and how far apart?

When controlling the instability in the normalized revenue distribution, optimality means finding the set of tax brackets with the associated tax liability as close as possible to that given by (4.1). The problem is defined as follows,

$$\min_b \|T_{t+1} - T(b)\|, \tag{4.8}$$

where $b = (b_0, \ldots, b_5)$ and $T_{t+1}$ given by (4.1). $T(b)$ is the tax liability associated with the set of tax brackets $(b)$ and defined by,

$$\begin{aligned} T(z,b) &= (I(z) - b_i)\, r_i + \sum_{k=1}^{i} r_{k-1}\,(b_k - b_{k-1}), \\ &= I(z)\, r_i - \sum_{k=1}^{i} b_k\,(r_k - r_{k-1}) - r_0\, b_0, \end{aligned} \tag{4.9}$$

when $b_i \leq I(z) \leq b_{i+1}$. The optimal brackets minimize the distance between those two functions. The norm used here can be any norm of functions. Numerical tests show that the norm on $L^2(f_t)$ gives the best result when the problem is put in the following form,

$$\min_b \int_0^\infty f_t(z)\{T_t(z)\, g_t(z) - T(z,b)\, g_{t+1}(z)\}^2\, dz. \tag{4.10}$$

Here, we minimize the $L^2$ distance between $f_t$ and $f_{t+1}$ while using the weight $f_t$. This is to emphasize the importance of controlling the instability which is at its highest level at the peak of $f_t$.

The problem (4.10) is a classical convex quadratic minimization problem that admits at least one minimum, Chong and Zak (2001). The first order optimality condition is given by,

$$\int_0^\infty -2\, \frac{\partial T}{\partial b_i}(z,b)\, g_{t+1}(z)\, f_t(z)\{T_t(z)\, g_t(z) - T(z,b)\, g_{t+1}(z)\}\, dz = 0 \quad i = 1, \ldots, 5. \tag{4.11}$$

For $b_i \leq I(z) \leq b_{i+1}$, we have,

$$\frac{\partial T}{\partial b_j}(z,b) = \begin{cases} -(r_j - r_{j-1}) & \text{if } j \leq i, \\ 0 \quad \text{if } j > i. \end{cases} \tag{4.12}$$

It follows that,

$$\int_{z_{b_i}}^\infty g_{t+1}(z)\, f_t(z)\{T_t(z)\, g_t(z) - T(z,b)\, g_{t+1}(z)\}\, dz = 0, \tag{4.13}$$

for $i = 1, \ldots, 5$ and where $z_{b_i}$ is the income level corresponding to the bracket threshold $b_i$, that is $I(z_{b_i}) = b_i$. Note the number of brackets is not a control variable.

In (4.13), the marginal tax rates are given by the previous year rates and $b_i$ are the unknown minimizers. As $z_{b_i}$ depends on $b_i$, (4.13) yields a system of nonlinear equations that is solved numerically. The results for year 1996 are given in Figure 4.4. The optimal brackets $b^{*b}_{t+1}$ and nondimensionalized bracket from data, denoted $b^{*d}_{t+1}$ for 1997 and $b^*_t$ for 1996, are displayed on the left. On the right, the corresponding revenue $f^{*b}_{t+1}$ using the optimal brackets $b^{*b}_{t+1}$, the revenue density from data in 1996 $f^*_t$ and in 1997 $f^{*d}_{t+1}$ are given. As expected, the range of the

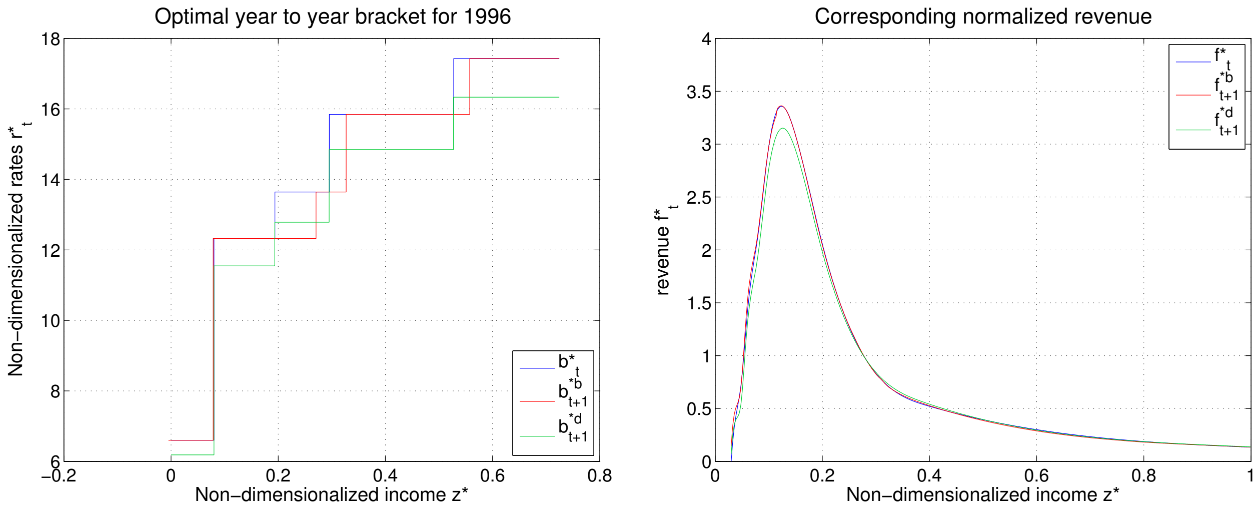



Figure 4.4: *Year to year optimal brackets and the resulting normalized revenue distribution for 1996.*

optimal bracket around the peak is widened compared to the current 1996 bracket. The more taxpayers in the bracket the higher the revenue contribution, the less discrepancy around the peak and thus the instability is contained. Higher brackets are shifted accordingly.

Figure 4.5 shows on the left the time series of optimal tax brackets that control the following year curve to the current one, from 1996 to 2007. The dashed lines are the brackets from data and the solid lines are the optimal brackets. This confirms the general trend that the first bracket is not changed. The second bracket threshold that contains the peak of the revenue distribution is increased. The remaining higher brackets are shifted more or less by the same amount. The top bracket threshold is shifted but it tends to be constant over time except for the years during EGTRRA where larger shifts in the second bracket were needed to contain the peak. The ratio of the controlled revenue to the revenue from data for 1996 is displayed on the right in Figure 4.5. It shows the relative change from year to year in the revenue in 1966. Clearly, the control is improved around the peak of the revenue distribution. Above 0.5 on the nondimensionalized income axis, the control is weaker. Intuitively, the top bracket alone is not strong enough to control for the revenue in this range. However, this is located on the wide bins of the revenue data where the cubic spline estimation is not reliable and one cannot draw a tangible conclusion. Moreover, the number of returns level on this range is very small, in the order of $10^{-6}$, which make the ratio numerically sensitive. In other words, there is a trade-off between controlling for the instability at the peak of $f_t^*(z^*)$ and at the tail.

In this section, we used the year 1996 data as an illustration but similar results hold for the control of instability using tax brackets for remaining years. The corresponding figures are given in Figure B.3 in Appendix B.

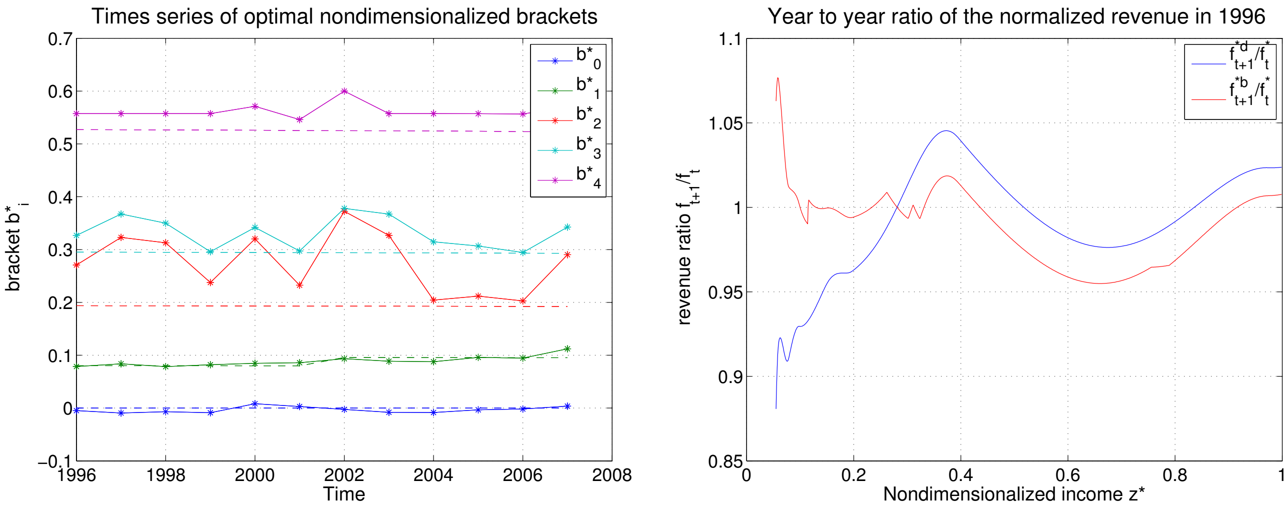



Figure 4.5: *Optimal inflation adjusted brackets time series from 1996 to 2007 and year to year ratio of the revenue for 1996 to 1997.*

We note that the marginal rates here are constant from 1996 to 1997 because the required revenue in this case is the total revenue obtained in 1997 from the data. That is, $R_* = \int_0^\infty f^c_{t+1}(z)\,dz = R_0^{t+1} = \int_0^\infty f^d_{t+1}(z)\,dz$. For the case where $R_* \neq R_0^{t+1}$, the marginal rates change as discussed in section 4.3.2 above. Optimal brackets only implies constant nondimensionalized marginal tax rates but may involve changes in nominal marginal tax rates as well. Furthermore, as the control of instability implies there are small errors between the controlled revenue and the data, $f^c_(z) \simeq f^*_t(z)$ then $\int_0^\infty f^c(z)\,dz \neq R_*$ but this can be corrected by requiring a slightly higher revenue $R'_* = R_* + \epsilon$ where $\epsilon$ is larger than the area between the controlled revenue and the data.

### 4.3.4 Year to year optimal marginal tax rates

To control the instability, the optimal set of marginal tax rates is obtained through the same objective function. The problem is set as follows,

$$\min_r \int_0^\infty f_t(z)\{T_t(z)\,g_t(z) - T(z,r)\,g_{t+1}(z)\}^2\,dz. \tag{4.14}$$

The tax liability, $T(z,r)$ is the same as defined in (4.9) where the brackets are fixed as given by the inflation indexation and the marginal rates are the unknown minimizers.

In this case, the first order optimality condition is given by,

$$\int_0^\infty -2\,\frac{\partial T}{\partial r_i}(z,b)\,g_{t+1}(z)\,f_t(z)\{T_t(z)\,g_t(z) - T(z,b)\,g_{t+1}(z)\}\,dz = 0 \quad i = 1,\dots,5. \tag{4.15}$$

Moreover, for $b_i \le I(z) \le b_{i+1}$,

$$\frac{\partial T}{\partial r_j}(z,b) = \begin{cases} (b_j - b_{j-1}) & \text{if } j < i, \\ (I(z) - b_j) & \text{if } j = i, \\ 0 \quad \text{if } j > i. \end{cases} \tag{4.16}$$

It follows that,

$$\int_{z_{b_i}}^{z_{b_{i+1}}} (I(z) - b_i)\, H(z,b)\, dz + \int_{z_{b_{i+1}}}^{\infty} (b_i - b_{i-1})\, H(z,b)\, dz = 0, \tag{4.17}$$

for $i = 1, \ldots, 5$ and where $H(z,b) = g_{t+1}(z)\, f_t(z)\{T_t(z)\, g_t(z) - T(z,b)\, g_{t+1}(z)\}$ and $z_{b_i}$ as before.

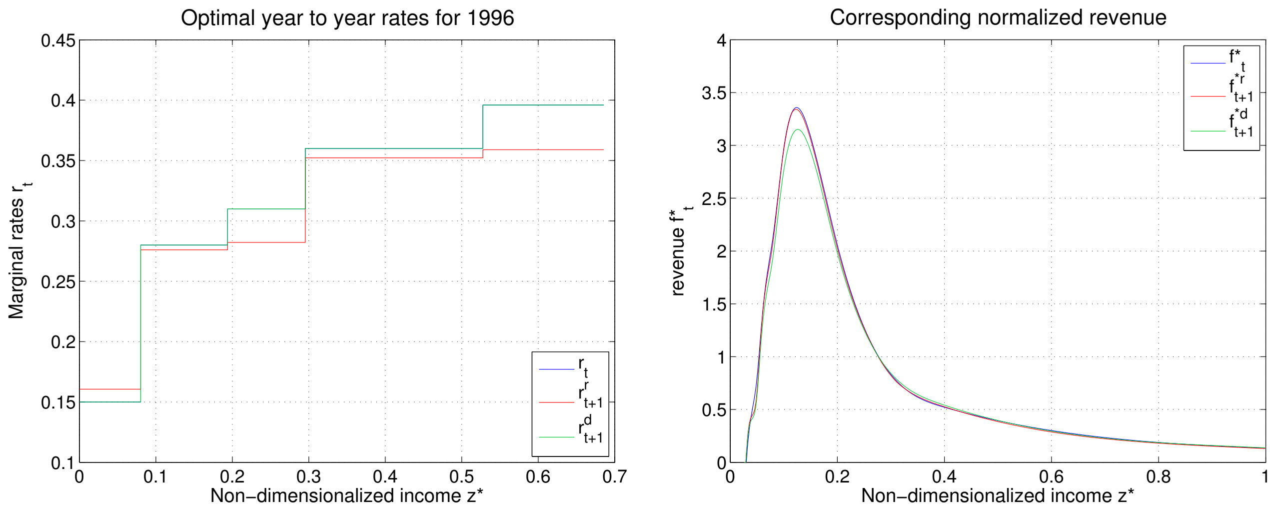



Figure 4.6: *Year to year optimal marginal tax rates and the resulting normalized revenue distribution for 1996.*

As $b_i$ is given by the inflation indexation of the previous year brackets, (4.17) is a linear system that can be easily solved to obtain the optimal marginal rates. Figure 4.6 displays the results for the year 1996. Notations are the same as in the previous section, a superscript $^r$ indicates a control variable/function whereas a superscript $^d$ indicates a variable estimated from data. The optimal marginal tax rates are displayed on the left. We observe that the optimal nominal marginal rates $r^r_{t+1}$ are decreased from their 1996 level labeled $r_t$ except for the bottom rate which increased. This is to be expected as shown in the previous chapter that marginal rates reductions brought by EGTRRA temporarily restored the instability. Decreased marginal rates implies increased relative revenue contribution at lower income ranges that restores the peak.

On the other hand, increased marginal rates on the top bracket will result in a decreased revenue contribution in higher income ranges as the average rate in this region is practically the

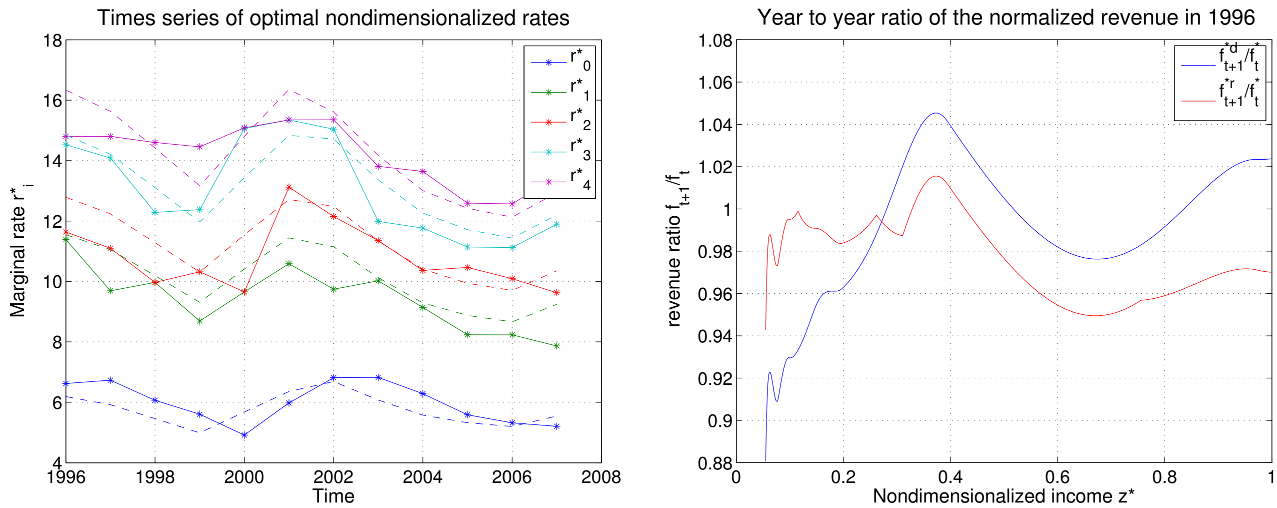



Figure 4.7: *Optimal marginal rates time series from 1996 to 2007 and year to year ratio of the revenue for 1996.*

same as the marginal rate. We note that the instability control here is not good comparable to the case of optimal brackets only, discussed in the previous section. The controlled revenue $f^{*r}_{t+1}$ coincides with the 1996 curve $f^*_t$ for most of the lower income ranges.

We plot on the left of Figure 4.7 the time series of optimal marginal rates to control the following year revenue to the current year one. The same notation as for the optimal brackets is used. That is, dashed lines indicate actual data whereas solid lines are used for the optimal values. It confirms the general trend whereby marginal rates are lowered to control the instability except for the lowest and top marginal rates which follow the opposite trend. The year to year revenue ratio is plotted on the right. Again, we achieve control below the nondimensionalized income 0.5. Again, above this value, there is no tangible conclusion to make as the revenue curve themselves are not properly estimated. Surprisingly, this figure also shows that there is no apparent superiority of the control of tax rates over tax brackets. We will discuss this further at a later section.

As before, the year 1996 is used here as an illustration but similar results hold for the remaining years and are reported in Figure B.4 given in Appendix B.

### 4.3.5 Combined optimal brackets and marginal rates

This is the case where both the tax brackets and marginal tax rates are to be optimally determined. The problem is defined as follows,

$$\min_{r,b} \int_0^\infty f_t(z)\{T_t(z)\, g_t(z) - T(z,r,b)\, g_{t+1}(z)\}^2 \, dz. \tag{4.18}$$

$T(z,r,b)$ is the same as defined in (4.9) with both brackets and marginal rates being the unknown minimizers. The nonlinear equations given by (4.13) and (4.17) are both simultaneously solved to give the optimal set of brackets and corresponding marginal tax rates.

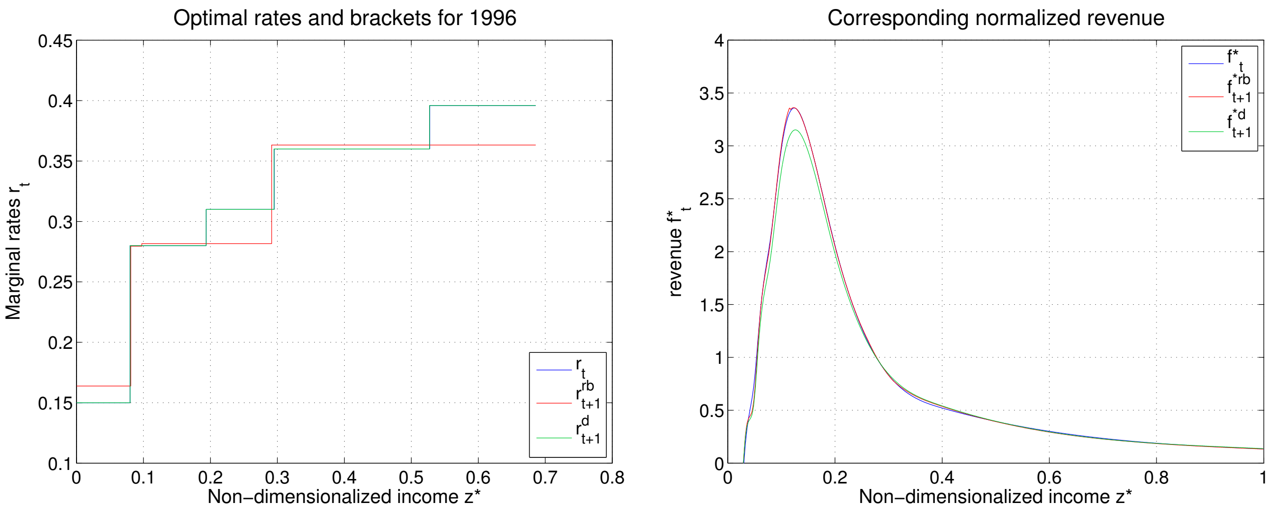



Figure 4.8: *Year to year optimal marginal tax rates and brackets together with the resulting normalized revenue distribution for 1996.*

Figure 4.8 displays the results for year 1996. The same notation is used, a superscript $^{r}b$ indicates a control variable whereas a superscript $^{d}$ indicates a variable estimated from data. The optimal brackets and marginal tax rates are displayed on the left. We observe a combination of the theories discussed in the two previous sections. The widening of brackets and reductions of rates become a complete removal of some of the brackets. As a result, some taxpayers would have their marginal rates decreased, others are a slightly higher than their original rates.

The revenue curves are displayed on the right. We achieve a control of the instability, with fewer brackets. The controlled revenue $f^{*rb}_{t+1}$ is superposed with the 1996 revenue curve $f^*_t$ at all income $z$. We plot in Figure 4.9 right upper corner the ratio of the controlled revenue and revenue from data to the 1996 revenue curve. It shows that near perfect control is achieved near the peak. The ratio is close to one near the peak but it deviates more and more as we move along the income axis. Again, this range $(0.5 \le z^* \le 1)$ falls in the wide income bin of revenue data where our cubic spline estimation is not reliable so that no tangible conclusion can be drawn. The remaining figures in Figure 4.9 shows the time series of optimal brackets (top left) and marginal tax rates (bottom row) to control the following year to the current one from 1996 to 2007. The top bracket sometimes disappear outside the range, they are omitted as they correspond to brackets that are optimally merged. Meaning, the rates are the same below and above them. The same for marginal rates are observed, they are reduced compared to the data (in dashed lines) to control for the instability except for the years where the EGTRRA rate reductions were phased in. For those years, the reverse trend is observed as expected. Similar results hold for the control of instability for the remaining years. They are reported in Figure B.5 in Appendix B.

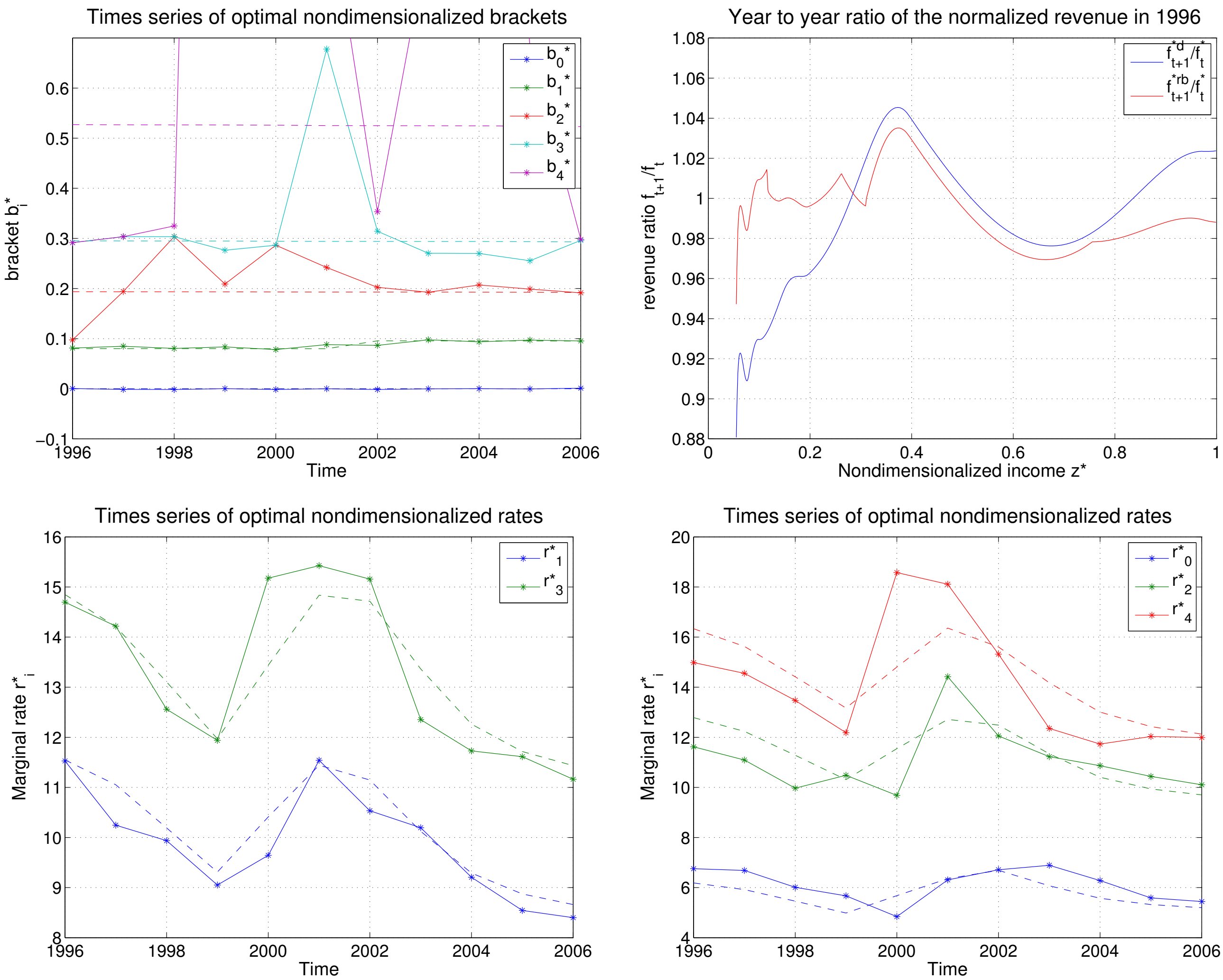



Figure 4.9: *Optimal marginal tax rates and brackets time series from 1996 to 2007 and year to year ratio of the revenue for 1996 to 1997.*

Controlling with both tax brackets and marginal tax rates is the best control to keep the shape of the relative revenue contribution constant. Figure 4.10 shows on the right the different year to year ratios using the three controls discussed so far for 1996. The notations are the same, a superscript $^b$ indicates a control variable using the brackets only, $^r$ indicates the control using the rates only and $^{rb}$ using both the rates and marginal rates. Clearly, the control using marginal rates only achieves the worst control. Using brackets only or both rates and brackets get similar control performance around the peak but using the combination is best around the tail. This is expected since it generalizes both tax brackets only and marginal tax rates only. However, its major disadvantage is that it increases the dimension of the problem considerably and this requires a very good initial iteration of the optimization for it to succeed.

On the left of Figure 4.10, we compare the different optimal rates and brackets using the three controls discussed for 1996. Using the brackets only as control yields higher nominal marginal tax rates. The required revenue, that is the total revenue of 1997, increased the nominal rates compared to their 1996 nominal level. Surprisingly, the optimal rates only and using both rates and brackets are very similar and differ only by a small amount at the top and bottom brackets. This confirms key results in the litterature that taxpayers at the top brackets are sensitive to marginal tax rates changes in the short term, Goolsbee (2000).

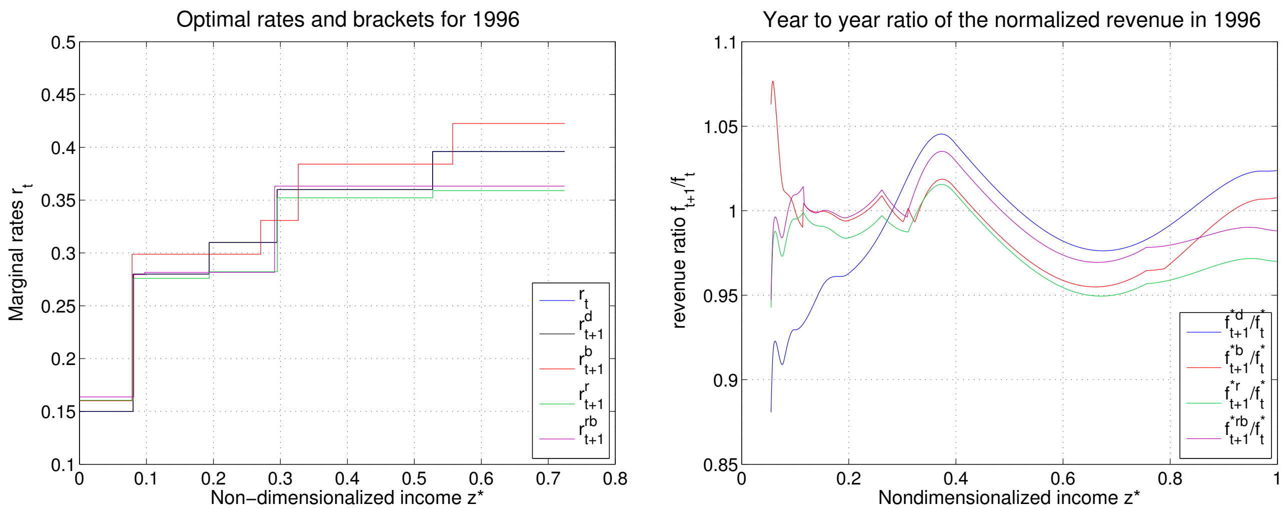



Figure 4.10: *Comparison of optimal marginal tax rates and brackets as well as year to year ratio of the revenue, for 1996.*

## 4.4 Dynamics on g and Implied Control

The analysis in the previous section relies on the assumption that we are provided with $g_{t+1}$, the distribution of number of returns for next year. This assumes that the approximation procedure to obtain $g_{t+1}$ does not involve and is independent of future tax liability (or equivalently tax brackets and/or marginal rates). In this section, we investigate the effect on the optimal individual tax payment when $g_{t+1}$ does depend on changes in brackets and/or marginal tax rates.

Intuitively, individuals respond to changes in brackets and/or marginal tax rates variably according to their respective elasticities. However, the main philosophy of the thesis is to investigate behavior at the aggregate group level. As seen in chapter 3, there are undoubtedly taxpayers responses to changes in tax parameters but they are not the main effect that drive changes in the revenue distribution and can perhaps be ignored. Hence, our goal is to quantify the extent to which changes in number of returns through changes in tax parameters affect the revenue and changes the optimal tax parameters in turn. As before, optimality here means tax parameters that control the instability in the revenue distribution.

Total differentiation on $g$ at fixed $z$ yields,

$$dg = \frac{\partial g}{\partial T} dT + \frac{\partial g}{\partial t} dt. \tag{4.19}$$

Here we decomposed changes in $g$ at fixed $z$ into tax liability effect and time effect. The tax liability effect is the response on $g$ when the tax liability per individual is changed whereas the time effect is the change in $g$ when the tax liability is not changed. That is the effet of everything else than $T$.

In discrete terms, we have,

$$\Delta g = \frac{\partial g}{\partial T} \Delta T + \frac{\partial g}{\partial t} \Delta t. \tag{4.20}$$

That is,

$$g_{t+1} = g_t + \frac{\partial g}{\partial T} \Delta T + \frac{\partial g}{\partial t}. \tag{4.21}$$

(4.21) is obviously an oversimplified model of $g_{t+1}$ as a result of a simple linearization of the true process. There are many other variables affecting the distribution of number returns but they are omitted in the equation. For control purposes, only the variables of interest and only those that have the strongest effect are of importance. Those are variables that can be physically altered to put such a strong control that it annihilates other smaller adverse effects on the dynamics.

For ease of notation, let us denote $\tilde{g}_{t+1} = g_t + \frac{\partial g}{\partial t}$ so that $g_{t+1} = \tilde{g}_{t+1} + \frac{\partial g}{\partial T} \Delta T$. Now, the control of instability $T_{t+1}\, g_{t+1} = T_t\, g_t$ implies,

$$T_{t+1}\,[\tilde{g}_{t+1} + \frac{\partial g}{\partial T} \Delta T] = T_t\, g_t. \tag{4.22}$$

That is,

$$(T_t + \Delta T)\,[\tilde{g}_{t+1} + \frac{\partial g}{\partial T} \Delta T] = T_t\, g_t. \tag{4.23}$$

Hence,

$$\frac{\partial g}{\partial T} \Delta T^2 + [\tilde{g}_{t+1} + \frac{\partial g}{\partial T} T_t]\, \Delta T - T_t\,(g_t - \tilde{g}_{t+1}) = 0. \tag{4.24}$$

This is a quadratic equation in $\Delta T$ whose solutions are given by,

$$\begin{cases} \Delta T = \dfrac{-1}{2\,\frac{\partial g}{\partial T}} \left\{ [\tilde{g}_{t+1} + \dfrac{\partial g}{\partial T} T_t] + \sqrt{[\tilde{g}_{t+1} + \dfrac{\partial g}{\partial T} T_t]^2 + 4\, T_t\,(g_t - \tilde{g}_{t+1})\, \dfrac{\partial g}{\partial T}} \right\}, \\ \Delta T = \dfrac{-1}{2\,\frac{\partial g}{\partial T}} \left\{ [\tilde{g}_{t+1} + \dfrac{\partial g}{\partial T} T_t] - \sqrt{[\tilde{g}_{t+1} + \dfrac{\partial g}{\partial T} T_t]^2 + 4\, T_t\,(g_t - \tilde{g}_{t+1})\, \dfrac{\partial g}{\partial T}} \right\} \end{cases} \tag{4.25}$$

The first solution blows up to infinity when $\frac{\partial g}{\partial T} \longrightarrow 0$ and is discarded. The change required in the individual tax liability to control the instability is therefore given by,

$$\Delta T = \frac{-1}{2\,\frac{\partial g}{\partial T}} \left\{ [\tilde{g}_{t+1} + \frac{\partial g}{\partial T} T_t] - \sqrt{[\tilde{g}_{t+1} + \frac{\partial g}{\partial T} T_t]^2 + 4\, T_t\,(g_t - \tilde{g}_{t+1})\, \frac{\partial g}{\partial T}} \right\} \tag{4.26}$$

The corresponding number of returns is given by,

$$\begin{aligned} g_{t+1} &= \frac{1}{2} \left\{ [\tilde{g}_{t+1} - \frac{\partial g}{\partial T} T_t] + \sqrt{[\tilde{g}_{t+1} + \frac{\partial g}{\partial T} T_t]^2 + 4\, T_t\,(g_t - \tilde{g}_{t+1})\, \frac{\partial g}{\partial T}} \right\} \\ &= \frac{1}{2}\,[\tilde{g}_{t+1} - \frac{\partial g}{\partial T} T_t] + \frac{1}{2} \sqrt{[\tilde{g}_{t+1} + \frac{\partial g}{\partial T} T_t]^2 - 4\, T_t\, \frac{\partial g}{\partial t}\, \frac{\partial g}{\partial T}} \end{aligned} \tag{4.27}$$

The individual tax liability that controls for the instability is $T_{t+1} = T_t\, \frac{g_t}{g_{t+1}}$. It is checked in the appendix that $T_{t+1} = T_t + \Delta T$ using (4.26) and $T_{t+1} = T_t\, \frac{g_t}{g_{t+1}}$ is the same. One can conclude here that including changes in tax parameters effects in changes in the number of returns does not alter the tax liability formula to control for the instability. It finds the appropriate changes in tax liability to counter or compensate changes in the number of returns that are independent of tax parameters as measured by $\frac{\partial g}{\partial t}$.

The second equality in (4.27) clearly shows that the number of returns that contain the instability in the revenue distribution is an average correction in the time changes in $g$ given by $\tilde{g}_{t+1}$ by the tax parameters effect $\frac{\partial g}{\partial T} T_t$. In some sense, the average correction is the right combination of how much the tax parameter effects should contribute to changes in $g$ to obtain no changes in the normalized revenue distribution. Clearly, as the tax parameters effect vanishes, $\frac{\partial g}{\partial T} \longrightarrow 0$, the optimal number of returns converges to its natural time evolution $g_{t+1} \longrightarrow \tilde{g}_{t+1}$.

It is proven in Appendix B.3.1 that $\Delta T \longrightarrow T_t \frac{(g_t - \tilde{g}_{t+1})}{\tilde{g}_{t+1}}$ as $\frac{\partial g}{\partial T} \longrightarrow 0$. It follows that, $T_{t+1} = T_t + \Delta T = T_t \frac{g_t}{\tilde{g}_{t+1}}$ when $\frac{\partial g}{\partial T} \longrightarrow 0$ as expected. When $g$ is unresponsive to changes in the tax parameters, the individual tax liability that contains the instability in the revenue distribution is the current tax liability multiplied by the inverse of the natural growth (independent of changes in tax parameters) of the distribution of number of returns.

## 4.4.1 Incidence on $\mathbf{I^{eff}}$

The taxable income of a taxpayer given its gross income is dependent on its deductions (and credits). However, the effective taxable income depends on changes in tax parameters as it depends on $f$ and $g$. We shall derive the resulting model of $I^{eff}$ as it is required in the process of deriving the corresponding optimal brackets and/or marginal rates as seen in the previous section.

For $b_i \le I^{eff}(z) \le b_{i+1}$, we have,

$$\begin{aligned} I^{eff}(z) &= b_i + \frac{1}{r_i}\,[\frac{f(z)}{g(z)} - \sum_{k=1}^{i} r_{k-1}(b_k - b_{k-1}),], \\ &= b_i + \frac{1}{r_i}\,[T_{t+1} - \sum_{k=1}^{i} r_{k-1}(b_k - b_{k-1}),], \end{aligned} \tag{4.28}$$

where $T_{t+1} = T_t + \Delta T$ with $\Delta T$ as defined in (4.26). Clearly, $I^{eff}$ depends on $T_{t+1}$. Although $I^{eff}$ is not involved in the derivation of optimal tax liability, (4.28) has to be used in numerical implementations to solve for the optimal brackets and rates.

## 4.4.2 Optimality Conditions

With $g_{t+1} = \tilde{g}_{t+1} + \frac{\partial g}{\partial T} \Delta T$, we derive the new optimality conditions of the control of instability problem given by (4.10), (4.14) and (4.18).

For the optimal brackets only, defined by (4.10), the optimality conditions become:

$$\int_{z_{b_i}}^{\infty} f_t(z)\,(g_{t+1}(z)+T(z,b)\frac{\partial g}{\partial T}(z))\,\{T_t(z)\,g_t(z)-T(z,b)\,g_{t+1}(z)\}\,dz=0, \tag{4.29}$$

for $i=1,\ldots,5$ and where $z_{b_i}$ is the income level corresponding to the bracket threshold $b_i$, that is $I(z_{b_i})=b_i$.

For the optimal marginal rates only problem (4.14), the optimality conditions are as before given by (4.17),

$$\int_{z_{b_i}}^{z_{b_{i+1}}} (I(z)-b_i)\,H(z,b)\,dz+\int_{z_{b_{i+1}}}^{\infty} (b_i-b_{i-1})\,H(z,b)\,dz=0, \tag{4.30}$$

for $i=1,\ldots,5$ and where

$$H(z,b)=f_t(z)\,(g_{t+1}(z)+T(z,b)\frac{\partial g}{\partial T}(z))\,\{T_t(z)\,g_t(z)-T(z,b)\,g_{t+1}(z)\}$$

and $z_{b_i}$ as before.

The combined optimal brackets and marginal rates optimality conditions are given by (4.29) and (4.30) as simultaneous equations.

Clearly, the optimality conditions change by adding a linear factor $T(z,b)\,\frac{\partial g}{\partial T}(z)$ from the previous optimality conditions. This is a correction of the effect of $g_{t+1}$ on the objective function as a number of returns responses to changes in the tax parameters. Both (4.29) and (4.30) are now nonlinear equations solved numerically.

### 4.4.3 Estimation of $\frac{\partial g}{\partial T}$ and $\frac{\partial g}{\partial t}$

Historically, the elasticity of taxable income has been used to estimate taxpayers responsiveness to tax rate changes. There is a wide range of methodologies and estimates for the elasticity of taxable income in the literature. See Gruber and Saez (2002) for a comprehensive review of the results and Giertz (2008) for a non-technical summary of methodologies involved.

In principle, one could derive and extrapolate $\frac{\partial g}{\partial T}$ from those estimates. We argue that this is not advisable for the following reasons. Firstly, the underlying principle in estimating the elasticity of taxable income is that it is a measure of how taxpayers respond to changes in tax rates. Therefore, this is a behavior at the individual level as opposed to $g$ which is the distribution of number of returns that captures the collective group behavior. Not only the response in groups is not always the straightforward addition of individual behaviors but also the distribution of number of returns speaks for what is happening and encompasses all possible responses such as tax evasion and frauds. Feldstein (1995) argues that the elasticity of taxable income achieves the same scope as compared to the elasticity of labour income. The number of returns responses clearly estimate tax avoidances and evasions better.

Secondly, there is no consensus on the estimation of this parameter. Gruber and Saez (2002) emphasize the need for convergence. The early estimates of Lindsey (1987) followed by the work of Feldstein (1995) found large estimate above one. Since then, a number of studies found smaller and smaller estimates. To site a few, see Navratil (1995), Auten and Carroll (1999), Sammartino and Weiner (1997), Goolsbee (2000), Saez (2003) and Kopczuk (2003). The most cited paper on this subject is Gruber and Saez (2002) which found an estimate of 0.4. This suggests a consensus that this is the accepted number but recent studies of Webber (2011) disputes this and gives an estimate above one. Various methods each with their own merits gave various estimates. This renders the parameter unusable as no tangible conclusion can be agreed.

Thirdly, most papers would give one or few estimates for the overall taxpayers population. We assume that responses at each income could be different continuously. Therefore the response in the number of returns should be evaluated at each income level, not just a few. Lastly, the literature particularly focused on the elasticity of the taxable income with respect to tax rates. We estimate the response of the number of returns due to changes in the individual tax liability. Changes could stem from changes in brackets and other tax parameters but not exclusively the marginal tax rates.

We use an ordinary least square estimation method on (4.20) at each income $z$ on the 12 years time series of $\Delta g$ and $\Delta T$. The linear coefficient is the estimate of $\frac{\partial g}{\partial T}$ and the intercept is the estimate of $\frac{\partial g}{\partial t}$.

Figure 4.11 shows the $\frac{\partial g}{\partial T}$ and $\frac{\partial g}{\partial t}$ along the income axis. As expected they are both small except for $\frac{\partial g}{\partial T}$ around the peak. Note that this does not contradict the literature that says that high income earners have higher elasticities of taxable income. Elasticities are scaled value which will be non-zero using $\frac{\partial g}{\partial T}$ as $g$ itself is vanishing at higher incomes.

Naturally, with more time series data and using more sophisticated data mining and filtering techniques, those estimates could be improved. This is beyond the scope of our study because of the data limitations and no particular estimation technique can have any superiority on others. One of the objectives in this section is to show that reasonable control of the instability can be achieved and small errors in $\frac{\partial g}{\partial T}$, $\frac{\partial g}{\partial t}$ do not affect the optimal brackets and/or marginal rates much.

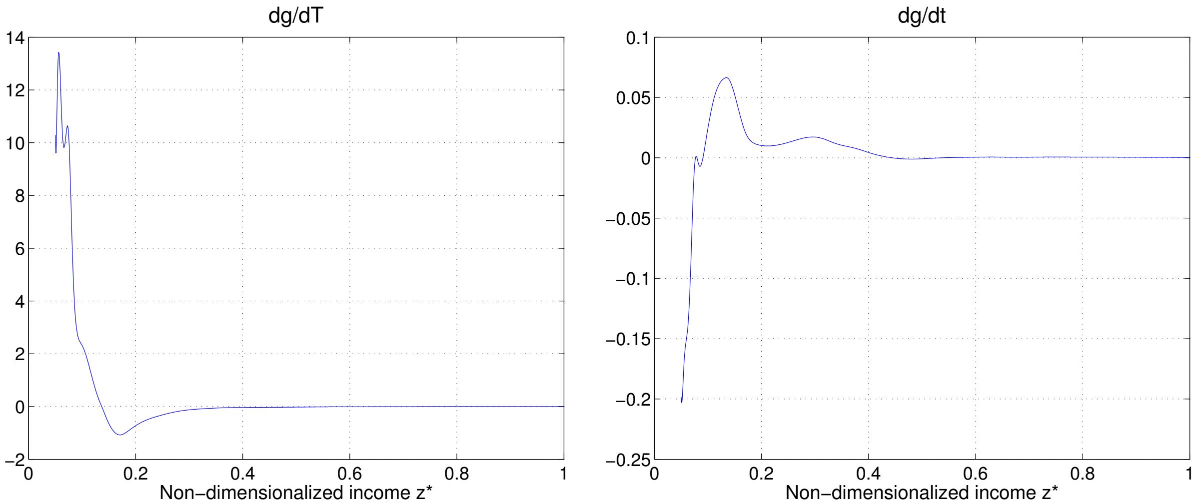



Figure 4.11: *Estimates of* $\frac{\partial g}{\partial T}$ *and* $\frac{\partial g}{\partial t}$ *varying with income.*

### 4.4.4 Results for 1996

Here we compare the optimal tax brackets and/or marginal tax rates with actual and modeled $g_{t+1}$ as well as the one using $\tilde{g}_{t+1}$ where $g$ is assumed independent of changes in $T$.

Figure 4.12 depicts the various tax brackets in the optimal tax brackets only problem. Superscripts $^{c1}$ and $^{c2}$ respectively indicate a control variable using the model on $g_{t+1}$ and the number of returns from data $g^{*d}_{t+1}$. On the left, the tax brackets from data in 1996 (denoted $b^*_t$) is compared with the tax brackets obtained when $g_{t+1}$ is from the 1997 data (denoted $b^{*c2}_{t+1}$) and the tax brackets where $g_{t+1}$ depends on the tax liability (denoted $b^{*c2}_{t+1}$). As $\frac{\partial g}{\partial T}$ is large in lower income range, the feedback correction in $g_{t+1}$ takes over the widening of the bracket range when there is no response in the number of returns. As $\frac{\partial g}{\partial T}$ is decreasing as income is larger, the brackets do not change. On the right, the corresponding revenues show that there is reasonable control achieved through the model of $g_{t+1}$. Naturally, $g_{t+1}$ from the 1997 data would have been wrong if $g_t$ does indeed depend on the tax liability. This is the main reason of the control discussed in this section. Similar results hold for the remaining years. They are reported in Figure B.6 in Appendix B.4.1.

Figure 4.13 shows the optimal marginal rates in the optimal marginal rates only problem. We use the same notation as before. We observe the opposite behavior to the optimal brackets only problem. As displayed on the left, the optimal marginal rates are similar in lower income ranges where $\frac{\partial g}{\partial T}$ is large. On the contrary, we have higher marginal rates in high income range where $\frac{\partial g}{\partial T}$ is small. Again, we achieve a reasonable control as showed on the right. The results for the remaining years are reported in Figure B.7 in Appendix B.4.2

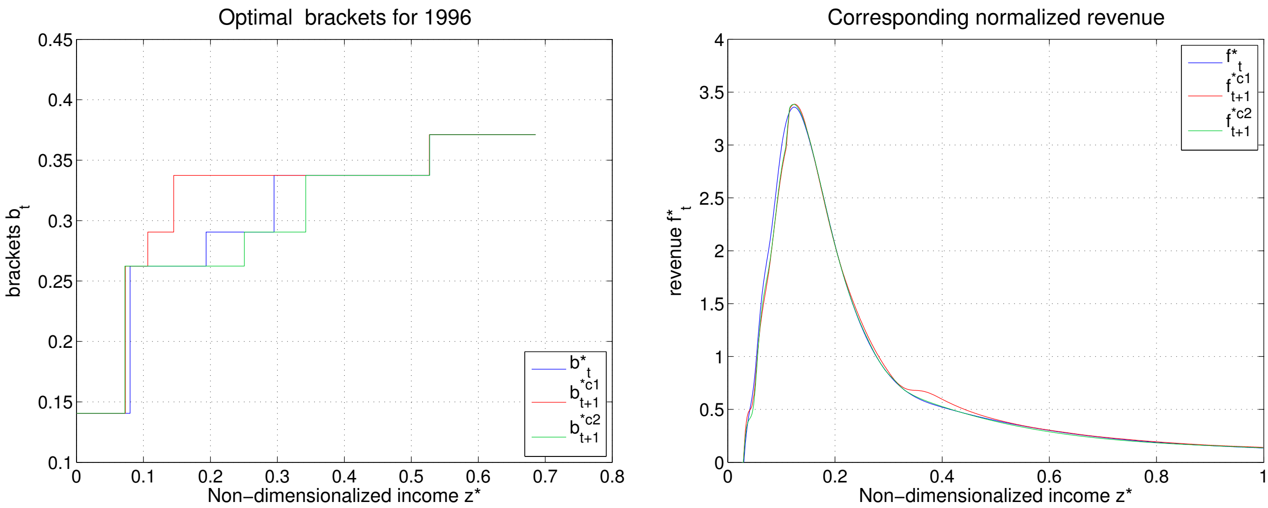



Figure 4.12: *Comparison of year to year optimal brackets from the known $g_{t+1}$ (in green) and modeled $g_{t+1}$ (in red) together with the resulting normalized revenue distribution for 1996 to 1997.*

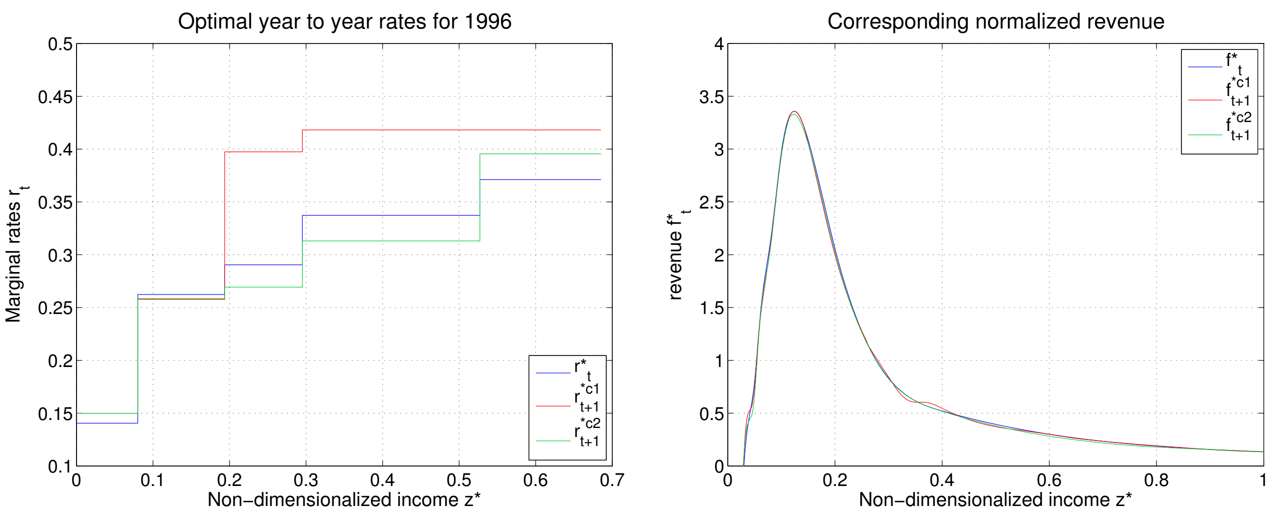



Figure 4.13: *Comparison of year to year optimal marginal rates from the known $g_{t+1}$ (in green) and modeled $g_{t+1}$ (in red) together with the resulting normalized revenue distribution for 1996 to 1997.*

Figure 4.14 depicts the various tax brackets and marginal rates in the combined optimal tax brackets and marginal rates problem. The same notation as before is used. The optimal marginal tax rates, displayed on the left, are higher than both the 1996 rates and the optimal rates while using the 1997 number of returns. A reasonable control is achieved on the right. We note that in this case the problem is higher dimensional than the previous two and it challenges convergence but with no apparent improvement on the control. Therefore, it is only advisable to implement this procedure when the previous two yield unacceptable control results. Similar results hold for the remaining years. They are reported in Figure B.8 in Appendix B.4.3.

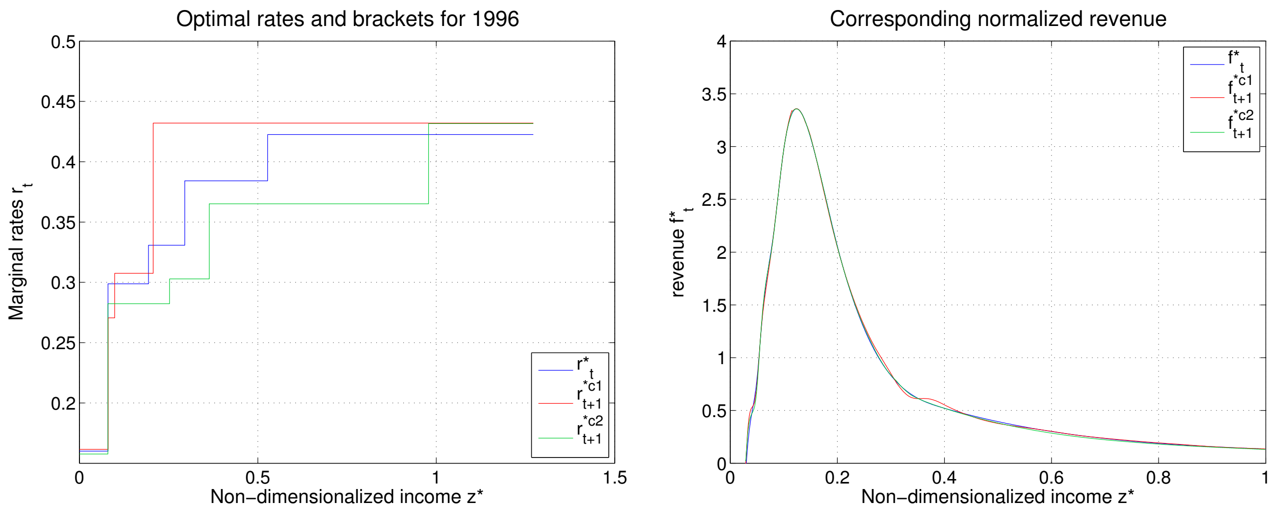



Figure 4.14: *Comparison of year to year optimal marginal rates and brackets known* $g_{t+1}$ *(in green) and modeled* $g_{t+1}$ *(in red) with the resulting normalized revenue distribution for 1996 to 1997.*

## 4.5 Conclusion

This chapter has discussed the control of instability in the relative revenue contribution. Simulation on tax data shows that if left unchanged, current tax policy would lead to a collapse of the revenue peak and would transform the revenue into a qualitatively different distribution with its mode at high income range. This potential collapse of a source of tax revenue is an instability. We control for this instability by deriving the tax liability per individual at each income that keeps constant the shape of the distribution especially at its peak.

The simple case where the future distribution of number of returns is known is treated first. The optimal tax brackets and/or marginal tax rates that minimize the change in revenue are derived. The optimal tax brackets problem is nonlinear unlike the optimal marginal tax rates problem. Both are solved numerically. The combined optimal tax brackets and marginal tax rates yields the best control but is stringed with numerical difficulties as the dimension of the problem is higher. The optimal brackets only problem gives the least satisfactory result. We conclude that the current practice of indexing the tax brackets by inflation rate could be improved by implementing the optimal marginal tax rates problem only. This is a linear problem that yields a unique solution.

The optimal tax liability formula to control for the instability does not change when the distribution of number of returns is replaced by a model that depends on changes in tax liability. The resulting optimal number of returns is an average correction to the number of returns responses due to time. Both the optimal tax liability and number of returns converge to the simple case when the response in the number of returns due to changes in tax liability vanishes. An ordinary least squares estimation shows that the number of returns responses to tax liability changes are decreased for high income and is increased for lower income. As a result, the optimal tax bracket and/or marginal tax rate changes in lower income as compared to when we completely ignore tax liability effect on the number of returns. This is an improvement in control in the sense that if the number of returns depends on the tax liability, using the number of returns from data would have been wrong as there needs to be a feedback on the control.

There are two obvious limitations to our theory. First, the Taylor expansion that is the basis of the model given in (4.1) is only valid when $\Delta T$ is small. That is, only small marginal tax parameters changes should be considered. Otherwise, the higher order terms should be included. Alternatively, a multi-year control process has to be used, and perhaps dividing larger changes into year to year smaller changes. Second, the number of returns responses to changes in tax liability $\frac{\partial g}{\partial T}$ could depend on the tax liability itself. For example, the supply-side economics argues that responses at high marginal tax rates are opposite of those of low marginal rates for the same level of income. Again, this restricts the domain of usability of our theory to local marginal changes only. Extensions to those two features are put to future work.

# Chapter 5

# Constant Fairness and Control of the Shape of the Revenue Distribution

## 5.1 Introduction

The equity-efficiency trade-off is considered to be one of the fundamental aspects of any tax design, Harberger (1964); Musgrave (1976); Feldstein (1999); Saez (2001). Concerns for equity advocates tax progression whereby those who are able to pay more should pay more. Meanwhile, tax distorts economic decisions whereby higher taxes constitute a disincentive to economic activities. Therefore the question arises as to how progressive a tax system should be?

In light of our discovery in a previous chapter that there is a hidden instability in the relative contribution to the revenue distributions, how is that instability connected to fairness or economic effectiveness? Common sense would suggest that a hidden instability is directly linked to economic inefficiency of the tax system because year by year increased relative contribution by the rich to the revenue distribution would necessarily imply distortions to economic decisions. Therefore, the natural question to ask is, can we associate optimality in the sense of controlling the instability, to optimality in fairness?

There are a number of researches that deal with the fairness of the tax system, Musgrave (1976); Young (1990). It is conventionally understood from these that the effective tax rate constitutes an acceptable measure of the degree of fairness of a tax system, GAO (2005); CBO (2007a). This is given by,

$$\bar{r}(z) = \frac{T(z)}{z}, \tag{5.1}$$

that is, tax paid as a proportion of gross income.

Economic theory deals with the optimality of a tax system. The Mirrlees framework constitutes the modern theory but is not quantitative (lack of explicit knowledge of utility and welfare function) while the theory of optimal tax reforms is recognized as not implying optimal taxation. From data, the average tax rates and the distribution of revenue in the previous chapter can be regarded as acceptable because there is no widespread racial criticism of their general shape. What is clear from the litterature we are aware of, is that the year to year change of average tax rate might hide a progressive, slow distortion. By our simulation, this itself could lead to dramatic social change.

While waiting for a clear and accepted theory, the aim of this chapter is to control for the instability while requiring a constant change in fairness. In other words, we want to find the optimal way of controlling the instability with equal change in fairness for all taxpayers. That is, a status quo control of instability that preserves relative fairness. This chapter is organized as follows. In section 2, we define the appropriate measure of fairness to our setting. In section 3, we derive the year to year control of the instability with constant change in fairness. In section 4, extensions and sensitivity to the number of returns distribution dynamics are discussed. We conclude in section 5.

## 5.2 Constant Fairness Over Tax Reforms

The ability-to-pay and the benefits-received principles are the two underlying principles that defines the equity of a tax system. The benefits-received principle links payment of a specific tax to benefits of the use of such taxes. It is not the object of this chapter. The ability-to-pay principle advocates that those who are able to pay more should contribute more. This is the main reason used for progressive taxation.

Intuitively, the fairness of a tax system is fundamentally subjective. The definition of fairness (5.1) has the benefit of simplicity. However, different individuals with the same gross income are subject to different deductions and exemptions. We are concerned that hidden effects such as the instability of $f_t^*$ will act on all individuals. In this case it is natural to subtract off known biases and to use effective taxable income $I^{eff}$ rather than gross income. We therefore choose to use the definition of fairness by the average tax rate, given by,

$$\bar{r}(z) = \frac{T(z)}{I^{eff}(z)}. \tag{5.2}$$

In spite of the disagreement as to which rate captures best the equity of a tax system, Our principle is not to define fairness or what fairness should be. We observe and control the current policy to not change with regard to relative fairness. That is, the control of instability should be done in such a way that the ratio of tax burden between any two taxpayers is constant prior and after reform. The best way to achieve this is to use the average tax rate. Deductions and exemptions play an important role in the horizontal equity of the tax system, Young (1990). More importantly, they are current policy that we cannot discard.

## 5.3 Year to year control of Instability with constant change in fairness

As seen in the previous chapter, year to year control of the instability in the normalized revenue distribution implies the tax liability given by,

$$T_{t+1}(z) = T_t(z)\,\frac{g_t(z)}{g_{t+1}(z)}. \tag{5.3}$$

On the other hand, requiring year to year constant change in fairness implies,

$$\frac{T_{t+1}(z)}{I_{t+1}(z)} - \frac{T_t(z)}{I_t(z)} = c, \tag{5.4}$$

for some constant $c \in \mathbb{R}$.

Naturally, our goal is to enforce both objectives. There is a trade-off between constant change in fairness and control of instability. In fact, enforcing control of instability results in a level of change in fairness. Hence our only option is to optimally enforce constant change in fairness while controlling for the instability. Equation (5.4) suggests that the change in fairness as measured by the constant c has to be enforced to achieve optimal control of instability. In other words, we look for the optimal constant change in fairness c that minimizes the distance between consecutive normalized revenue distributions. The problem is defined by,

$$\min_c \|[T_t \frac{I_{t+1}}{I_t} + c\, I_{t+1}]\, g_{t+1} - T_t\, g_t\|. \tag{5.5}$$

The resulting optimal tax liability is given by,

$$T_{t+1} = T_t \frac{I_{t+1}}{I_t} + c^*\, I_{t+1}, \tag{5.6}$$

where $c^*$ is the minimizer of (5.5).

As expected, we achieve a less satisfactory control of the instability as a result of the compromise to achieve constant change in fairness over the distribution. Optimal $c^*$ defines the level of constant change in fairness that provides the least change in the normalized revenue distributions. As in the previous chapter the norm of function used here is to indicate that both tax liability from the control of instability and the one that provides constant change in fairness is close to each other at all incomes $z$. Using the $L^2$ norm with the revenue $f_t$ as a weight, we have,

$$\min_c \int_0^\infty f_t(z) \left\{ [T_t(z) \frac{I_{t+1}(z)}{I_t(z)} + c\, I_{t+1}(z)]\, g_{t+1}(z) - T_t(z)\, g_t(z) \right\}^2 dz. \tag{5.7}$$

This is a convex quadratic optimization problem. One can easily derive the first order condition to obtain the minimizer given by,

$$c^* = \frac{\int_0^\infty f_t(z)\, I_{t+1}(z)\, g_{t+1}(z)\, T_t(z)\, [g_t(z) - \frac{I_{t+1}(z)}{I_t(z)}\, g_{t+1}(z)]\, dz}{\int_0^\infty f_t(z)\, I_{t+1}^2(z)\, g_{t+1}^2(z)\, dz}. \tag{5.8}$$

Rearranging terms, we have,

$$c^* = \frac{\int_0^\infty [\frac{T_t(z)\, g_t(z)}{I_{t+1}(z) g_{t+1}(z)} - r_t(z)]\, f_t(z)\, I_{t+1}^2(z)\, g_{t+1}^2(z)\, dz}{\int_0^\infty f_t(z)\, I_{t+1}^2(z)\, g_{t+1}^2(z)\, dz}. \tag{5.9}$$

Thus, the optimal constant change in fairness is the average year to year change in the average tax rate to obtain the normalized revenue $f_t = T_t(z)\, g_t(z)$ of year $t$. The average is taken using the density $f_t(z)\, I_{t+1}^2(z)\, g_{t+1}^2(z)$, that is the distribution of variance in future total taxable income. Hence, $c$ is the change in average tax rate that optimally keeps the normalized revenue close to the current one and adjusts the disparities across future taxable income.

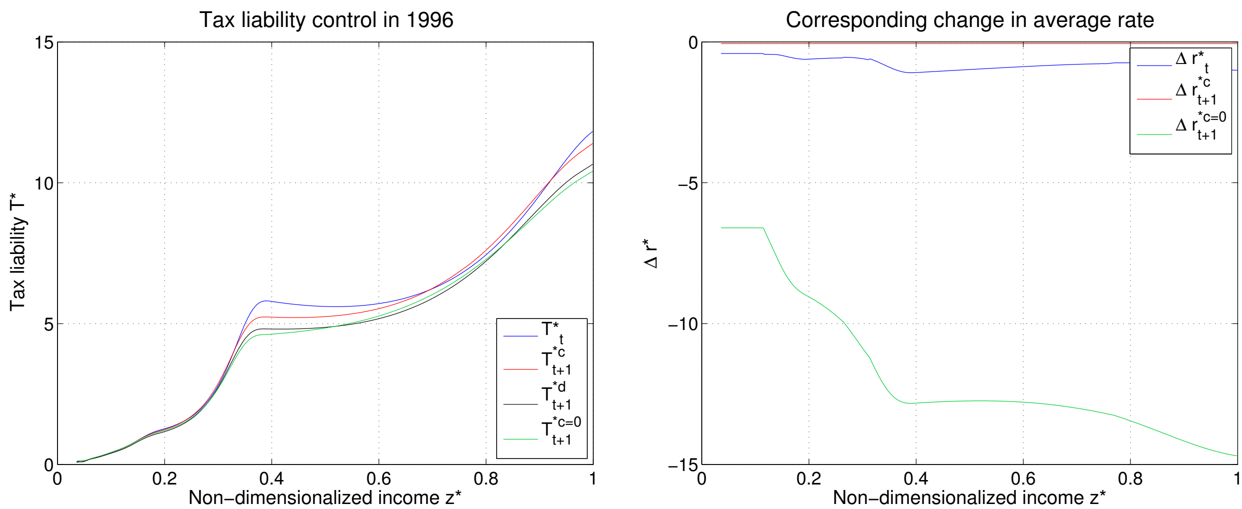



Figure 5.1: *Tax liability and change in fairness for 1996.*

Figure 5.1 shows on the left the tax liability $T_t^*$ from data for 1996 and the tax liability control $T_{t+1}^{*c}$ from (5.6), for instability with constant change in fairness. The tax liability $T_{t+1}^{*d}$ from data for 1997 as well as the tax liability $T^{*c=0}_{t+1}$ that controls solely the instability with no regards to fairness, are also plotted. The corresponding change in average tax rate is plotted on the right. It clearly shows that controlling for the instability alone has a substantial equity cost. This is an important result because this inequality is a measure of the hidden annual increase in tax burden, under the policy that tax brackets be adjusted by inflation only. Imposing fairness results in almost no change in the average tax rate ($c^* = -0.0613$ for 1996). However, this implies a larger change in the tax liability itself as it is very sensitive to change in average rate which in turn results in a less control of the instability. This will be discussed further in the next section. Similar observations on tax liability and average rates as for 1996 hold for the remaining years. The corresponding figures are put in Figure C.1 in Appendix C.

### 5.3.1 Year to year optimal tax parameters

Practical tax policy has to be given in terms of tax parameters that the government can change. As is done in the previous chapter, to control for the instability while enforcing constant change in fairness, we look for the optimal set of marginal tax rates and/or tax brackets with the corresponding tax liability given by (5.6).

Intuitively, as (5.6) is the result of the optimization problem (5.7), minimizing the distance between future tax liability and (5.6) is suboptimal. Setting the overall problem, that is simultaneously finding the optimal c (probably different from (5.9) ) and tax parameters, in theory should yield a better control of the instability in the revenue distribution. However, the problem (5.7) is well defined and has only one minimizer, that is c. One has to introduce tax parameters into the problem. There are two options to do this. The first one is to put the emphasize on the instability and extract fairness. That is, we optimally find the tax parameters that control for the instability and the optimal change in fairness that approximates the corresponding tax liability. The problem is set as follows,

$$\min_{r,b,c} \int_0^\infty f_t(z) \left\{ [T_t(z)\, g_t(z) - T(z,r,b)\, g_{t+1}(z)]^2 + [r_t + c - \frac{T(z,r,b)}{I_{t+1}}]^2 \right\} dz, \tag{5.10}$$

where the tax liability $T(z,r,b)$ is given by,

$$\begin{aligned} T(z,r,b) &= (I(z) - b_i)\, r_i + \sum_{k=1}^{i} r_{k-1}\,(b_k - b_{k-1}), \\ &= I(z)\, r_i - \sum_{k=1}^{i} b_k\,(r_k - r_{k-1}) - r_0\, b_0, \end{aligned} \tag{5.11}$$

for $b_i \leq I(z) \leq b_{i+1}$.

The corresponding first order condition for optimality yields,

$$c^* = \int_0^\infty f_t(z)\,[\frac{T(z,r^*,b^*)}{I_{t+1}} - r_t]\,dz, \tag{5.12}$$

where $c^*, r^*, b^*$ denote respectively the optimal constant change in fairness, marginal rates and brackets. Thus, the optimal constant change in fairness is the average change in average tax rate. The average is now taken directly with respect to the distribution of revenue. Clearly, the optimal c here is the best c that approximates $T(z,r^*,b^*)$ along the revenue distribution, since, the control of instability is the primary objective and $T(z,r^*,b^*)$ is obtained by containing changes in the normalized revenue. It follows that, the optimal tax liability will poorly fit the optimal $c^*$. In other words, constant change in fairness is secondary to control of instability and is poorly achieved.

Computing $c^*, r^*, b^*$ involves solving the simultaneous equations given by (5.12) and the remaining optimality conditions,

$$\begin{aligned}\int_0^\infty f_t(z)\,\frac{\partial T}{\partial \theta}(z,r,b)\{g_{t+1}(z)\,[T_t(z)\,g_t(z) - T(z,r,b)\,g_{t+1}(z)] + \\ \frac{1}{I_{t+1}}\,[r_t + c - \frac{T(z,r,b)}{I_{t+1}}]\Big\}\,dz = 0,\end{aligned} \tag{5.13}$$

for $\theta = r_0, \ldots, r_4, b_0, \ldots, b_4$ with $\frac{\partial T}{\partial \theta}(z,r,b)$ given by,

$$\frac{\partial T}{\partial b_j}(z,b) = \begin{cases} -(r_j - r_{j-1}) & \text{if } j \leq i, \\ 0 \quad \text{if } j > i \end{cases} \tag{5.14}$$

and

$$\frac{\partial T}{\partial r_j}(z,b) = \begin{cases} (b_j - b_{j-1}) & \text{if } j < i, \\ (I(z) - b_j) & \text{if } j = i, \\ 0 \quad \text{if } j > i \end{cases} \tag{5.15}$$

for $b_i \leq I(z) \leq b_{i+1}$.

(5.13) is linear for the case of rate or bracket only problems. However it is nonlinear when the both marginal rates and brackets are sought. In this case, it is solved numerically.

The alternative is to emphasize the constant change in fairness and extract the corresponding optimal tax parameters. That is, we optimally find the optimal constant change in fairness c that controls for the instability and the optimal tax parameters that approximates the corresponding tax liability. The problem is set as follows,

$$\min_{r,b,c} \int_0^{\infty} f_t(z) \left\{ [(T_t \frac{I_{t+1}}{I_t} + c\, I_{t+1})\, g_{t+1}(z) - T_t(z)\, g_t(z)]^2 \right. \\ \left. + [r_t + c - \frac{T(z,r,b)}{I_{t+1}}]^2 \right\} dz, \tag{5.16}$$

In this case, the first order condition yields,

$$c^* = \frac{\int_0^{\infty} f_t(z) \left\{ I_{t+1}^2(z)\, g_{t+1}^2(z)\, [\frac{T_t(z)\, g_t(z)}{I_{t+1}(z)\, g_{t+1}(z)} - r_t] + [\frac{T(z,r^*,b^*)}{I_{t+1}(z)} - r_t] \right\} dz}{\int_0^{\infty} f_t(z)\, [1 + I_{t+1}^2(z)\, g_{t+1}^2(z)]\, dz}. \tag{5.17}$$

This is an average of the true optimal $c$ given by (5.9) and the optimal $c$ that approximates the tax liability $T(z, r*, b*)$. In some sense, there is a correction term to account for approximation error of the tax liability with constant change in fairness by the tax liability formula involving tax parameters.

The remaining optimality conditions are given by,

$$\int_0^{\infty} f_t(z) \frac{\partial T}{\partial \theta}(z,r,b) \frac{1}{I_{t+1}} [r_t + c - \frac{T(z,r,b)}{I_{t+1}}]\, dz = 0, \tag{5.18}$$

with $\frac{\partial T}{\partial \theta}(z,r,b)$ as defined in (5.14) and (5.15). As before, the overall system of equations provided by the optimality conditions is linear for the marginal rates only or brackets only problem. It is nonlinear, and is solved numerically, for the case of combined marginal rates and brackets.

### 5.3.2 Subminimum benchmark

Heuristic approach on the US data between 1996 to 2008 shows that experimentally, a two step nested minimization of (5.16) yields a solution not far off the optimal solutions. First, the optimal $c$ to minimize instability with constant change in fairness is obtained through (5.8). The second step consists of using $c^*$ to find the optimal tax parameters that fits constant change in fairness. That is,

$$\min_{r,b} \int_0^{\infty} f_t(z)\, [r_t + c^* - \frac{T(z,r,b)}{I_{t+1}}]^2\, dz, \tag{5.19}$$

with $c^*$ given by (5.8). This reduces the dimension of the problem by one and considerably reduces its complexity. Despite having no theoretical justification it offers an alternative set of tax parameters that bounds achievable control of instability with constant change in fairness. We explore some properties of the optimal marginal rates and brackets based on this formulation.

The tax liability for $b_i \le I(z) \le b_{i+1}$ is given by,

$$\begin{aligned} T(z,r,b) &= (I(z)-b_i)\,r_i + \sum_{k=1}^{i} r_{k-1}\,(b_k - b_{k-1}), \\ &= I(z)\,r_i - \sum_{k=1}^{i} b_k\,(r_k - r_{k-1}) - r_0\,b_0. \end{aligned} \tag{5.20}$$

Hence,

$$\bar{r}(z,r,b) = \frac{T(z,r,b)}{I(z)} = r_i - \frac{\alpha_i(r,b)}{I(z)}, \tag{5.21}$$

where $\alpha_i(r,b) = \sum_{k=1}^{i} b_k\,(r_k - r_{k-1}) + r_0\,b_0$.

Now, with $b_0 = 0$, we have $\alpha_i(r+c,b) = \alpha_i(r,b)$. It follows that,

$$\begin{aligned} \bar{r}(z,r+c,b) &= r_i + c - \frac{\alpha_i(r+c,b)}{I(z)} \\ &= r_i + c - \frac{\alpha_i(r,b)}{I(z)} \\ &= \bar{r}(z,r,b) + c. \end{aligned} \tag{5.22}$$

Thus, by keeping the brackets the same (within inflation indexation as provided by the nondimensionalization procedure) and increasing each marginal tax rate by the optimal $c^*$, the resulting average tax rate will be increased by the same $c^*$. Hence, assuming no change in the individual taxable income, the optimal parameters that control for the instability with constant change in fairness is obtained by increasing each marginal rate by the optimal $c^*$ and keeping each bracket constant.

This result relies heavily on the assumption that taxable income is constant from year to year. If this is not satisfied, we have,

$$r_t + c^* - \frac{T(z,r,b)}{I_{t+1}(z)} = (r_i^t + c^* - r_i) - [\frac{\alpha_i^t(r,b)}{I_t(z)} - \frac{\alpha_i(r,b)}{I_{t+1}(z)}], \tag{5.23}$$

Now, as $z \to \infty$, $I_t(z), I_{t+1}(z) \to \infty$, this implies that the highest marginal rate satisfies $r_4^* = r_4^t + c^*$. Similar argument around the origin implies that the lowest marginal rate also satisfy $r_0^* = r_0^t + c^*$. The optimality conditions of (5.16) implies,

$$\int_0^\infty \frac{f_t(z)}{I_{t+1}(z)}\,\frac{\partial T}{\partial \theta}(z,r,b)\,[r_t + c^* - \frac{T(z,r,b)}{I_{t+1}(z)}]\,dz = 0,$$

where $\frac{\partial T}{\partial \theta}(z,r,b)$ is as defined in (5.14) and (5.15). We have,

$$\begin{cases} \displaystyle\int_{z_{b_i}}^{z_{b_{i+1}}} (I_{t+1}(z) - b_i)\,H(z,r,b)\,dz + \int_{z_{b_{i+1}}}^{\infty} (b_{i+1} - b_i)\,H(z,r,b)\,dz = 0, \\ \displaystyle\int_{z_{b_i}}^{\infty} H(z,r,b)\,dz = 0, \end{cases} \tag{5.24}$$

for $i=1,\ldots,4$ and where $H(z,r,b)=\frac{f_t(z)}{I_{t+1}(z)}\,[r_t+c^*-\frac{T(z,r,b)}{I_{t+1}(z)}]$ and $z_{b_i}$ defined by $I_t(z_{b_i})=b_i$. (5.24) is nonlinear in r and b but for the case of optimal brackets only problem, from the second equality in (5.24), we have,

$$\int_{z_{b_i}}^{z_{b_{i+1}}} H(z,r,b)\,dz=0, \quad i=1,2,3,4. \tag{5.25}$$

Let us assume $r^*=r_t+c^*$. From (5.25) for $i=1$, we have,

$$\int_{z_{b_1}}^{z_{b_2}} \frac{f_t(z)}{I_{t+1}(z)}\,[\frac{\alpha_1^t(r,b)}{I_t(z)}-\frac{\alpha_1(r,b)}{I_{t+1}(z)}]\,dz=0. \tag{5.26}$$

Since $\alpha_1=b_1(r_1-r_0)$, we have,

$$b_1^*=b_1^t\frac{\int_{z_{b_1}}^{z_{b_2}}\frac{f_t(z)}{I_{t+1}(z)\,I_t(z)}\,dz}{\int_{z_{b_1}}^{z_{b_2}}\frac{f_t(z)}{I_{t+1}^2(z)}\,dz}. \tag{5.27}$$

Similarly,

$$b_2^*=b_2^t\frac{\int_{z_{b_2}}^{z_{b_3}}\frac{f_t(z)}{I_{t+1}(z)\,I_t(z)}\,dz}{\int_{z_{b_2}}^{z_{b_3}}\frac{f_t(z)}{I_{t+1}^2(z)}\,dz}+\frac{1}{(r_2^t-r_1^t)}\frac{\int_{z_{b_2}}^{z_{b_3}}\frac{f_t(z)}{I_{t+1}(z)}\,[\frac{\alpha_1(r,b^t)}{I_t(z)}-\frac{\alpha_1(r,b^*)}{I_{t+1}(z)}]\,dz}{\int_{z_{b_2}}^{z_{b_3}}\frac{f_t(z)}{I_{t+1}^2(z)}\,dz}, \tag{5.28}$$

$$b_3^*=b_3^t\frac{\int_{z_{b_3}}^{z_{b_4}}\frac{f_t(z)}{I_{t+1}(z)\,I_t(z)}\,dz}{\int_{z_{b_3}}^{z_{b_4}}\frac{f_t(z)}{I_{t+1}^2(z)}\,dz}+\frac{1}{(r_3^t-r_2^t)}\frac{\int_{z_{b_3}}^{z_{b_4}}\frac{f_t(z)}{I_{t+1}(z)}\,[\frac{\alpha_2(r,b^t)}{I_t(z)}-\frac{\alpha_2(r,b^*)}{I_{t+1}(z)}]\,dz}{\int_{z_{b_3}}^{z_{b_4}}\frac{f_t(z)}{I_{t+1}^2(z)}\,dz}, \tag{5.29}$$

and

$$b_4^*=b_4^t\frac{\int_{z_{b_4}}^{\infty}\frac{f_t(z)}{I_{t+1}(z)\,I_t(z)}\,dz}{\int_{z_{b_4}}^{\infty}\frac{f_t(z)}{I_{t+1}^2(z)}\,dz}+\frac{1}{(r_4^t-r_3^t)}\frac{\int_{z_{b_4}}^{\infty}\frac{f_t(z)}{I_{t+1}(z)}\,[\frac{\alpha_3(r,b^t)}{I_t(z)}-\frac{\alpha_3(r,b^*)}{I_{t+1}(z)}]\,dz}{\int_{z_{b_4}}^{\infty}\frac{f_t(z)}{I_{t+1}^2(z)}\,dz}. \tag{5.30}$$

Thus, for the general case where the taxable income changes with time, the optimal bracket $b_i^*$ associated with constant change in all marginal rates is the current bracket $b_i^t$ increased by a factor,

$$\frac{\int_{z_{b_i}}^{z_{b_{i+1}}}\frac{f_t(z)}{I_{t+1}(z)\,I_t(z)}\,dz}{\int_{z_{b_i}}^{z_{b_{i+1}}}\frac{f_t(z)}{I_{t+1}^2(z)}\,dz}$$

measuring the degree of relative discrepancy between $I_{t+1}(z)$ and $I_t(z)$ and shifted by the effect of changes in lower brackets on the overall average tax rate, as given by the second term on the right hand side of (5.28), (5.29) and (5.30). This is clearer if we consider the case where $I_t = I_{t+1}$ in bracket $i$ but $I_t \neq I_{t+1}$ in the lower brackets. We have,

$$b_i^* = b_i^t + \sum_{k=1}^{i-1} (b_k^* - b_k^t) \frac{(r_k - r_{k-1})}{(r_i - r_{i-1})}. \tag{5.31}$$

Bracket $i$ is shifted by the changes in lower brackets according to its proportion from the difference in marginal rates applied in each bracket. Intuitively, the change in marginal rates for taxpayers that creep into bracket $i$ after the new bracket $b_i^*$ is in effect, would be $(r_i - r_{i-1})$. Hence, when the taxable income stays the same in one bracket then that effect has to be retracted consistently as only the constant increase in the marginal rate is necessary in that case. This is represented by the second term on the right hand side of (5.31).

### 5.3.3 Numerical implementations

In Figure 5.2, we present various optimal marginal rates, the corresponding normalized revenue and year to year change in the average tax rates for 1996. A superscript $^c$ indicates a control variable/function obtained by solving the optimal control of instability with constant change in fairness as defined in (5.9). A superscript $^b$ indicates a variable/function obtained by solving the suboptimal nested optimization problem discussed in the previous section. The superscript $^{c=0}$ indicates the control variable/function obtained by controlling the instability with no regard to change in fairness, as obtained in the previous chapter.

The optimal brackets and marginal rates for 1996 are plotted on the top left of Figure 5.2. Surprisingly, the optimal marginal rates differ very little from the 1996 applied rates. The highest change is of the order of 1.5% at the bottom rate. However, the optimal brackets are not the same. The second and fourth brackets are wider than the 1996 applied brackets and the 1996 third bracket is removed. The suboptimal benchmark rates are very close to the optimal ones. They agree over all income scale except on a short range in the third bracket that does not exist in the optimal solution but is present for the suboptimal one.

The resulting normalized revenue curves are plotted on the top right of Figure 5.2. A satisfactory control of the instability is achieved. Compared to the 1996 normalized revenue from data, the control curve is slightly lower at its peak and higher in the upper tail. The trade-off between control of instability and requiring constant change in fairness is confirmed. As expected there is minimal discrepancy between the optimal and suboptimal resulting revenue curves.

We plot in the bottom row of Figure 5.2, the corresponding change in average tax rates. Both average rates from data (the applied policy in 1996-1997) and from the sole control of instability

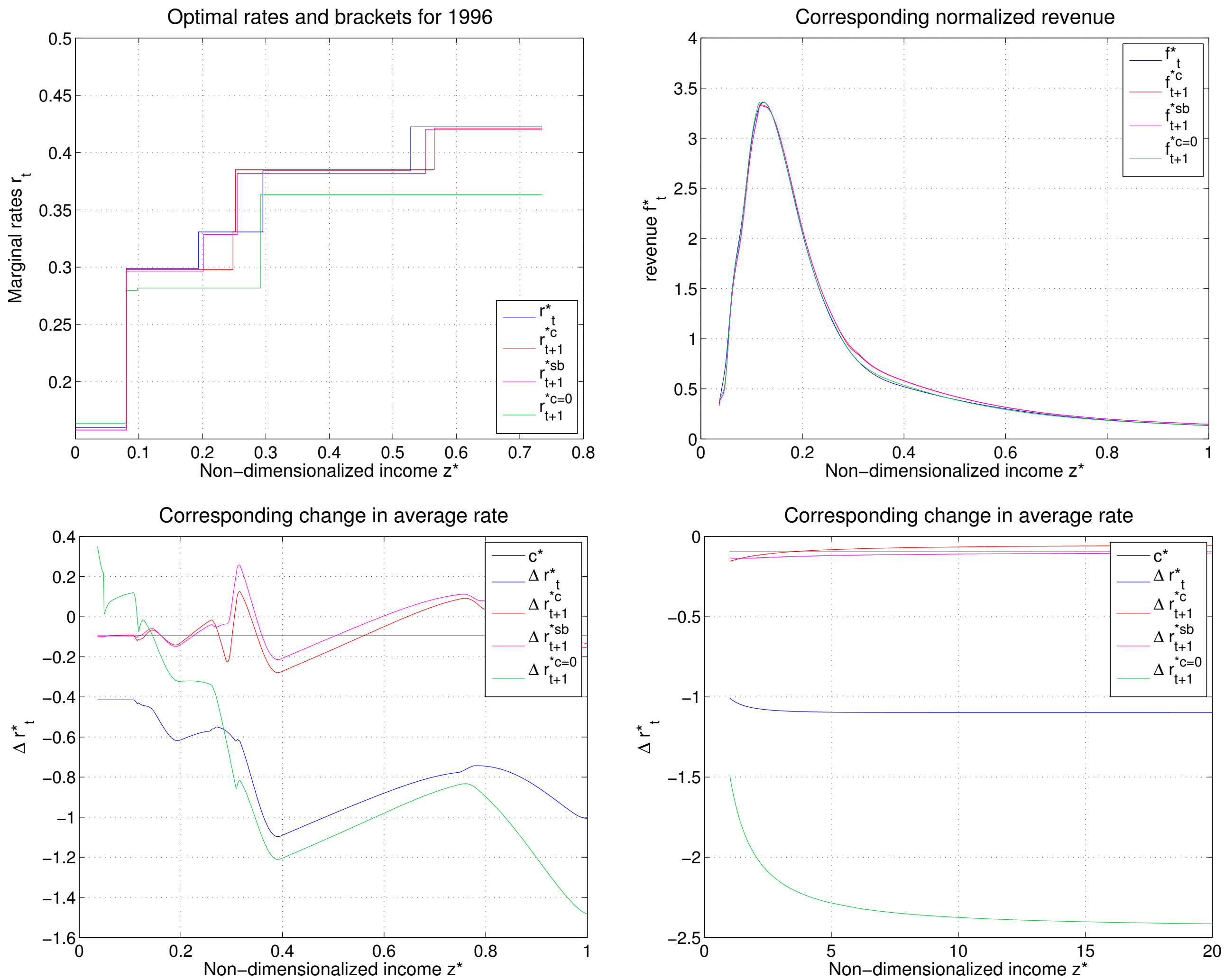



Figure 5.2: *Year to year optimal marginal tax rates and brackets, the resulting normalized revenue distribution and change in average tax rates for 1996.*

show strongly decreased fairness. Higher income range taxpayers are relatively favored by those policies. The change in average tax rate from control of instability with constant change in fairness, varies around its optimal theoretical $c^*$ value confirming that a reasonable control is achieved. The degree of variation of this control over the income axis improves at the very high income, as plotted on the lower right of Figure 5.2, than that of lower income plotted on the left. This will be discussed further below.

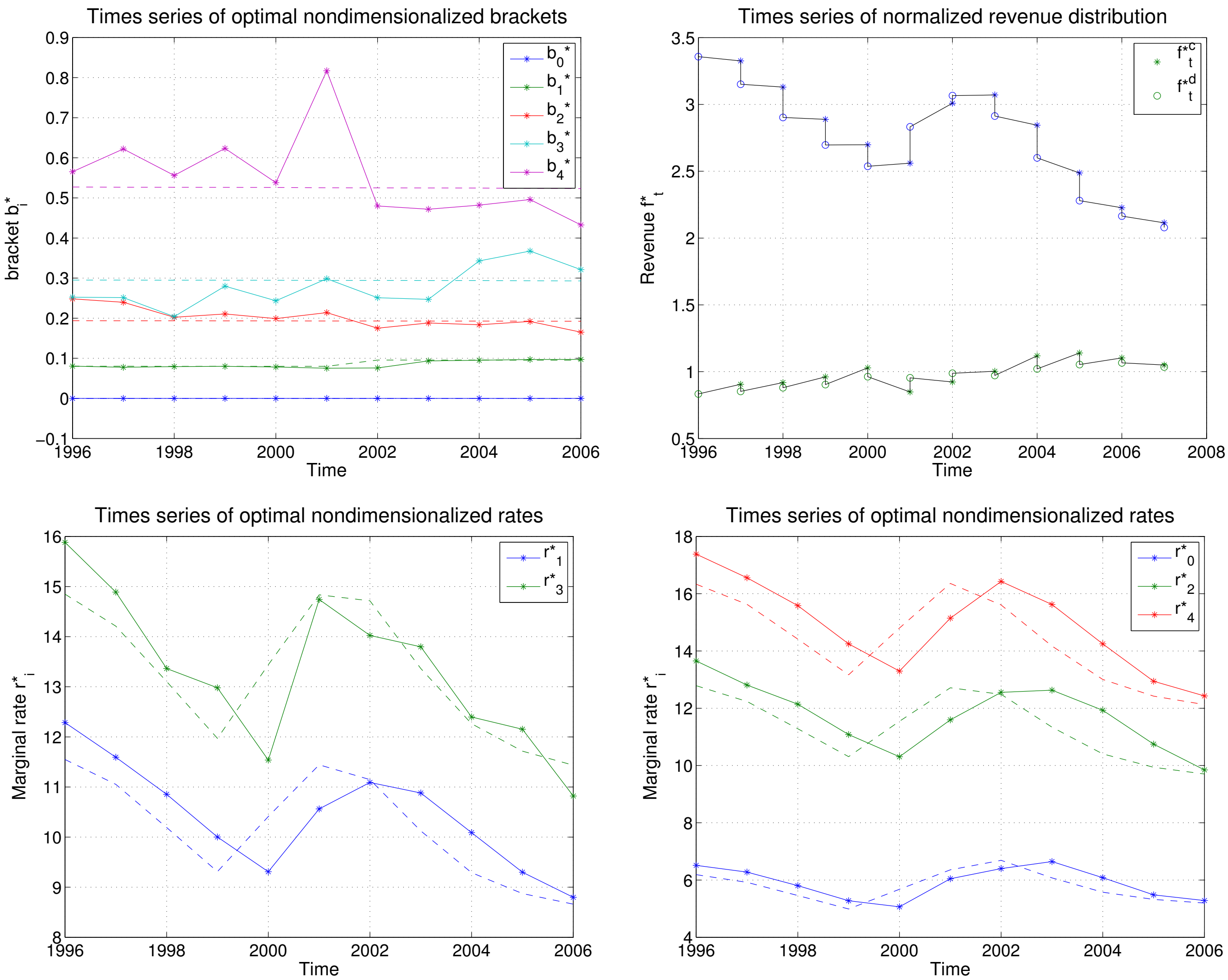



Figure 5.3: *Time series of optimal tax brackets, marginal tax rates and the normalized revenue distribution at its peak (in blue) and at $z^* = 1$ (in green) from 1996 to 2007.*

We plot in Figure 5.3 the time series from 1996 to 2007 of the optimal brackets, marginal rates and the normalized revenue distribution at the mode and at $z^* = 1$, to see if the properties of the control observed in 1996 is a general trend over the years. Nondimensionalized brackets are displayed on the top left. The first and second bracket are very close to the data (dashed lines). The third bracket only disappears from 1996 to 1998 and become less than the data until

2004 when EGTRRA transition rules are completely enabled. The top bracket is higher than the data before 2002 and lower henceforth. This could explain the general trend in the normalized revenue displayed on the top right of the figure. That is, the controlled revenue at $z^* = 1$ is almost always higher than the revenue estimated from data. Naturally, this is as a result of the compromise between constant change in fairness and control of instability. A good control of the peak is achieved. Note that they are year to year controls and that control of instability is seen by almost constant-time segments in normalized revenue at the mode. The difference of the curves at the peak is not very significant.

Time series of the nondimensionalised marginal tax rates are displayed in the bottom row of Figure 5.3. Here the trend is clear. Marginal rates are higher prior and after to EGTRRA and lower when EGTRRA was phasing in. The difference is small which makes the changes in nominal rates even smaller. The control of instability with constant change in fairness results in smaller changes in the marginal tax rates. This is expected as optimal $c^*$ is small and the suboptimal control shows that $c^*$=constant across all marginal rates is at least as good as the optimal marginal rates.

### 5.3.4 Uniformity of constant change in fairness

The bottom row of Figure 5.2 and Figure C.3 in Appendix C.1.3 show that although a better control of the change in average tax rate is achieved, this control is not uniform over the income axis. It is observed that the change in average tax rates is closer and more uniform at lower income, around the peak of the revenue distribution and at the upper tail of the revenue distribution where the Pareto distribution is used as an approximation. This suggests that the non-uniformity of the control in the middle range is due to the cubic spline estimation of the revenue/number of returns distribution where the wide bins rendered them unreliable.

To investigate this claim, we reduce the domain of control of the instability with constant change in fairness to incomes near the peak and in the Pareto tail. We plot in Figure 5.4 the change in average tax rate for the control of instability with constant change in fairness. In the top left corner is the case where the control is done near the peak only. A very good control is achieved. This control slightly changes the lower brackets and marginal rates.

Similarly, on the top right, when controlling the instability with constant change in fairness over the very high income range, at the Pareto tail of the revenue distribution, a uniform and very close control to the optimal $c^*$ change in average rate is achieved. This control also slightly changes the upper brackets and marginal rates.

Constant change in fairness over those two regions is a significant result. This is evidence that with improved data, that is, with fine bin structure over the whole range up to Pareto at $z^* = 1$, very good control of both instability and fairness can be achieved.

By controlling this instability we maintain a tax policy. Of course, this distribution is subjective, with a trade-off between acceptance by poorer and richer taxpayers. By controlling fairness, that is, constant year to year proportional change in tax burden across all income, we extend an objective method for deciding change to tax brackets. The shape of $f_t^*$ continues to be the clear and transparent responsibility of politicians.

We plot in the bottom row of Figure 5.4 the change in average tax rates to control for instability with constant change in fairness for both the lower and upper tail of the revenue distribution. The resulting change in average tax rate is not changed at the lower tail as displayed on the left. It is slightly higher than the optimal $c^*$ at the upper tail. Uniformity is achieved but the control is not as good as the theoretical $c^*$. This is expected as constant change in fairness now spans both the mode and Pareto tail of the distribution.

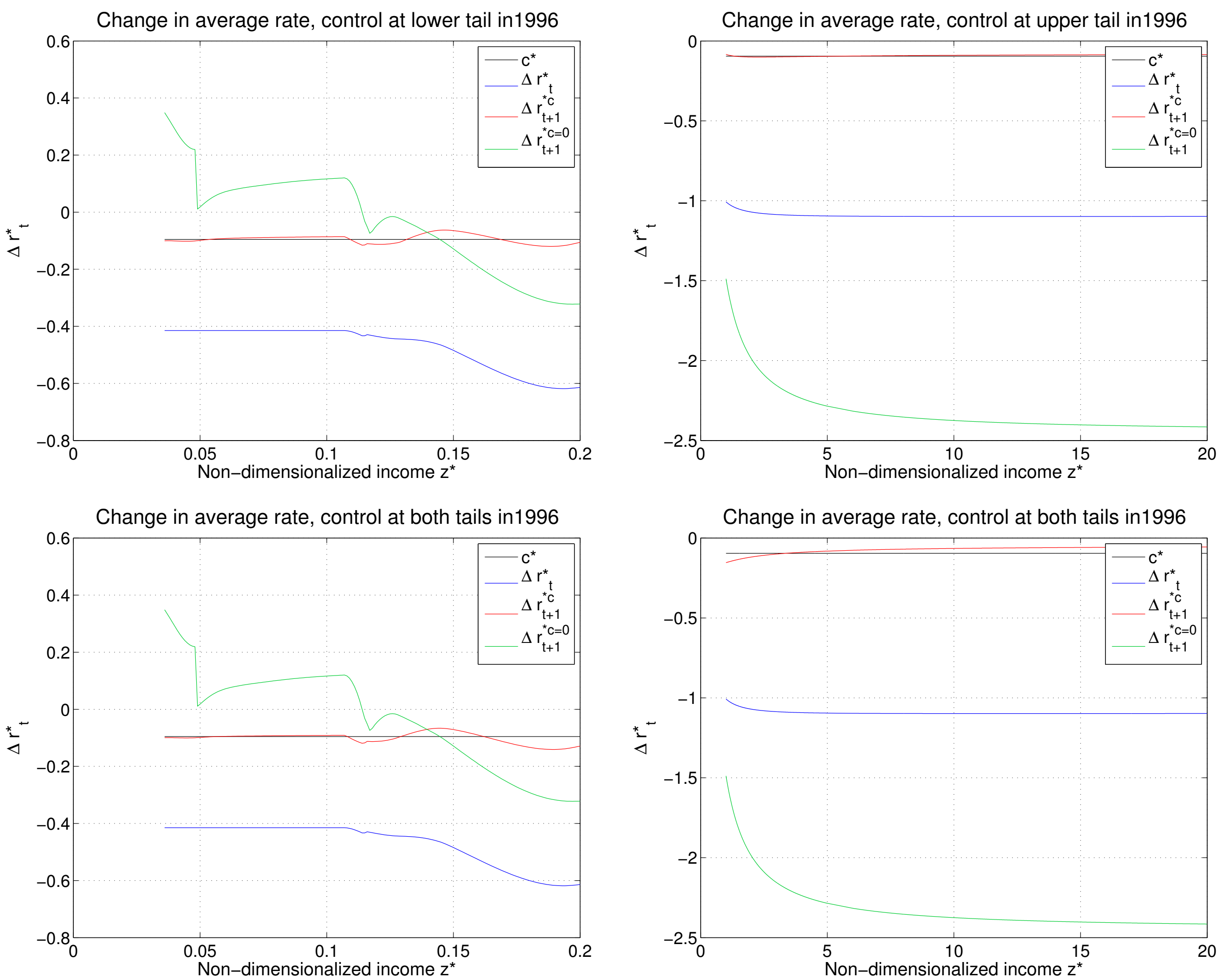



Figure 5.4: *Change in average tax rates for control of instability with constant change in fairness in 1996.*

## 5.4 Dynamics on g and Implied Control

We consider the general case where the distribution of number of returns $g_{t+1}$ depends on the tax liability. In other words, change in tax liability caused by change in tax parameters is reflected in the number of returns. We use the simple model developed in the previous chapter. That is,

$$\begin{aligned} g_{t+1} &= g_t + \frac{\partial g}{\partial T} \Delta T + \frac{\partial g}{\partial t}, \\ &= \tilde{g}_{t+1} + \frac{\partial g}{\partial T} \Delta T. \end{aligned} \tag{5.32}$$

Recall that with constant change in fairness, $T_{t+1} = (r_t + c)\, I_{t+1}$. Hence, $\Delta T = r_t \Delta I + c\, I_{t+1}$. It follows that,

$$\begin{aligned} g_{t+1} &= \tilde{g}_{t+1} + \frac{\partial g}{\partial T} (r_t \Delta I + c\, I_{t+1}) \\ &= \tilde{g}_{t+1} + r_t \Delta I \frac{\partial g}{\partial T} + c\, I_{t+1} \frac{\partial g}{\partial T} \\ &= \bar{g}_{t+1} + c\, I_{t+1} \frac{\partial g}{\partial T}, \end{aligned} \tag{5.33}$$

where $\bar{g}_{t+1} = \tilde{g}_{t+1} + r_t \Delta I \frac{\partial g}{\partial T} = g_t + \frac{\partial g}{\partial t} + r_t \Delta I \frac{\partial g}{\partial T}$, the future number of returns if there is no change in average tax rate.

The control of instability with constant change in fairness now becomes:

$$\min_c \int_0^\infty f_t(z) \left\{ (r_t(z) + c)\, I_{t+1}(z)\, [\bar{g}_{t+1} + c\, I_{t+1}(z) \frac{\partial g}{\partial T}] - T_t(z)\, g_t(z) \right\}^2 dz. \tag{5.34}$$

The corresponding optimality condition is given by,

$$\begin{aligned} &\int_0^\infty f_t(z)\, I_{t+1}(z)\, [\bar{g}_{t+1} + c\, I_{t+1}(z) \frac{\partial g}{\partial T} + \frac{\partial g}{\partial T} (r_t(z) + c)\, I_{t+1}(z)] \\ &\qquad \left\{ (r_t(z) + c)\, I_{t+1}(z)\, [\bar{g}_{t+1} + c\, I_{t+1}(z) \frac{\partial g}{\partial T}] - T_t(z)\, g_t(z) \right\} dz = 0. \end{aligned} \tag{5.35}$$

The optimal tax parameters problem is now given by,

$$\begin{aligned} \min_{r,b,c} \int_0^\infty f_t(z) \Big\{ &[(r_t + c)\, I_{t+1}\, g_{t+1}(z) - T_t(z)\, g_t(z)]^2 \\ &+ [r_t + c - \frac{T(z, r, b)}{I_{t+1}}]^2 \Big\} dz, \end{aligned} \tag{5.36}$$

with $g_{t+1} = \bar{g}_{t+1} + c\, I_{t+1} \frac{\partial g}{\partial T}$.

Since, (5.35) is a cubic polynomial with no obvious root and (5.36) is nonlinear, one has to resort to numerical methods to find the optimal tax parameters.

As is the case of control of instability, when $\frac{\partial g}{\partial T} \longrightarrow 0$, the optimal tax liability and parameters converge to their values when $g_{t+1}$ is independent of changes in tax liability.

### 5.4.1 Results for 1996

Here we compare the optimal tax brackets and/or marginal tax rates with actual and modeled $g_{t+1}$ as well as the one using $\tilde{g}_{t+1}$ where $g$ is assumed independent of changes in $T$.

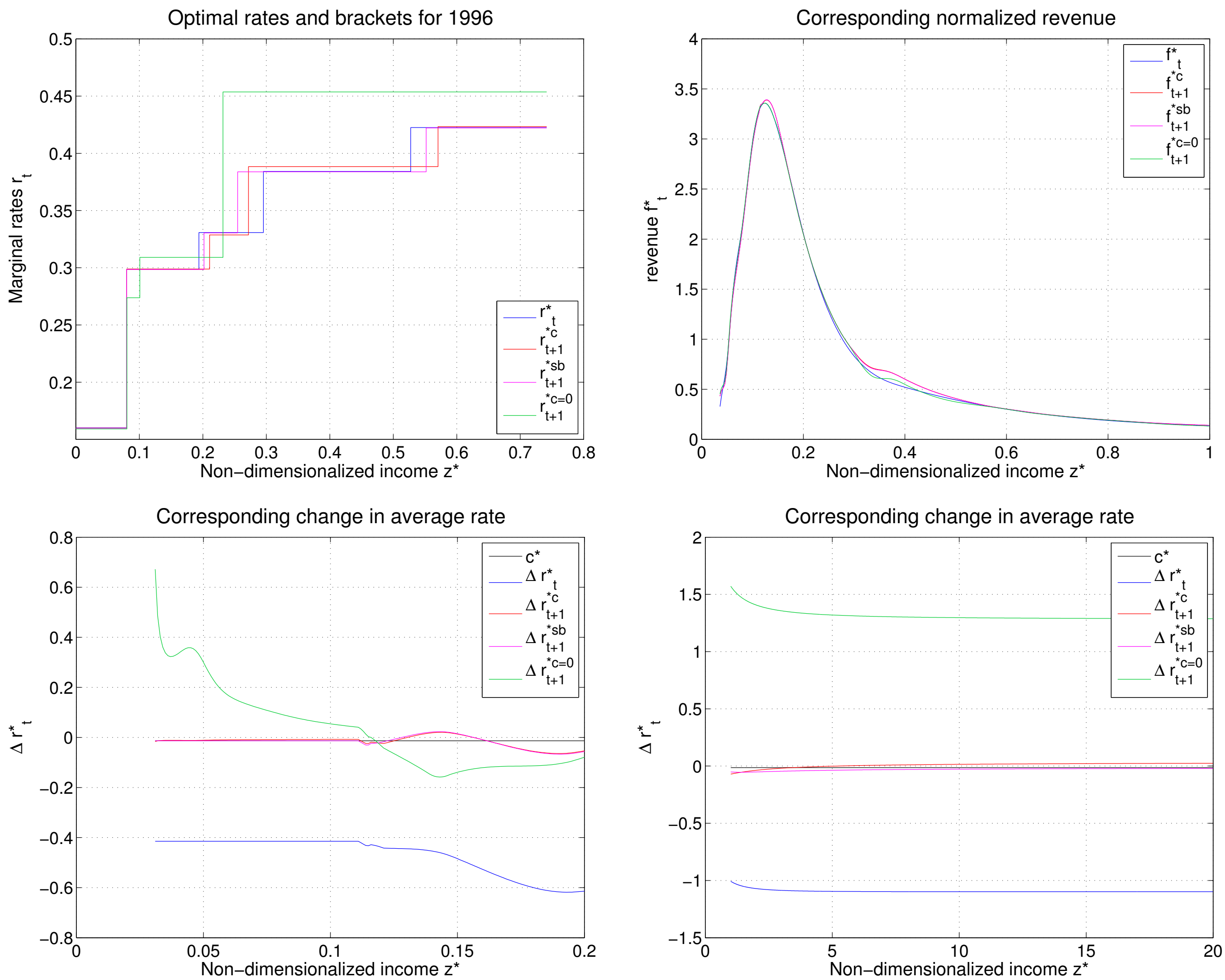



Figure 5.5: *Year to year optimal marginal tax rates and brackets, the resulting normalized revenue distribution and change in average tax rates for 1996, using a model on* $g_{t+1}$.

The optimal brackets and marginal rates for 1996 are plotted on the top left of Figure 5.5. Conrol for instability only (green) demands radical change. Otherwise, the optimal marginal rates differ very little from the 1996 applied rates as a result of the shift by $c^*$ which is small ($c = 0.082$ for 1996). The optimal brackets are not the same. As before, the suboptimal benchmark rates are very close to the optimal ones. They agree over all income scale except on the third bracket where the optimal rate is higher.

The resulting normalized revenue curves are plotted on the top right of Figure 5.5. As expected, the control is less robust but it is satisfactory near the peak where income bins are small. Compared to the 1996 normalized revenue from data, the mode is slightly shifted to the right and the peak is higher. There is no apparent discrepancy between the optimal and suboptimal resulting revenue curves.

We plot at the bottom row of Figure 5.5, the corresponding change in average tax rates. As before average rates from data, that is the applied policy in 1996-1997 show strongly decreased changes. On the other hand, the sole control of instability is improved and thus it decreases on the first bracket but converges afterwards. The change in average tax rates from the control of instability with constant change in fairness approaches its optimal theoretical $c^*$ value confirming that a reasonable control is achieved.

We plot in Figure 5.6 the time series from 1996 to 2007 of the optimal brackets, marginal rates and the normalized revenue distribution at the mode and at $z^* = 1$. Nondimensionalized brackets are displayed on the top left, normalized revenue on the top right and optimal nondimensionalised marginal tax rates are displayed in the bottom row. Both optimal brackets and marginal rates have not changed much from their optimal values when the model in $g$ is independent of changes in tax liabilities. The highest difference in the bracket is $0.0376$ at the second bracket threshold in 1996. For the marginal rate it is $0.2918$, the third marginal rate in 1997. One concludes here that the optimal parameters are unresponsive to dependence in tax liabilities of the number of returns when one controls for the instability with constant change in fairness. The difference at the peak is improved but this is no indication of the performance of the control. We have seen for example for 1996 that the mode is shifted and the peak increased but the difference at the mode of 1996 curve is small.

From the point of view of control, we may ignore changes in $g_t(z)$.

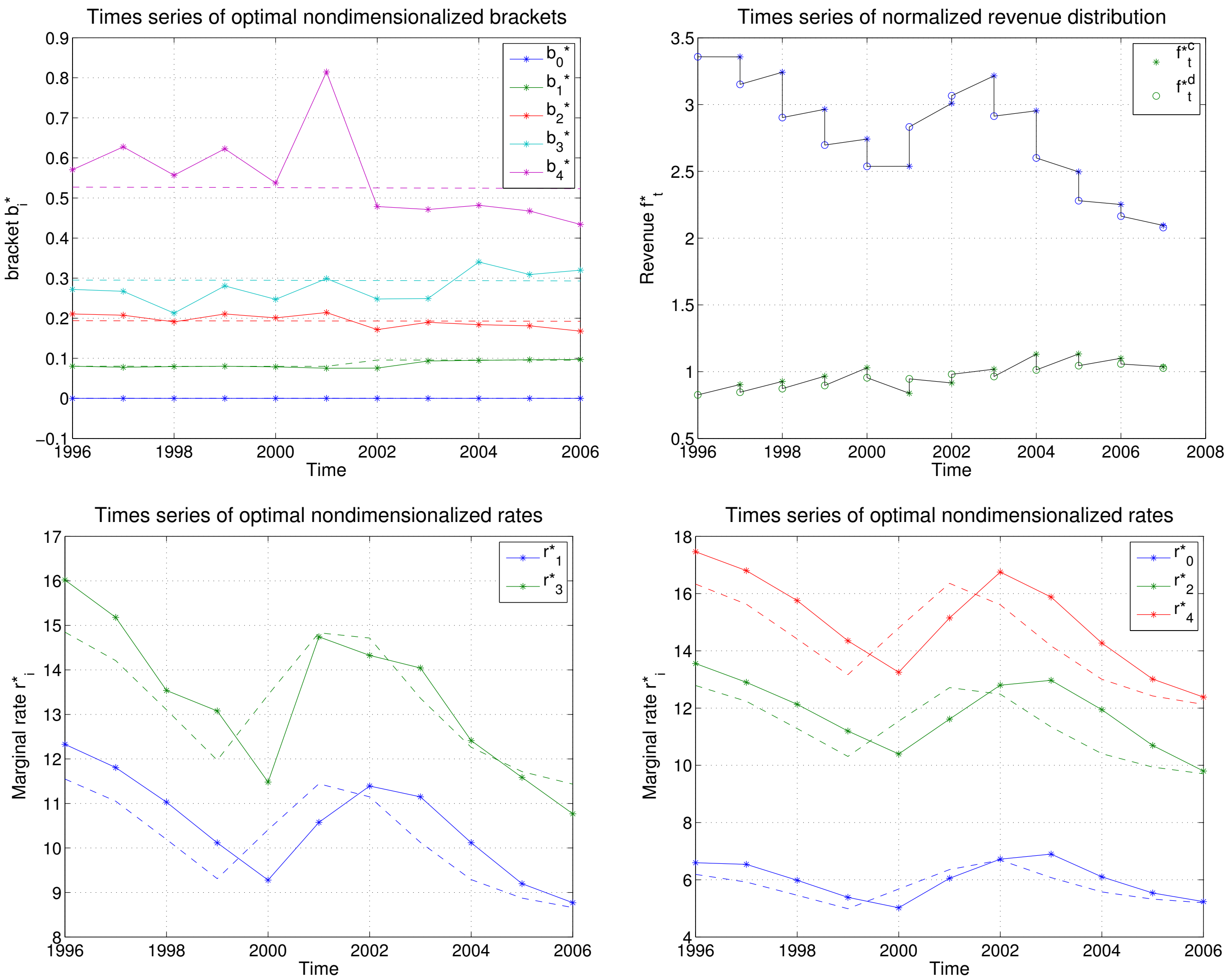



Figure 5.6: *Time series of optimal tax brackets, marginal tax rates and the normalized revenue distribution at its peak (in blue) and at $z^* = 1$ (in green) from 1996 to 2007, using a model on* $g_{t+1}$.

## 5.5 Conclusion

The sole control of instability in the normalized revenue distribution leads to substantial changes in fairness of the tax system as measured the average tax rate. In fact, controlling for instability alone has favored the rich as their average rate decreased the more you move into higher income ranges. Controlling for instability means restoring the peak of the normalized revenue distribution which in turn implies reduction of relative contribution at higher incomes. Hence, there is a trade-off between controlling the instability and resolving fairness imbalances.

This chapter focused on deriving the control of instability with constant change in fairness. As expected, less satisfactory control is achieved when constant change in fairness is required. It is an improvement in the sense that a potential collapse of the distribution will take much longer time to occur while no taxpayer is treated differently relative to each other from their treatment before the control. The unreliable estimate of the revenue curve in the middle income range where only wide income bins are provided in the data impedes on our ability to derive an overall uniform constant change in fairness. It is shown however that with appropriate data specification, one can in theory obtain reasonable constant change in fairness that control the instability.

The optimal marginal rates turn out to be very close to the statutory rates from data. This is a desirable property since no major reforms are therefore needed as opposed to the changes to marginal rates required to obtain near perfect control of instability alone. Of course, government policy may choose to institute preferences over those two opposing effects by adding more weight on one relative to the other in the optimization procedure.

We treat the general case of tax liability dependence of the distribution of number of returns. Higher dimensional nonlinear optimization is involved here but a reasonable control of instability with constant change in fairness is achieved. Time series of the optimal rates and brackets show that the optimal parameters are insensitive to the dependence of the number of returns to changes in tax liability. In other words, for the US tax system between 1996 and 2008, the control of instability with constant change in fairness could be done by just using a simple growth rate of the number of returns.

# Chapter 6

# Conclusion

In this thesis, we introduced time dependency to income taxation as an empirical additional tool to complement existing practices. These require that tax brackets be adjusted by inflation rate (Figure is empirical evidence of that). Where rates and/or brackets are adjusted, expected revenues are confirmed by microsimulation. In this thesis we, first, find an instability that is hidden in the nominal IRS data, we suppose that this is unacceptable and that a more sophisticated practice is required. This is to control tax revenue in some acceptable distribution by mathematical control of rates, brackets and fairness.

We assume that the normalized distribution of revenue changes with time as the number of returns, various tax parameters and other variables affecting the revenue, change with time. The first part of the thesis up to and including chapter 3 looked at the data history of those variables and its consequences on the revenue. The second part, chapter 4 and 5, consists of finding the optimal tax parameters to control the revenue to a desired shape. The main results of this thesis are summarized as follows:

- Our time-control approach differs from the *theoretical* Mirrleesian optimal income taxation in the following ways:
    * He considers general nonlinear taxation. This is difficult to implement. We restrict ourselves to stepwise marginal tax rates as is done in practice.
    * He uses theoretical utility and welfare functions. We use the distribution of revenue and number of returns directly from raw data.
    * Mirrlees theory explores the causes of revenue distribution by defining welfare function. Our theory does not refer to human behavior.

- Our time-dependent approach within control theory provides tax rates and tax brackets . These can be input for microsimulation as is currently done for considerations of policy change. Inversely, the output data of microsimulation can be input in our time-dependent control of revenue shape and fairness, as defined in chapters 4 and 5. Such a feedback, will help ensure that the instability does not develop. Microsimulation and our theory complement each other.

- Normalization formally admits comparative study of $f_t(z)$. Nondimensionalization facilitates comparison but also minimize the risk of numerical errors involved. With different methodologies, our results are consistent with the CBO (2007a) reports.

- Year to year comparison of normalized revenue distribution reveals an instability. The peak of the revenue decreases by as much as 30% over four years, pre- and post EGTRRA. The EGTRRA rate reductions reverse this trend. We conclude that the policy of bracket inflation indexing with constant marginal tax rates, enables the instability. Change in marginal rates can restore the status quo.

- Year to year control of the shape of $f_t^*$ is achievable. If $g_{t+1}$ is given, as we have found to be empirically reasonable, the following results hold:

  * Control by tax brackets only, widens the second bracket (ie near the peak) and increases the second bracket on average by 24%. The remaining brackets are shifted upwards accordingly. That is an increase of approximately eight times more than the inflation indexation.
  * Control by marginal rates only reduces the marginal rates by 5% on average.
  * Control using both brackets and marginal rates combines two brackets into one and thus reduces the marginal rates for some taxpayers.

  If $g_t$ is modelled, we have similar results:

  * Control by tax brackets only, changes the brackets on average by 42%
  * Control by marginal rates only, reduces the marginal rates by 16% on average.
  * Control using both brackets and marginal rates with a model on $g_{t+1}$ also combines two brackets into one and thus reduces the marginal rates for some taxpayers.

- Control of instability alone results in a huge discrepancy in change of fairness over the income axis. As a result, instability and change in fairness must be simultaneously controlled.

- Year to year control of the instability with constant fairness is achievable:
  * If $g_{t+1}$ is given, the optimal brackets change on average by 4% and the marginal rates by 0.7%.
  * If $g_{t+1}$ is modelled, the optimal brackets change on average by 5% and the marginal rates by 1%.

Regarding the management of joint returns income tax, the analysis of this thesis suggests that bracket adjustment to inflation rate, only, is not optimal as it shifts the distribution of the burden to the rich. That may be an acceptable, "time-progressive" tax policy, but if it is not transparent, it will only deceive one or both of the government and the electorate. We suggest publishing year to year comparisons of the distribution of normalized revenue and change in fairness together with the distribution of number of returns. We believe that control of the shape of normalized revenue distribution is an obligation to transparency and to effective stable tax management.

## 6.1 Future work

There are several directions to which this research could be extended to:

- The analysis performed in this thesis should be carried to other filing categories such as single taxpayers and heads of household. Corporate tax should be investigated as well.
- The overall analysis given in this thesis is impeded by the wide bins on high income range. Before any serious use of this control, the raw data must be improved and control reassessed.
- Our analysis should be extended to the study of personnal income tax returns in other countries. It is relevant to whether the instability is a universal phenomenon and not just particular to the US tax system.
- The general result here is that control of instability with or without concern for fairness is achievable. However from the above summary of results, changes in brackets and rates can be quite large (24-42%, 5-15%, respectively). An extension to our theory is to consider gradual control to a target, whereby large changes are achieved by optimally dividing control over a number of smaller annual steps. This is a well understood problem in control theory.

- The evidence of our normalized distributions is that the US tax system is not stable. Uncontrolled redistribution takes place. The economic causes of this are of future interest. We note the very complicated system of exemptions, deductions and credits in US tax policy ( see Chapter 1). Further, tax revenue results from the interplay of the effectiveness of the IRS, taxpayers behavior and compliance, tax policy and economic conditions. It is of interest to develop a precise measure of each of these effects independently of each other, at each income level.

# Appendix A

# US Joint Returns Tax Data

Here we give the various tables described in this chapter. Note that all tables are for married taxpayers filing jointly only. Tables for other filing status can be obtained from the IRS statistics of income.

## A.1 OBRA

These are the tax structure specifications from 1996 to 2000, prior to EGTRRA.

### A.1.1 Bracket Thresholds

| Marginal tax rates | 1996 | 1997 | 1998 | 1999 | 2000 |
|---|---|---|---|---|---|
| 15.0% | 0 | 0 | 0 | 0 | 0 |
| 28.0% | $40,100 | $41,200 | $42,350 | $43,050 | $43,850 |
| 31.0% | $96,900 | $99,600 | $102,300 | $104,050 | $105,950 |
| 36.0% | $147,700 | $151,750 | $155,950 | $158,550 | $161,450 |
| 39.6% | $263,750 | $271,050 | $278,450 | $283,150 | $288,350 |

Table A.1: *Tax bracket thresholds for married taxpayers filing jointly from 1996 to 2000.*

| | 1996 | 1997 | 1998 | 1999 | 2000 |
|---|---|---|---|---|---|
| Exemption | $2,550 | $2,650 | $2,700 | $2,750 | $2,800 |
| Beginning phaseout | $176,950 | $181,800 | $186,800 | $189,950 | $193,400 |
| Maximum phaseout | $299,450 | $304,300 | $309,300 | $312,450 | $315,900 |

Table A.2: *personal exemptions and phaseout intervals for married taxpayers filing jointly from 1996 to 2000.*

### A.1.2 Exemptions

## A.2 Deductions

| | 1996 | 1997 | 1998 | 1999 | 2000 |
|---|---|---|---|---|---|
| Standard deduction | $6,700 | $6,900 | $7,100 | $7,200 | $7,350 |
| Itemized deduction phaseout | $117,950 | $121,200 | $124,500 | $126,600 | $128,950 |

Table A.3: *Standard deductions and Itemized deductions phaseout levels for married taxpayers filing jointly from 1996 to 2000.*

## A.3 Combined EGTRRA and JGTRRA

This is the tax structure combining the first part of EGTRRA and JGTRRA applied from 2001 to 2008.

The exemption phaseout and phaseout rate are still $150,000, 25% respectively for married taxpayers.

### A.3.1 Marginal Tax Rates

### A.3.2 Brackets

Note, we leave blank the bracket threshold corresponding to the new 10% bracket in 2001.

| 2001 | 2002 | 2003 | 2004 | 2005 | 2006 | 2007 | 2008 |
|---|---|---|---|---|---|---|---|
| 15.0% | 10.0% | 10.0% | 10.0% | 10.0% | 10.0% | 10.0% | 10.0% |
| 15.0% | 15.0% | 15.0% | 15.0% | 15.0% | 15.0% | 15.0% | 15.0% |
| 27.5% | 27.0% | 25.0% | 25.0% | 25.0% | 25.0% | 25.0% | 25.0% |
| 30.5% | 30.0% | 28.0% | 28.0% | 28.0% | 28.0% | 28.0% | 28.0% |
| 35.5% | 35.0% | 33.0% | 33.0% | 33.0% | 33.0% | 33.0% | 33.0% |
| 39.1% | 38.6% | 35.0% | 35.0% | 35.0% | 35.0% | 35.0% | 35.0% |

Table A.4: *Marginal tax rates from 2001 to 2008.*

| 2001 | 2002 | 2003 | 2004 | 2005 | 2006 | 2007 | 2008 |
|---:|---:|---:|---:|---:|---:|---:|---:|
| $0 | $0 | $0 | $0 | $0 | $0 | $0 | $0 |
| | $12,000 | $14,000 | $14,300 | $14,600 | $15,100 | $15,650 | $16,050 |
| $45,200 | $46,700 | $56,800 | $58,100 | $59,400 | $61,300 | $63,700 | $65,100 |
| $109,250 | $112,850 | $114,650 | $117,250 | $119,950 | $123,700 | $128,500 | $131,450 |
| $166,500 | $171,950 | $174,700 | $178,650 | $182,800 | $188,450 | $195,850 | $200,300 |
| $297,350 | $307,050 | $311,950 | $319,100 | $326,450 | $336,550 | $349,700 | $357,700 |

Table A.5: *Tax bracket thresholds for married taxpayers filing jointly from 2001 to 2008.*

### A.3.3 Exemptions

## A.4 Deductions

| | **2001** | **2002** | **2003** | **2004** | **2005** | **2006** | **2007** | **2008** |
|---|---|---|---|---|---|---|---|---|
| Exemption | \$2,900 | \$3,000 | \$3,050 | \$3,100 | \$3,200 | \$3,300 | \$3,400 | \$3,500 |
| Beginning phaseout | \$199,450 | \$206,000 | \$209,250 | \$214,050 | \$218,950 | \$225,750 | \$234,600 | \$239,950 |
| Maximum phaseout | \$321,950 | \$328,500 | \$331,750 | \$336,550 | \$341,450 | \$348,250 | \$357,100 | \$362,450 |

Table A.6: *personal exemptions and phaseout intervals for married taxpayers filing jointly from 2001 to 2008.*

| | **2001** | **2002** | **2003** | **2004** | **2005** | **2006** | **2007** | **2008** |
|---|---|---|---|---|---|---|---|---|
| Standard deduction | \$7,600 | \$7,850 | \$9,500 | \$9,700 | \$10,000 | \$10,300 | \$10,700 | \$10,900 |
| Itemized deduction phaseout | \$132,950 | \$137,300 | \$139,500 | \$142,700 | \$145,950 | \$150,500 | \$156,400 | \$159,950 |

Table A.7: *Standard deductions and Itemized deductions phaseout levels for married taxpayers filing jointly from 2001 to 2008.*

## A.5 Combined 1996-2008 Data

| 1996 | 1997 | 1998 | 1999 | 2000 | 2001 | 2002 | 2003 | 2004 | 2005 | 2006 | 2007 | 2008 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 0.15 | 0.15 | 0.15 | 0.15 | 0.15 | 0.15 | 0.15 | 0.15 | 0.15 | 0.15 | 0.15 | 0.15 | 0.15 |
| 0.28 | 0.28 | 0.28 | 0.28 | 0.28 | 0.28 | 0.27 | 0.25 | 0.25 | 0.25 | 0.25 | 0.25 | 0.25 |
| 0.31 | 0.31 | 0.31 | 0.31 | 0.31 | 0.31 | 0.30 | 0.28 | 0.28 | 0.28 | 0.28 | 0.28 | 0.28 |
| 0.36 | 0.36 | 0.36 | 0.36 | 0.36 | 0.36 | 0.35 | 0.33 | 0.33 | 0.33 | 0.33 | 0.33 | 0.33 |
| 0.40 | 0.40 | 0.40 | 0.40 | 0.40 | 0.39 | 0.39 | 0.35 | 0.35 | 0.35 | 0.35 | 0.35 | 0.35 |

Table A.8: *Individual Income Marginal tax rates per bracket given by Table A.9.*

| 1996 | 1997 | 1998 | 1999 | 2000 | 2001 | 2002 | 2003 | 2004 | 2005 | 2006 | 2007 | 2008 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 40,100 | 41,200 | 42,350 | 43,050 | 43,850 | 45,200 | 46,700 | 56,800 | 58,100 | 59,400 | 61,300 | 63,700 | 65,100 |
| 96,900 | 99,600 | 102,300 | 104,050 | 105,950 | 109,250 | 112,850 | 114,650 | 117,250 | 119,950 | 123,700 | 128,500 | 131,450 |
| 147,700 | 151,750 | 155,950 | 158,550 | 161,450 | 166,500 | 171,950 | 174,700 | 178,650 | 182,800 | 188,450 | 195,850 | 200,300 |
| 263,750 | 271,050 | 278,450 | 283,150 | 288,350 | 297,350 | 307,050 | 311,950 | 319,100 | 326,450 | 336,550 | 349,700 | 357,700 |

Table A.9: *Tax brackets for married taxpayers filing jointly from 1996 to 2008. Amounts in US$.*

| Income | 1996 | 1997 | 1998 | 1999 | 2000 | 2001 | 2002 | 2003 | 2004 | 2005 | 2006 | 2007 | 2008 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 0 | 0 | 13820 | 0 | 792 | 1028 | 9370 | 433 | 208 | 220 | 6153 | 1245 | 1714 | 414 |
| 5000 | 3067 | 1993 | 0 | 1527 | 6079 | 4402 | 484 | 1044 | 38 | 2334 | 1738 | 3005 | 182 |
| 10000 | 153698 | 114309 | 63749 | 44005 | 31828 | 19074 | 5887 | 518 | 450 | 277 | 864 | 364 | 477 |
| 15000 | 1188653 | 1027433 | 885217 | 776633 | 683331 | 538603 | 335875 | 147453 | 109220 | 72305 | 45712 | 25622 | 11368 |
| 20000 | 2830584 | 2489252 | 2040473 | 1843944 | 1672477 | 1360714 | 848895 | 637879 | 523031 | 458411 | 392681 | 307318 | 219958 |
| 25000 | 5151585 | 4379311 | 3443385 | 2901854 | 2790828 | 2412163 | 1452265 | 1126896 | 994708 | 903197 | 818671 | 674514 | 526061 |
| 30000 | 16980425 | 16118252 | 12869065 | 11315056 | 10548039 | 9181528 | 6814913 | 5323753 | 4442765 | 4096184 | 3686985 | 3368608 | 2605387 |
| 40000 | 24163912 | 23109155 | 20178531 | 18401266 | 17960927 | 16632679 | 12634169 | 10790799 | 8820366 | 8315362 | 7479399 | 6843230 | 5930846 |
| 50000 | 76721419 | 77576534 | 71667361 | 70612518 | 68289254 | 65933401 | 56428611 | 49760917 | 46158685 | 44267633 | 41652349 | 40518383 | 36443628 |
| 75000 | 60974084 | 65623501 | 68807911 | 71454589 | 76569821 | 74725024 | 70504909 | 62134478 | 61250430 | 59890569 | 60497999 | 61005220 | 57653746 |
| 100000 | 93027177 | 105848663 | 119263879 | 134330849 | 151573862 | 153375902 | 144238598 | 135715164 | 144194087 | 154053096 | 167569483 | 181369966 | 182723389 |
| 200000 | 75264903 | 84885453 | 94186359 | 108775972 | 122830392 | 114534178 | 106048961 | 102935356 | 117739773 | 134187048 | 148390523 | 164078817 | 164107909 |
| 500000 | 37037975 | 43754776 | 49055531 | 55882830 | 64224627 | 57580958 | 54534220 | 51487629 | 61461496 | 72161471 | 80207974 | 87510328 | 80212539 |
| 1000000 | 78957588 | 100044612 | 120033097 | 148330072 | 29558043 | 25527234 | 23123307 | 21527036 | 26369452 | 31934763 | 37344755 | 40761589 | 35595168 |
| 1500000 | – | – | – | – | 18923840 | 15522093 | 13486253 | 13048736 | 16464182 | 20287098 | 22380918 | 24614555 | 21379643 |
| 2000000 | – | – | – | – | 48322024 | 38400829 | 32385619 | 30947934 | 40352595 | 50977694 | 58580821 | 64779947 | 52419969 |
| 5000000 | – | – | – | – | 28406411 | 19943231 | 16442409 | 15833503 | 21186596 | 28347173 | 32721487 | 36047029 | 28414762 |
| 10000000 | – | – | – | – | 62171994 | 36269452 | 25730797 | 27978527 | 42692338 | 61119553 | 71934726 | 88165288 | 65832735 |

Table A.10: *Histogram joint returns income tax revenue data. Revenue amounts in thousands of dollar. Source IRS: Statistics of Income (2008), Table 1.2.*

| income | 1996 | 1997 | 1998 | 1999 | 2000 | 2001 | 2002 | 2003 | 2004 | 2005 | 2006 | 2007 | 2008 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 0 | 0 | 7792 | 0 | 967 | 590 | 1990 | 2005 | 23 | 39 | 344 | 261 | 1652 | 13 |
| 5000 | 1241 | 343 | 0 | 1983 | 1351 | 4183 | 1046 | 49 | 17 | 1058 | 1023 | 1138 | 655 |
| 10000 | 812267 | 625533 | 436181 | 333792 | 251158 | 169110 | 88780 | 3990 | 2049 | 670 | 1058 | 752 | 338 |
| 15000 | 1808437 | 1680988 | 1577940 | 1514341 | 1381021 | 1307926 | 1222199 | 831053 | 698016 | 557644 | 364714 | 253851 | 116621 |
| 20000 | 2449255 | 2168036 | 1836516 | 1692576 | 1581282 | 1443955 | 1287690 | 1199103 | 1086806 | 1061950 | 991346 | 902130 | 745647 |
| 25000 | 2933833 | 2595072 | 2196305 | 1893708 | 1814200 | 1647811 | 1413266 | 1245080 | 1133085 | 1081465 | 1020444 | 906688 | 827933 |
| 30000 | 6044639 | 5918174 | 5349047 | 4946432 | 4620366 | 4236090 | 3933032 | 3498278 | 2958767 | 2826553 | 2630643 | 2490405 | 2182737 |
| 40000 | 5864343 | 5752022 | 5595343 | 5282650 | 5233725 | 5082236 | 4728996 | 4375033 | 3748642 | 3587071 | 3283227 | 3132765 | 2854330 |
| 50000 | 11299683 | 11644182 | 11716060 | 11976409 | 11633334 | 11775034 | 11494860 | 11043208 | 10890037 | 10605815 | 10194158 | 10052491 | 9295400 |
| 75000 | 5029341 | 5520445 | 6130908 | 6469052 | 7040507 | 7223481 | 7536872 | 7737447 | 8069006 | 8185559 | 8496260 | 8752815 | 8478886 |
| 100000 | 4019938 | 4633004 | 5367156 | 6102311 | 6860666 | 7225680 | 7187168 | 7629783 | 8329227 | 9161709 | 10133172 | 11161429 | 11449186 |
| 200000 | 1046148 | 1220515 | 1383004 | 1605259 | 1813324 | 1733190 | 1634831 | 1716464 | 1996458 | 2317452 | 2642480 | 2946286 | 2976732 |
| 500000 | 183524 | 224009 | 258898 | 293271 | 334090 | 300837 | 289023 | 302800 | 371024 | 440361 | 497749 | 546882 | 487791 |
| 1000000 | 92988 | 120181 | 143371 | 170009 | 83063 | 71689 | 65254 | 67960 | 87098 | 105804 | 125784 | 138388 | 117540 |
| 1500000 | – | – | – | – | 37230 | 30385 | 26373 | 28792 | 37676 | 47311 | 53101 | 58464 | 49235 |
| 2000000 | – | – | – | – | 55010 | 43337 | 36735 | 40177 | 54571 | 69515 | 80183 | 89523 | 70387 |
| 5000000 | – | – | – | – | 14478 | 10041 | 8188 | 9109 | 12885 | 17375 | 20467 | 22762 | 17185 |
| 10000000 | – | – | – | – | 9239 | 5513 | 4169 | 4899 | 7702 | 10982 | 12828 | 14852 | 10716 |

Table A.11: *Histogram number of joint returns data. Source IRS: Statistics of Income (2008), Table 1.2.*

| 1996 | 1997 | 1998 | 1999 | 2000 | 2001 | 2002 | 2003 | 2004 | 2005 | 2006 | 2007 | 2008 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 5 | 5.1148 | 5.1941 | 5.3088 | 5.4876 | 5.6438 | 5.7331 | 5.8632 | 6.0195 | 6.2239 | 6.4239 | 6.6075 | 6.8614 |

Table A.12: *Time series of $z_0$, income non-dimensionalizing factor. Amounts in \$100000*

| 1996 | 1997 | 1998 | 1999 | 2000 | 2001 | 2002 | 2003 | 2004 | 2005 | 2006 | 2007 | 2008 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 |
| 0.0500 | 0.0489 | 0.0481 | 0.0471 | 0.0456 | 0.0443 | 0.0436 | 0.0426 | 0.0415 | 0.0402 | 0.0389 | 0.0378 | 0.0364 |
| 0.0600 | 0.0587 | 0.0578 | 0.0565 | 0.0547 | 0.0532 | 0.0523 | 0.0512 | 0.0498 | 0.0482 | 0.0467 | 0.0454 | 0.0437 |
| 0.0800 | 0.0782 | 0.0770 | 0.0753 | 0.0729 | 0.0709 | 0.0698 | 0.0682 | 0.0665 | 0.0643 | 0.0623 | 0.0605 | 0.0583 |
| 0.1000 | 0.0978 | 0.0963 | 0.0942 | 0.0911 | 0.0886 | 0.0872 | 0.0853 | 0.0831 | 0.0803 | 0.0778 | 0.0757 | 0.0729 |
| 0.1500 | 0.1466 | 0.1444 | 0.1413 | 0.1367 | 0.1329 | 0.1308 | 0.1279 | 0.1246 | 0.1205 | 0.1168 | 0.1135 | 0.1093 |
| 0.2000 | 0.1955 | 0.1925 | 0.1884 | 0.1822 | 0.1772 | 0.1744 | 0.1706 | 0.1661 | 0.1607 | 0.1557 | 0.1513 | 0.1457 |
| 0.4000 | 0.3910 | 0.3851 | 0.3767 | 0.3645 | 0.3544 | 0.3489 | 0.3411 | 0.3323 | 0.3213 | 0.3113 | 0.3027 | 0.2915 |
| 1.0000 | 0.9776 | 0.9626 | 0.9418 | 0.9112 | 0.8859 | 0.8721 | 0.8528 | 0.8306 | 0.8034 | 0.7783 | 0.7567 | 0.7287 |
| 2.0000 | 1.9551 | 1.9253 | 1.8837 | 1.8223 | 1.7719 | 1.7443 | 1.7055 | 1.6613 | 1.6067 | 1.5567 | 1.5134 | 1.4574 |
| – | – | – | – | 2.7335 | 2.6578 | 2.6164 | 2.5583 | 2.4919 | 2.4101 | 2.3350 | 2.2702 | 2.1861 |
| – | – | – | – | 3.6446 | 3.5437 | 3.4885 | 3.4111 | 3.3226 | 3.2134 | 3.1134 | 3.0269 | 2.9149 |
| – | – | – | – | 9.1115 | 8.8593 | 8.7213 | 8.5277 | 8.3064 | 8.0336 | 7.7834 | 7.5672 | 7.2871 |
| – | – | – | – | 18.2230 | 17.7185 | 17.4426 | 17.0555 | 16.6128 | 16.0671 | 15.5668 | 15.1344 | 14.5743 |

Table A.13: *Yearly inflation adjusted, non-dimensionalized income threshold.*

| **1996** | **1997** | **1998** | **1999** | **2000** | **2001** | **2002** | **2003** | **2004** | **2005** | **2006** | **2007** | **2008** |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 0.0088 | 0.0069 | 0.0053 | 0.0043 | 0.0034 | 0.0031 | 0.0021 | 0.0015 | 0.0011 | 0.0008 | 0.0006 | 0.0004 | 0.0003 |
| 0.0109 | 0.0083 | 0.0061 | 0.0046 | 0.0040 | 0.0038 | 0.0026 | 0.0021 | 0.0017 | 0.0013 | 0.0011 | 0.0008 | 0.0007 |
| 0.0359 | 0.0307 | 0.0229 | 0.0181 | 0.0150 | 0.0145 | 0.0121 | 0.0101 | 0.0075 | 0.0061 | 0.0050 | 0.0042 | 0.0035 |
| 0.0511 | 0.0440 | 0.0359 | 0.0295 | 0.0255 | 0.0263 | 0.0224 | 0.0204 | 0.0149 | 0.0124 | 0.0102 | 0.0086 | 0.0081 |
| 0.1624 | 0.1478 | 0.1274 | 0.1130 | 0.0969 | 0.1043 | 0.0999 | 0.0940 | 0.0779 | 0.0660 | 0.0568 | 0.0506 | 0.0496 |
| 0.1291 | 0.1250 | 0.1223 | 0.1144 | 0.1087 | 0.1182 | 0.1248 | 0.1174 | 0.1033 | 0.0892 | 0.0825 | 0.0762 | 0.0785 |
| 0.1969 | 0.2016 | 0.2120 | 0.2150 | 0.2151 | 0.2427 | 0.2553 | 0.2564 | 0.2433 | 0.2296 | 0.2284 | 0.2267 | 0.2489 |
| 0.1593 | 0.1617 | 0.1674 | 0.1741 | 0.1743 | 0.1812 | 0.1877 | 0.1944 | 0.1986 | 0.2000 | 0.2022 | 0.2051 | 0.2236 |
| 0.0784 | 0.0833 | 0.0872 | 0.0895 | 0.0912 | 0.0911 | 0.0965 | 0.0973 | 0.1037 | 0.1075 | 0.1093 | 0.1094 | 0.1093 |
| 0.1671 | 0.1906 | 0.2134 | 0.2375 | 0.0420 | 0.0404 | 0.0409 | 0.0407 | 0.0445 | 0.0476 | 0.0509 | 0.0509 | 0.0485 |
| – | – | – | – | 0.0269 | 0.0246 | 0.0239 | 0.0246 | 0.0278 | 0.0302 | 0.0305 | 0.0308 | 0.0291 |
| – | – | – | – | 0.0686 | 0.0608 | 0.0573 | 0.0585 | 0.0681 | 0.0760 | 0.0798 | 0.0810 | 0.0714 |
| – | – | – | – | 0.0403 | 0.0316 | 0.0291 | 0.0299 | 0.0357 | 0.0422 | 0.0446 | 0.0451 | 0.0387 |
| – | – | – | – | 0.0882 | 0.0574 | 0.0455 | 0.0528 | 0.0720 | 0.0911 | 0.0980 | 0.1102 | 0.0897 |

Table A.14: *Yearly normalized, non-dimensionalized revenue per income thershold in Table A.13.*

| **1996** | **1997** | **1998** | **1999** | **2000** | **2001** | **2002** | **2003** | **2004** | **2005** | **2006** | **2007** | **2008** |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 0.1219 | 0.1065 | 0.0917 | 0.0838 | 0.0752 | 0.0692 | 0.0635 | 0.0512 | 0.0453 | 0.0405 | 0.0335 | 0.0280 | 0.0218 |
| 0.0705 | 0.0616 | 0.0523 | 0.0448 | 0.0424 | 0.0389 | 0.0345 | 0.0313 | 0.0287 | 0.0270 | 0.0252 | 0.0219 | 0.0209 |
| 0.1454 | 0.1405 | 0.1274 | 0.1170 | 0.1080 | 0.1001 | 0.0960 | 0.0880 | 0.0749 | 0.0705 | 0.0649 | 0.0600 | 0.0550 |
| 0.1410 | 0.1366 | 0.1333 | 0.1249 | 0.1224 | 0.1201 | 0.1155 | 0.1101 | 0.0949 | 0.0895 | 0.0810 | 0.0755 | 0.0719 |
| 0.2717 | 0.2765 | 0.2790 | 0.2832 | 0.2720 | 0.2783 | 0.2806 | 0.2779 | 0.2758 | 0.2646 | 0.2514 | 0.2424 | 0.2343 |
| 0.1209 | 0.1311 | 0.1460 | 0.1530 | 0.1646 | 0.1707 | 0.1840 | 0.1947 | 0.2044 | 0.2042 | 0.2095 | 0.2110 | 0.2137 |
| 0.0967 | 0.1100 | 0.1278 | 0.1443 | 0.1604 | 0.1708 | 0.1755 | 0.1920 | 0.2110 | 0.2286 | 0.2499 | 0.2691 | 0.2885 |
| 0.0252 | 0.0290 | 0.0329 | 0.0380 | 0.0424 | 0.0410 | 0.0399 | 0.0432 | 0.0506 | 0.0578 | 0.0652 | 0.0710 | 0.0750 |
| 0.0044 | 0.0053 | 0.0062 | 0.0069 | 0.0078 | 0.0071 | 0.0071 | 0.0076 | 0.0094 | 0.0110 | 0.0123 | 0.0132 | 0.0123 |
| 0.0022 | 0.0029 | 0.0034 | 0.0040 | 0.0019 | 0.0017 | 0.0016 | 0.0017 | 0.0022 | 0.0026 | 0.0031 | 0.0033 | 0.0030 |
| – | – | – | – | 0.0009 | 0.0007 | 0.0006 | 0.0007 | 0.0010 | 0.0012 | 0.0013 | 0.0014 | 0.0012 |
| – | – | – | – | 0.0013 | 0.0010 | 0.0009 | 0.0010 | 0.0014 | 0.0017 | 0.0020 | 0.0022 | 0.0018 |
| – | – | – | – | 0.0003 | 0.0002 | 0.0002 | 0.0002 | 0.0003 | 0.0004 | 0.0005 | 0.0005 | 0.0004 |
| – | – | – | – | 0.0002 | 0.0001 | 0.0001 | 0.0001 | 0.0002 | 0.0003 | 0.0003 | 0.0004 | 0.0003 |

Table A.15: *Yearly, normalized, non-dimensionalized number of returns per income threshold as in Table A.13*

# Appendix B

# Control of instability

Here we present the figures of the remaining years for various control developed as well as some derivation of various proofs left in chapter 4.

## B.1 Effective Taxable income and deductions

Figure B.1 displays the effective taxable income and deductions as well as their theoretical values. Legends, colors and notations are the same as in Figure 4.3 in section 4.3.1 of chapter 4.

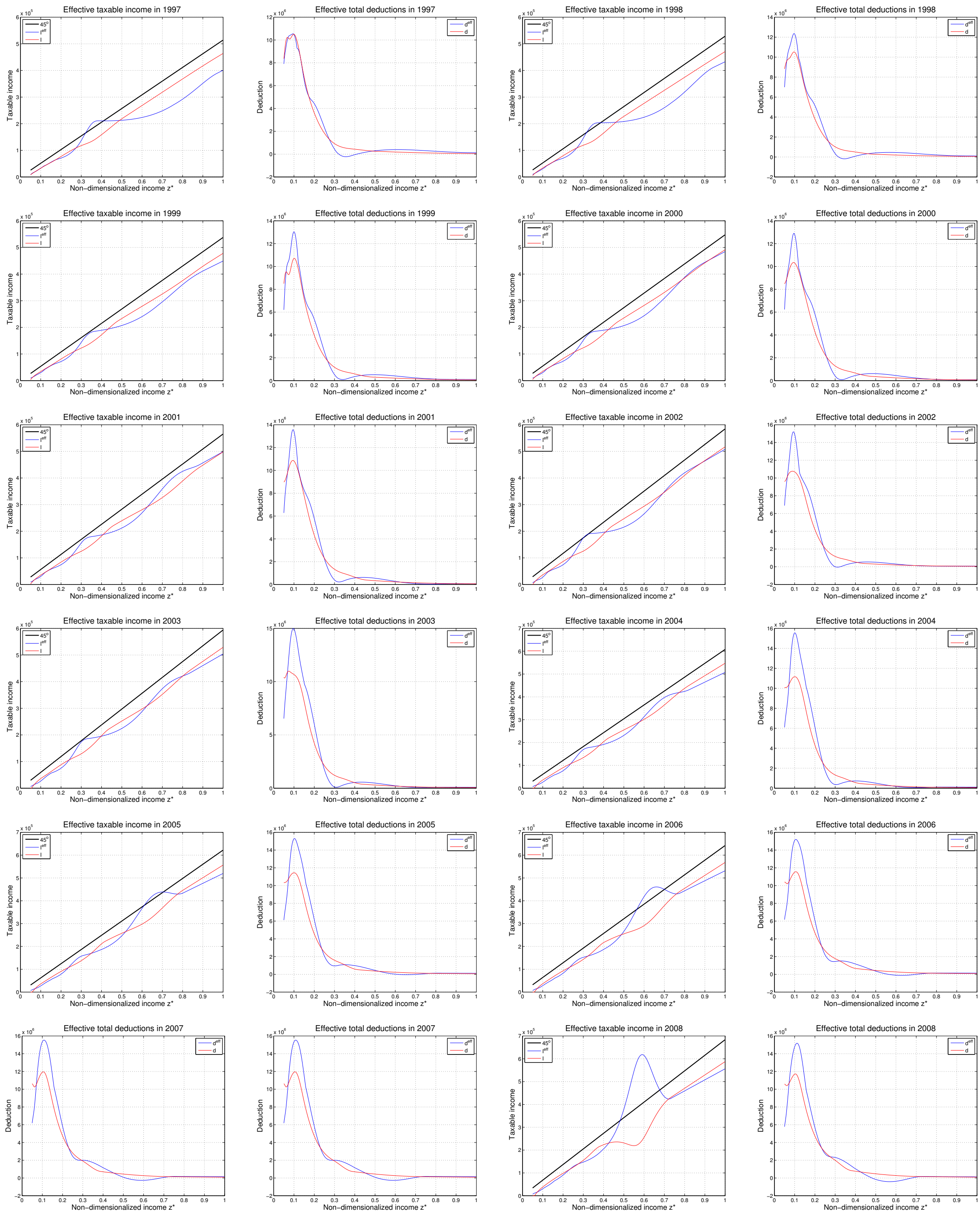


Figure B.1: *Effective taxable income and deductions from 1997 to 2008.*

## B.2 Control given $g_{t+1}$

we assume $g_{t+1}$ is as given by the data.

### B.2.1 Year to year tax liability control

Figure B.2 displays the tax liability required to control for the instability in the revenue distribution. It is compared with actual tax liability from data and the year to year ratio of the number of returns is displayed as well. Legends, colors and notations are the same as in Figure 4.2, explained in section 4.3 of chapter 4.

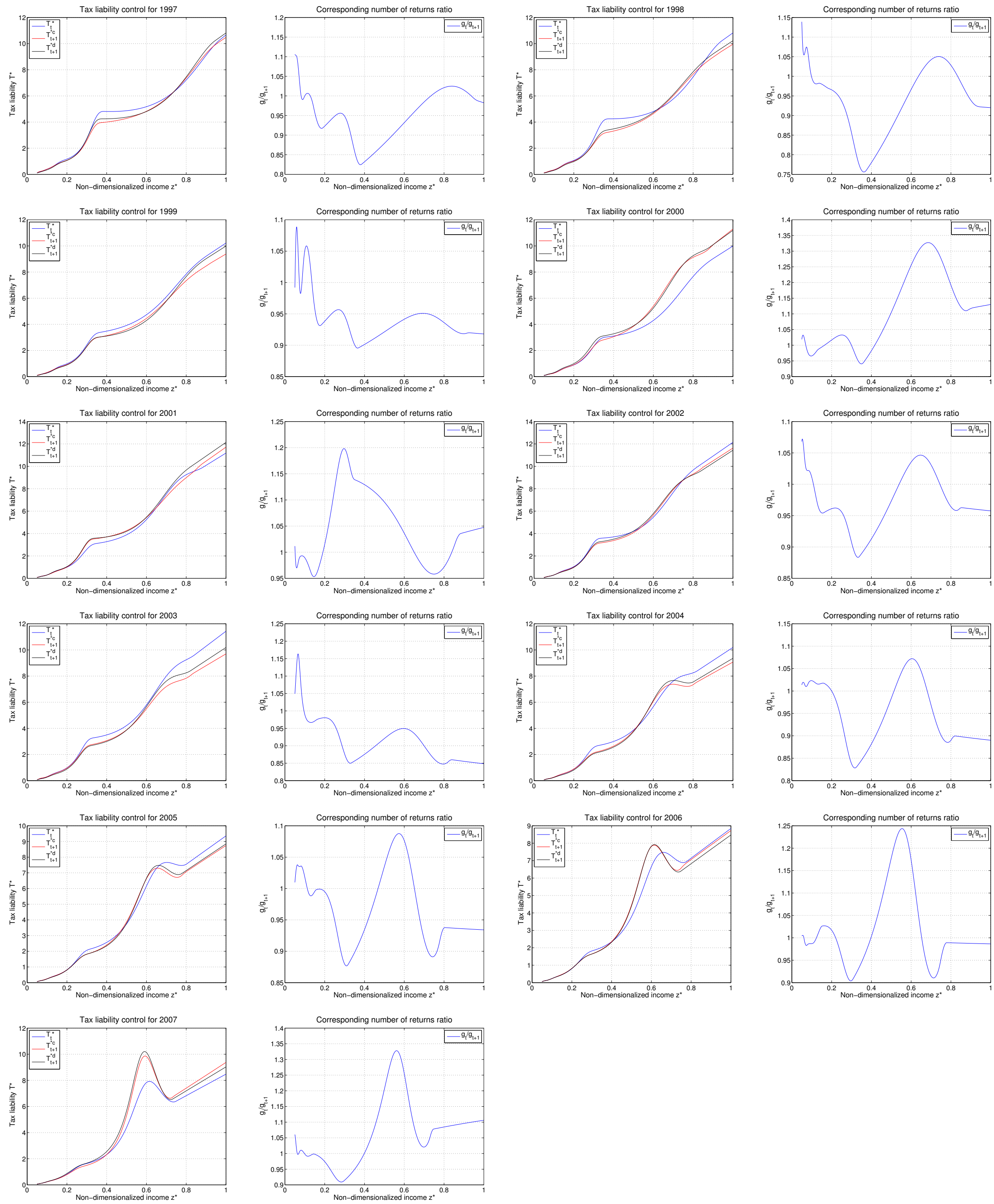


Figure B.2: *Tax liability control and corresponding number of returns ratio from 1997 to 2008.*

## B.2.2 Optimal brackets

Figure B.3 displays the optimal tax brackets to control for the instability in the revenue distribution. It is compared with actual brackets from data. The resulting revenue is plotted on the right. Legends, colors and notations are the same as in Figure 4.4, explained in section 4.3 of chapter 4.

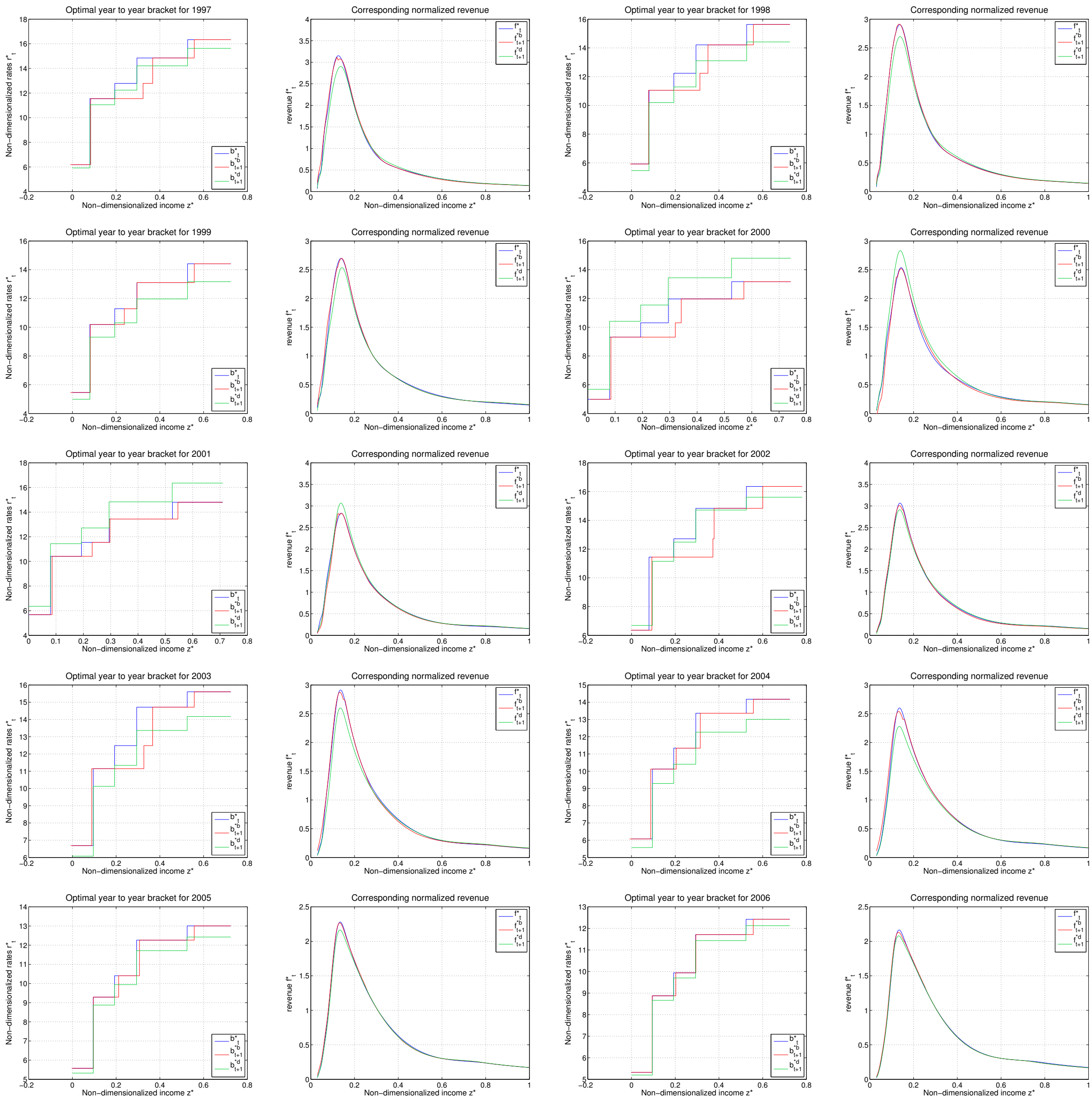



Figure B.3: *Optimal brackets from 1997 to 2006 and the corresponding normalized revenue.*

### B.2.3 Optimal rates

Figure B.4 displays the optimal marginal rates to control for the instability in the revenue distribution. It is compared with actual marginal rates from data. The corresponding revenue distribution is plotted on the right. Legends, colors and notations are the same as in Figure 4.6, explained in section 4.3 of chapter 4.

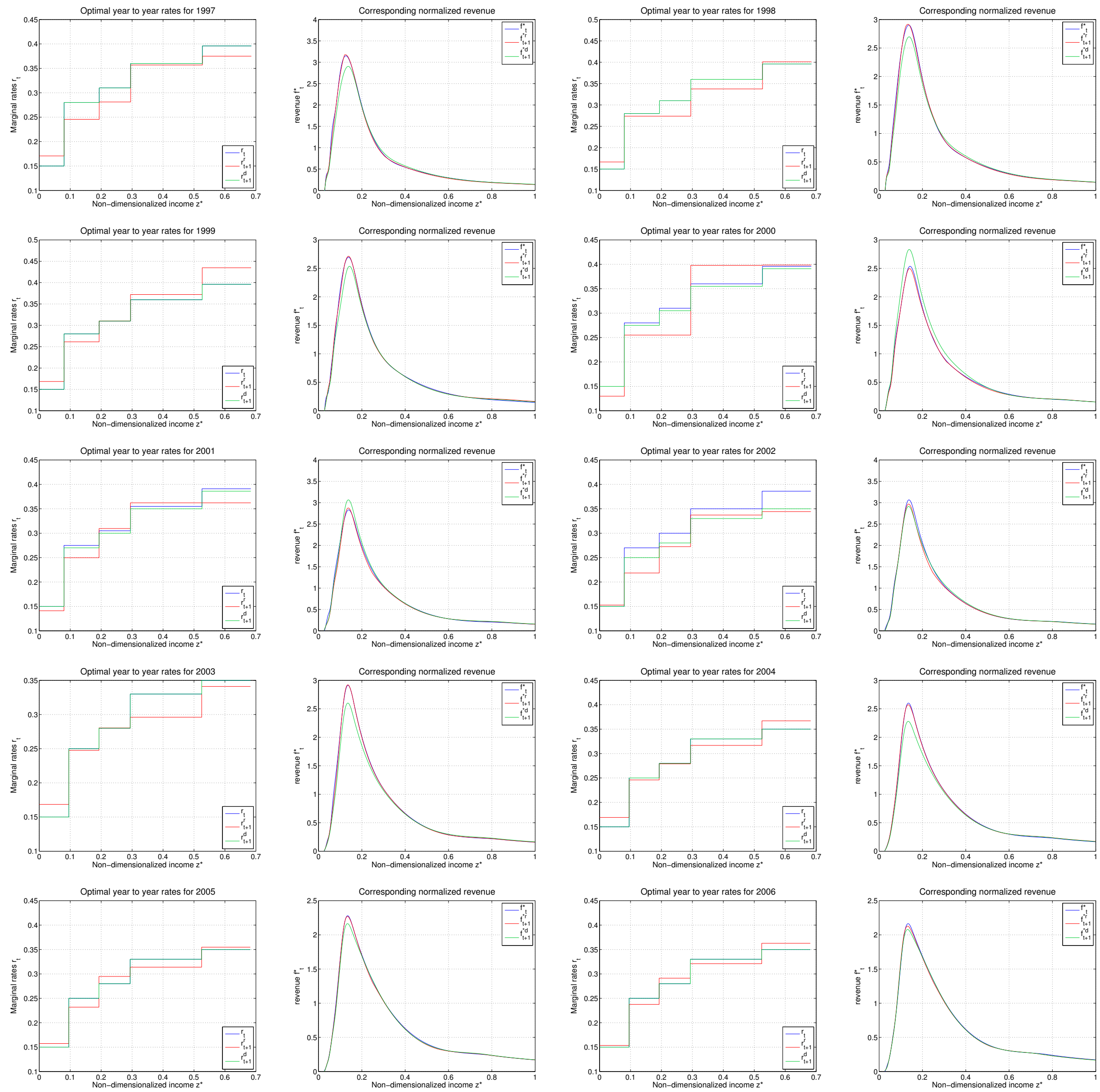



Figure B.4: *optimal marginal rates from 1997 to 2006 and the corresponding normalized revenue.*

### B.2.4 Optimal rates and brackets

Figure B.5 displays the optimal brackets and rates to control for the instability in the revenue distribution. They are compared with rates and brackets from data and the corresponding revenue distributions. Legends, colors and notations are the same as in Figure 4.8, explained in section 4.3 of chapter 4.

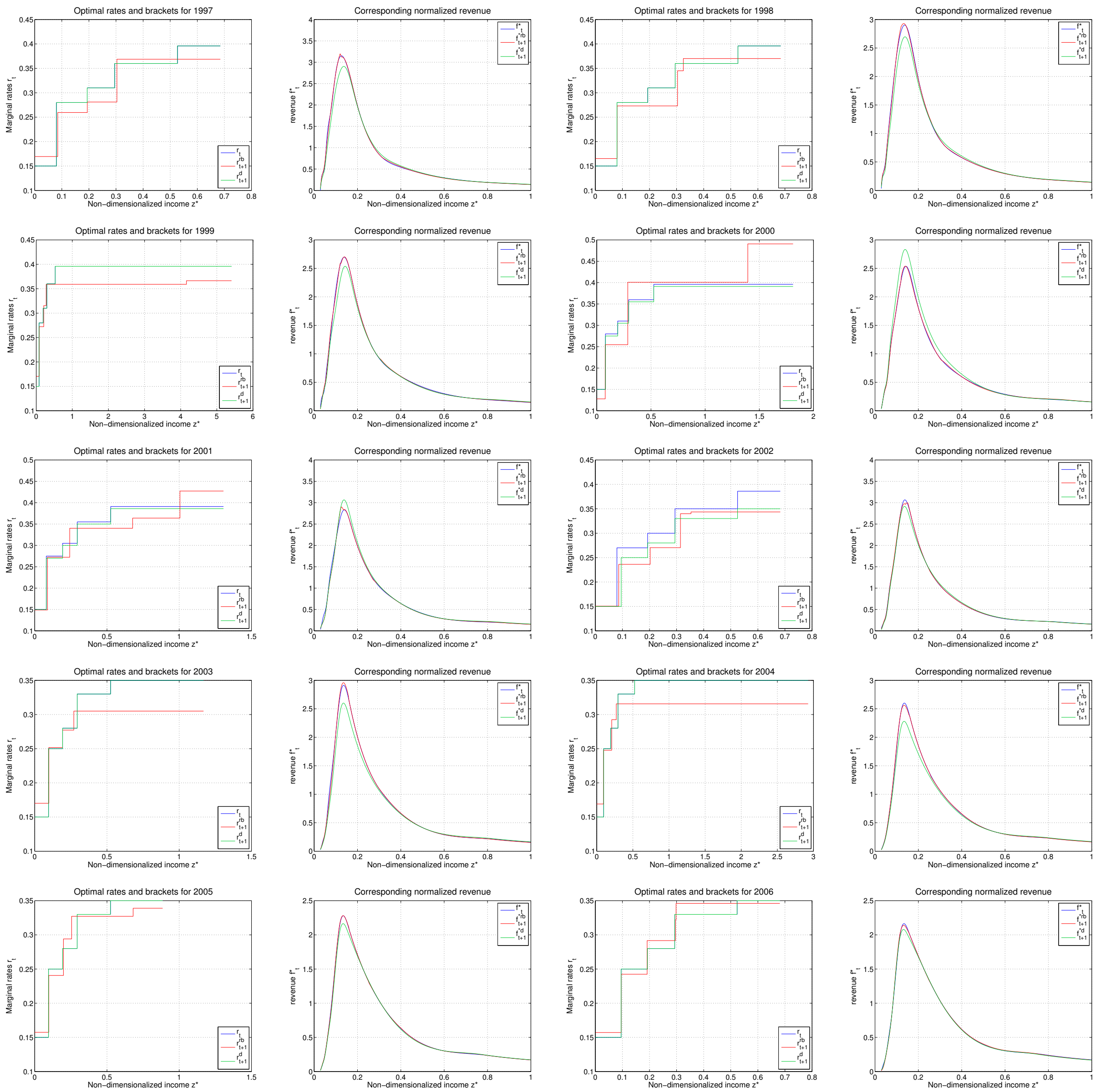



Figure B.5: *optimal marginal rates and brackets from 1997 to 2006 and corresponding normalized revenue.*

## B.3 Model for $g_{t+1}$

Here $g_{t+1}$ is given by (4.21)

### B.3.1 Proof that $T_{t+1} = T_t \frac{g_t}{g_{t+1}}$

For the control of instability, recall,

$$T_{t+1}\,[\tilde{g}_{t+1} + \frac{\partial g}{\partial T}\Delta T] = T_t\, g_t.$$

Since $\Delta T = T_{t+1} - T_t$, we have,

$$\frac{\partial g}{\partial T} T_{t+1}^2 + [\tilde{g}_{t+1} - \frac{\partial g}{\partial T} T_t]\, T_{t+1} - T_t\, g_t = 0. \tag{B.1}$$

This is a quadratic equation in $T_{t+1}$ whose solutions are given by,

$$\begin{cases} T_{t+1} = \dfrac{-1}{2\frac{\partial g}{\partial T}} \left\{ [\tilde{g}_{t+1} - \dfrac{\partial g}{\partial T} T_t] + \sqrt{[\tilde{g}_{t+1} - \dfrac{\partial g}{\partial T} T_t]^2 + 4\, T_t\, g_t \dfrac{\partial g}{\partial T}} \right\}, \\ T_{t+1} = \dfrac{-1}{2\frac{\partial g}{\partial T}} \left\{ [\tilde{g}_{t+1} - \dfrac{\partial g}{\partial T} T_t] - \sqrt{[\tilde{g}_{t+1} - \dfrac{\partial g}{\partial T} T_t]^2 + 4\, T_t\, g_t \dfrac{\partial g}{\partial T}} \right\} \end{cases} \tag{B.2}$$

Let us prove using $g_{t+1}$ of (4.27) that we recover (B.2) by setting $T_{t+1} = T_t \frac{g_t}{g_{t+1}}$. We have,

$$\begin{aligned} \frac{T_t\, g_t}{g_{t+1}} &= 2\, T_t\, g_t / \left\{ [\tilde{g}_{t+1} - \frac{\partial g}{\partial T} T_t] + \sqrt{[\tilde{g}_{t+1} + \frac{\partial g}{\partial T} T_t]^2 + 4\, T_t\, (g_t - \tilde{g}_{t+1}) \frac{\partial g}{\partial T}} \right\} \\ &= 2\, T_t\, g_t \frac{[\tilde{g}_{t+1} - \frac{\partial g}{\partial T} T_t] - \sqrt{[\tilde{g}_{t+1} + \frac{\partial g}{\partial T} T_t]^2 + 4\, T_t\, (g_t - \tilde{g}_{t+1}) \frac{\partial g}{\partial T}}}{[\tilde{g}_{t+1} - \frac{\partial g}{\partial T} T_t]^2 - [\tilde{g}_{t+1} + \frac{\partial g}{\partial T} T_t]^2 - 4\, T_t\, (g_t - \tilde{g}_{t+1}) \frac{\partial g}{\partial T}} \\ &= 2\, T_t\, g_t \frac{[\tilde{g}_{t+1} - \frac{\partial g}{\partial T} T_t] - \sqrt{[\tilde{g}_{t+1} + \frac{\partial g}{\partial T} T_t]^2 + 4\, T_t\, (g_t - \tilde{g}_{t+1}) \frac{\partial g}{\partial T}}}{-4\, \tilde{g}_{t+1} \frac{\partial g}{\partial T} T_t - 4\, T_t\, (g_t - \tilde{g}_{t+1}) \frac{\partial g}{\partial T}} \\ &= \frac{-1}{2\frac{\partial g}{\partial T}} \left\{ [\tilde{g}_{t+1} - \frac{\partial g}{\partial T} T_t] - \sqrt{[\tilde{g}_{t+1} + \frac{\partial g}{\partial T} T_t]^2 + 4\, T_t\, (g_t - \tilde{g}_{t+1}) \frac{\partial g}{\partial T}} \right\}. \end{aligned} \tag{B.3}$$

Now,

$$
\begin{aligned}
[\tilde{g}_{t+1} + \frac{\partial g}{\partial T} T_t]^2 + 4\, T_t\, (g_t - \tilde{g}_{t+1}) \frac{\partial g}{\partial T} &= \tilde{g}_{t+1}^2 + [\frac{\partial g}{\partial T} T_t]^2 + 2\, \tilde{g}_{t+1} \frac{\partial g}{\partial T} T_t \\
&\quad + 4\, T_t\, (g_t - \tilde{g}_{t+1}) \frac{\partial g}{\partial T}, \\
&= \tilde{g}_{t+1}^2 + [\frac{\partial g}{\partial T} T_t]^2 - 2\, \tilde{g}_{t+1} \frac{\partial g}{\partial T} T_t + 4\, T_t\, g_t \frac{\partial g}{\partial T}, \\
&= [\tilde{g}_{t+1} - \frac{\partial g}{\partial T} T_t]^2 + 4\, T_t\, g_t \frac{\partial g}{\partial T}
\end{aligned}
\tag{B.4}
$$

Thus, we have proven that $T_{t+1} = T_t \frac{g_t}{g_{t+1}}$.

### B.3.2 Proof that $\Delta T \longrightarrow T_t \frac{(g_t - \tilde{g}_{t+1})}{\tilde{g}_{t+1}}$ as $\frac{\partial g}{\partial T} \longrightarrow 0$

We have,

$$
\begin{aligned}
\Delta T &= \frac{-1}{2 \frac{\partial g}{\partial T}} \left\{ [\tilde{g}_{t+1} + \frac{\partial g}{\partial T} T_t] - \sqrt{[\tilde{g}_{t+1} + \frac{\partial g}{\partial T} T_t]^2 + 4\, T_t\, (g_t - \tilde{g}_{t+1}) \frac{\partial g}{\partial T}} \right\}, \\
&= \frac{-1}{2 \frac{\partial g}{\partial T}} \left\{ [\tilde{g}_{t+1} + \frac{\partial g}{\partial T} T_t] - [\tilde{g}_{t+1} + \frac{\partial g}{\partial T} T_t] \sqrt{1 + \frac{4\, T_t\, (g_t - \tilde{g}_{t+1}) \frac{\partial g}{\partial T}}{[\tilde{g}_{t+1} + \frac{\partial g}{\partial T} T_t]^2}} \right\}.
\end{aligned}
\tag{B.5}
$$

Using Taylor expansion as $\frac{\partial g}{\partial T} \longrightarrow 0$, we have,

$$
\begin{aligned}
\sqrt{1 + \frac{4\, T_t\, (g_t - \tilde{g}_{t+1}) \frac{\partial g}{\partial T}}{[\tilde{g}_{t+1} + \frac{\partial g}{\partial T} T_t]^2}} &= \left\{ 1 + \frac{4\, T_t\, (g_t - \tilde{g}_{t+1}) \frac{\partial g}{\partial T}}{[\tilde{g}_{t+1} + \frac{\partial g}{\partial T} T_t]^2} \right\}^{1/2}, \\
&= 1 + \frac{1}{2} \frac{4\, T_t\, (g_t - \tilde{g}_{t+1}) \frac{\partial g}{\partial T}}{[\tilde{g}_{t+1} + \frac{\partial g}{\partial T} T_t]^2} + O([\frac{\partial g}{\partial T}]^2).
\end{aligned}
\tag{B.6}
$$

It follows that,

$$
\begin{aligned}
\Delta T &= \frac{-1}{2 \frac{\partial g}{\partial T}} \left\{ \frac{-2\, T_t\, (g_t - \tilde{g}_{t+1}) \frac{\partial g}{\partial T}}{[\tilde{g}_{t+1} + \frac{\partial g}{\partial T} T_t]^2} + O([\frac{\partial g}{\partial T}]) \right\}, \\
&= \frac{T_t\, (g_t - \tilde{g}_{t+1})}{[\tilde{g}_{t+1} + \frac{\partial g}{\partial T} T_t]} + O([\frac{\partial g}{\partial T}]).
\end{aligned}
\tag{B.7}
$$

This proves the result: $\Delta T \longrightarrow T_t \frac{(g_t - \tilde{g}_{t+1})}{\tilde{g}_{t+1}}$ as $\frac{\partial g}{\partial T} \longrightarrow 0$.

## B.4 Results on control with a model on $g_{t+1}$

### B.4.1 Optimal brackets

Figure B.6 displays the optimal brackets to control for the instability in the revenue distribution, with a model on $g_{t+1}$. Legends, colors and notations are the same as in Figure 4.12, explained in section 4.4 of chapter 4. Optimal brackets for $g_{t+1}$ from data of the following year and model are compared with brackets for the current year. The corresponding revenue is displayed on its right. The optimal brackets are more or less on the same range and in general, we achieve a reasonable control of the instability.

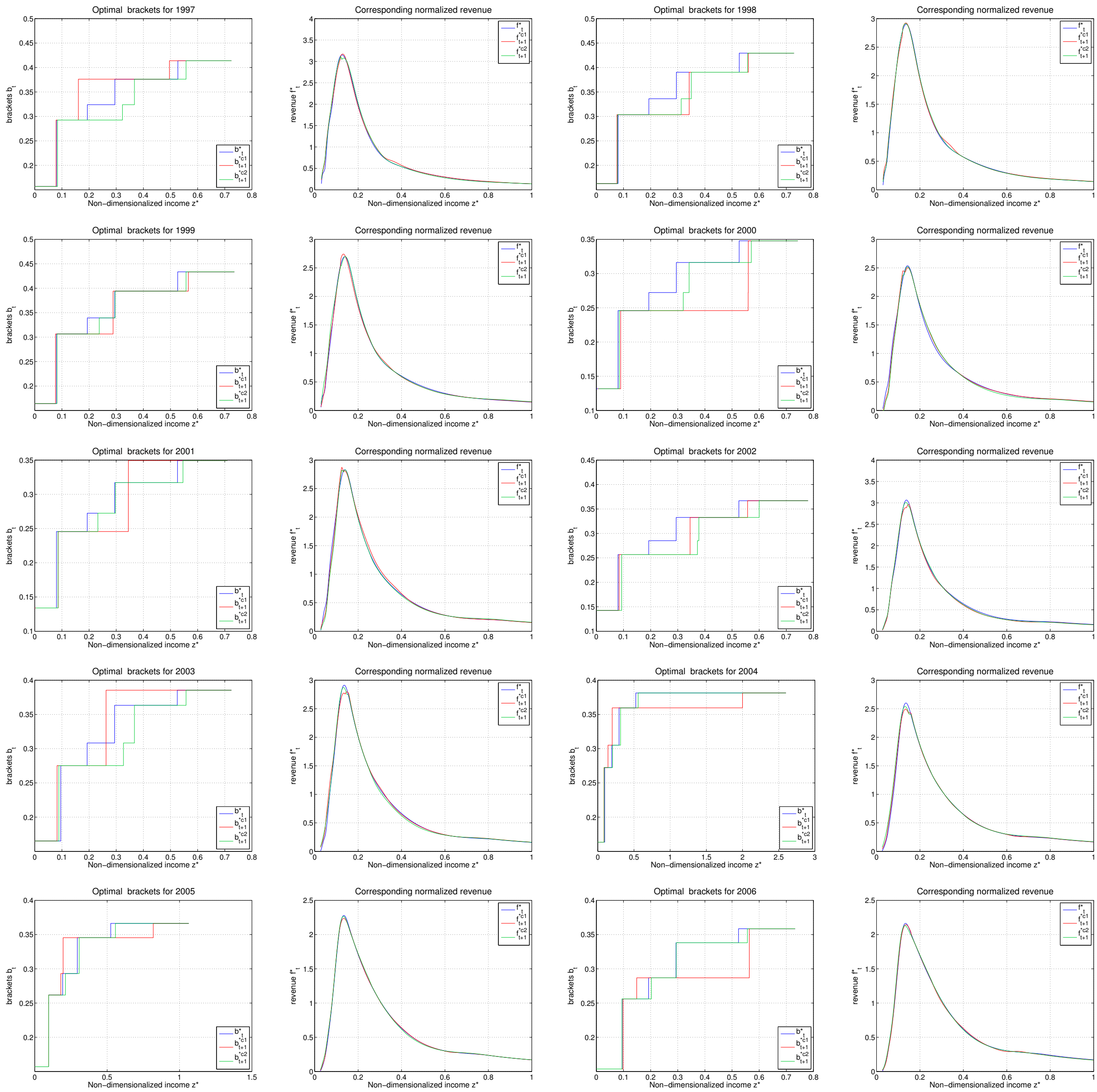



Figure B.6: *Optimal brackets from 1997 to 2006 and the corresponding normalized revenue.*

## B.4.2 Optimal rates

Figure B.7 displays the optimal marginal rates to control for the instability in the revenue distribution, with a model on $g_{t+1}$. Legends, colors and notations are the same as in Figure 4.13, explained in section 4.4 of chapter 4. Optimal rates for $g_{t+1}$ from data of the following year and model are compared with brackets for the current year. The corresponding revenue is displayed on its right. The optimal brackets are more or less on the same range and in general, we achieve a reasonable control of the instability.

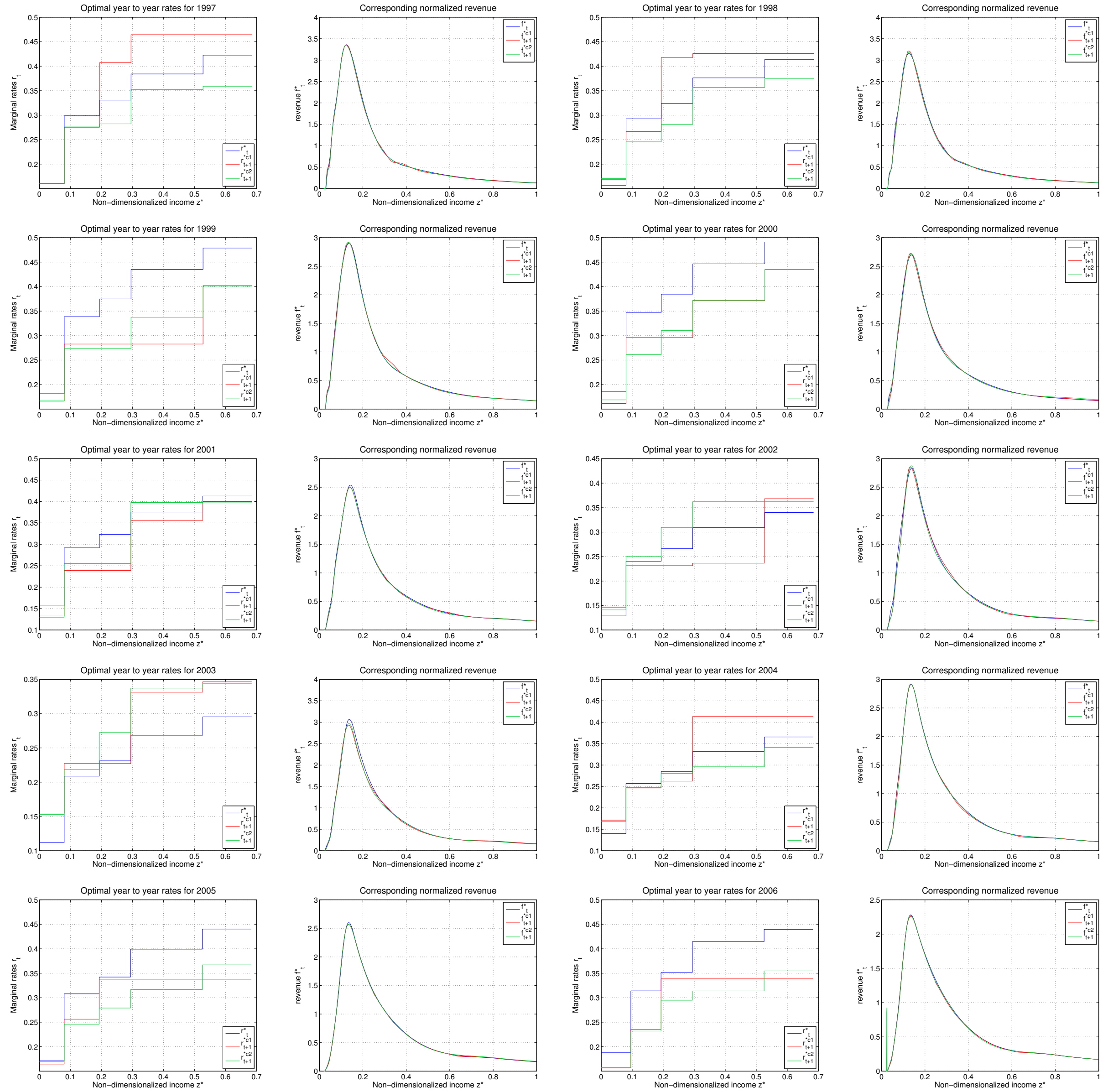



Figure B.7: *optimal marginal rates from 1997 to 2006 and the corresponding normalized revenue.*

### B.4.3 Optimal rates and brackets

Figure B.8 displays the optimal brackets and marginal rates to control for the instability in the revenue distribution, with a model on $g_{t+1}$. Legends, colors and notations are the same as in Figure 4.14, explained in section 4.4 of chapter 4. The corresponding revenue is displayed on the right.

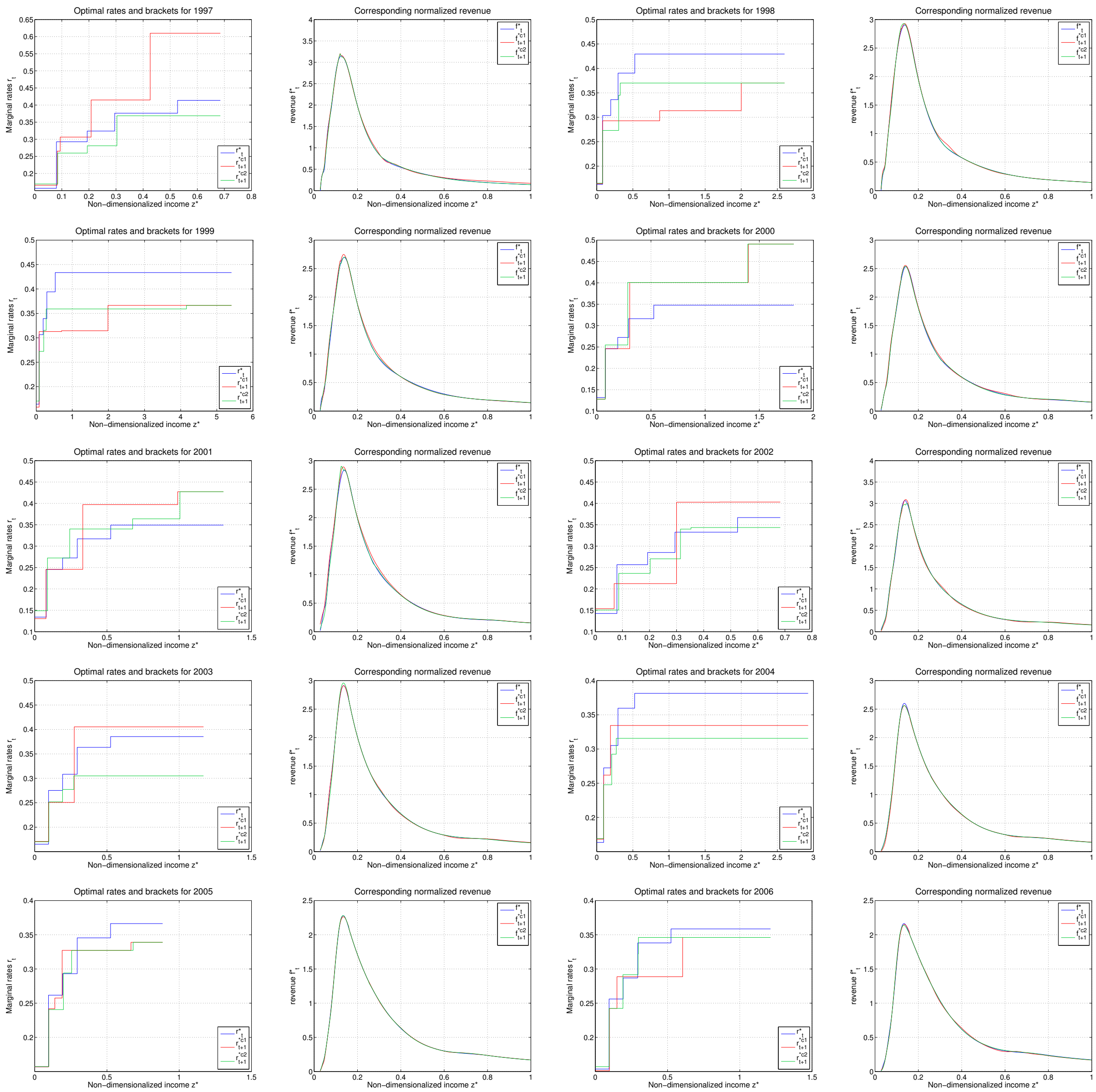



Figure B.8: *optimal marginal rates and brackets from 1997 to 2006 and corresponding normalized revenue.*

# Appendix C

# Control of instability with constant fairness

We display the remaining figures for year to year control of instability with constant fairness as derived in chapter 5 for 1997 to 2008.

## C.1 Control given $g_{t+1}$

we assume $g_{t+1}$ is as given by the data.

### C.1.1 Year to year tax liability control with constant change in fairness

Figure C.1 displays the tax liability required to control for the instability with constant change in fairness. It is compared with actual tax liability from data and tax liability to control the instability with no regard to fairness. Legends, colors and notations are the same as in Figure 5.1, explained in section 5.3 of chapter 5.

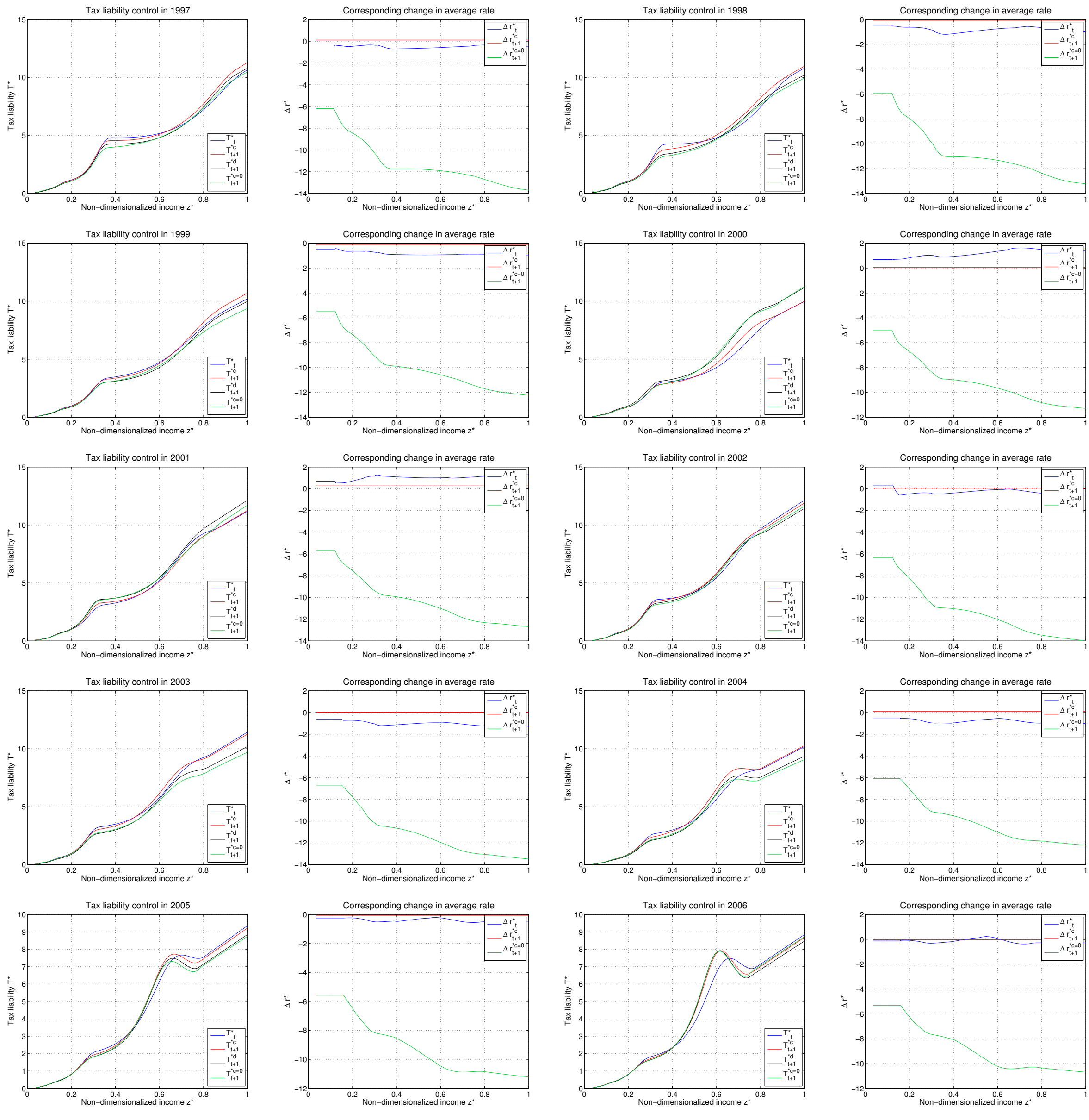



Figure C.1: *Tax liability control and corresponding change in average tax rate from 1997 to 2006.*

## C.1.2 Optimal tax parameters

Figure C.2 displays the optimal tax brackets and marginal rates required to control for the instability with constant change in fairness, year to year from 1997 to 2006. It is compared with actual brackets and rates from data, with optimal rates and brackets to control of the instability with no regard to fairness and with the suboptimal benchmark rates and brackets. The year to year corresponding normalized revenue is displayed as well. Legends, colors and notations are the same as in Figure 5.2, explained in section 5.3.3 of chapter 5.

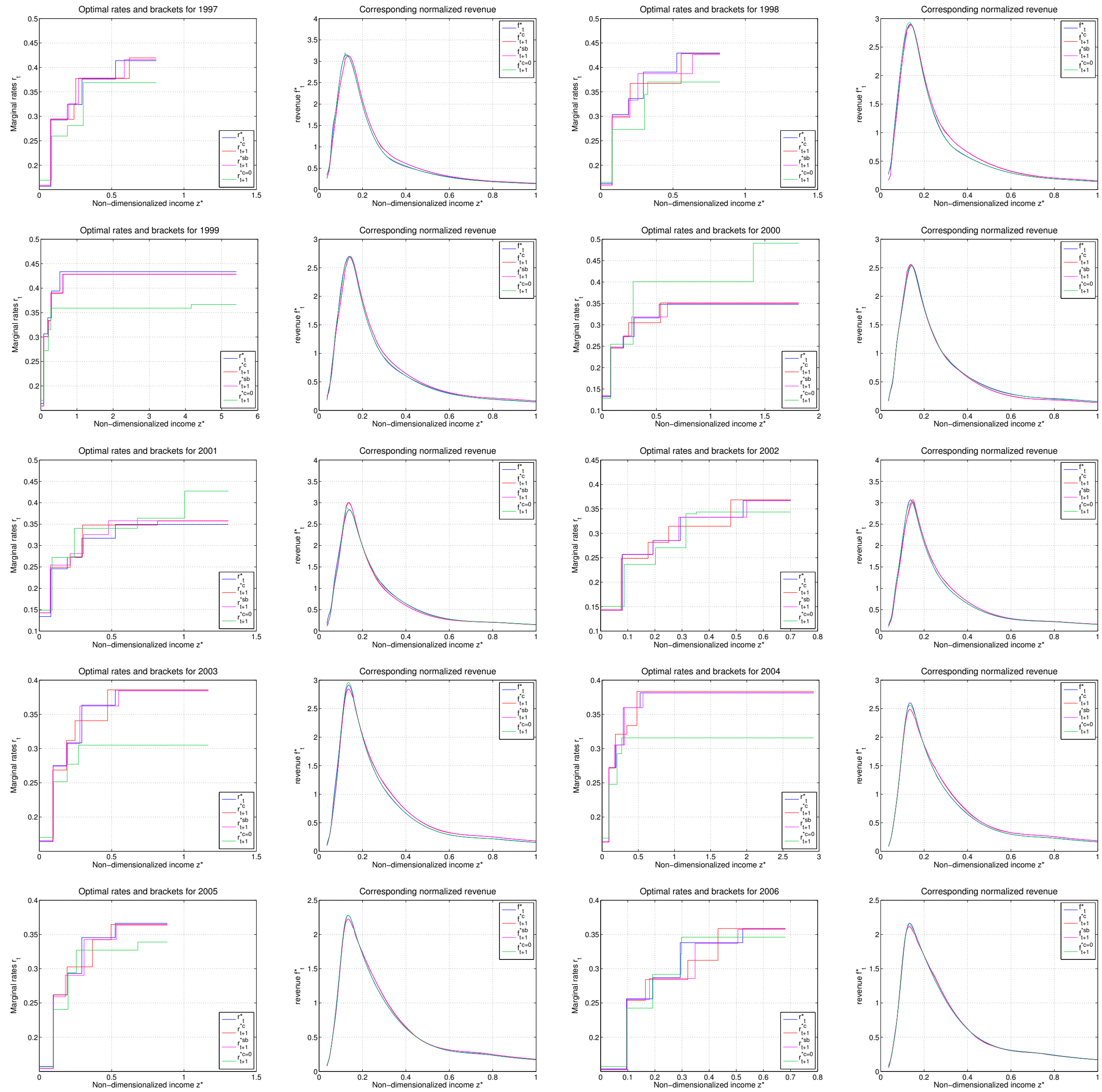



Figure C.2: *Optimal brackets and marginal rates from 1997 to 2006 and the corresponding normalized revenue.*

### C.1.3 Corresponding year to year change in average tax rates

Figure C.3 Changes in average tax rates from 1997 to 2006 when the rates and brackets in Figure C.2 are used. Changes in average rates for the control of instability with constant change in fairness, the corresponding suboptimal parameters, the optimal rates and brackets that control instability with no regard to fairness and the brackets and rates from data are plotted on both the low and high income ranges. Legends, colors and notations are the same as in Figure 5.2, explained in section 5.3.3 of chapter 5.

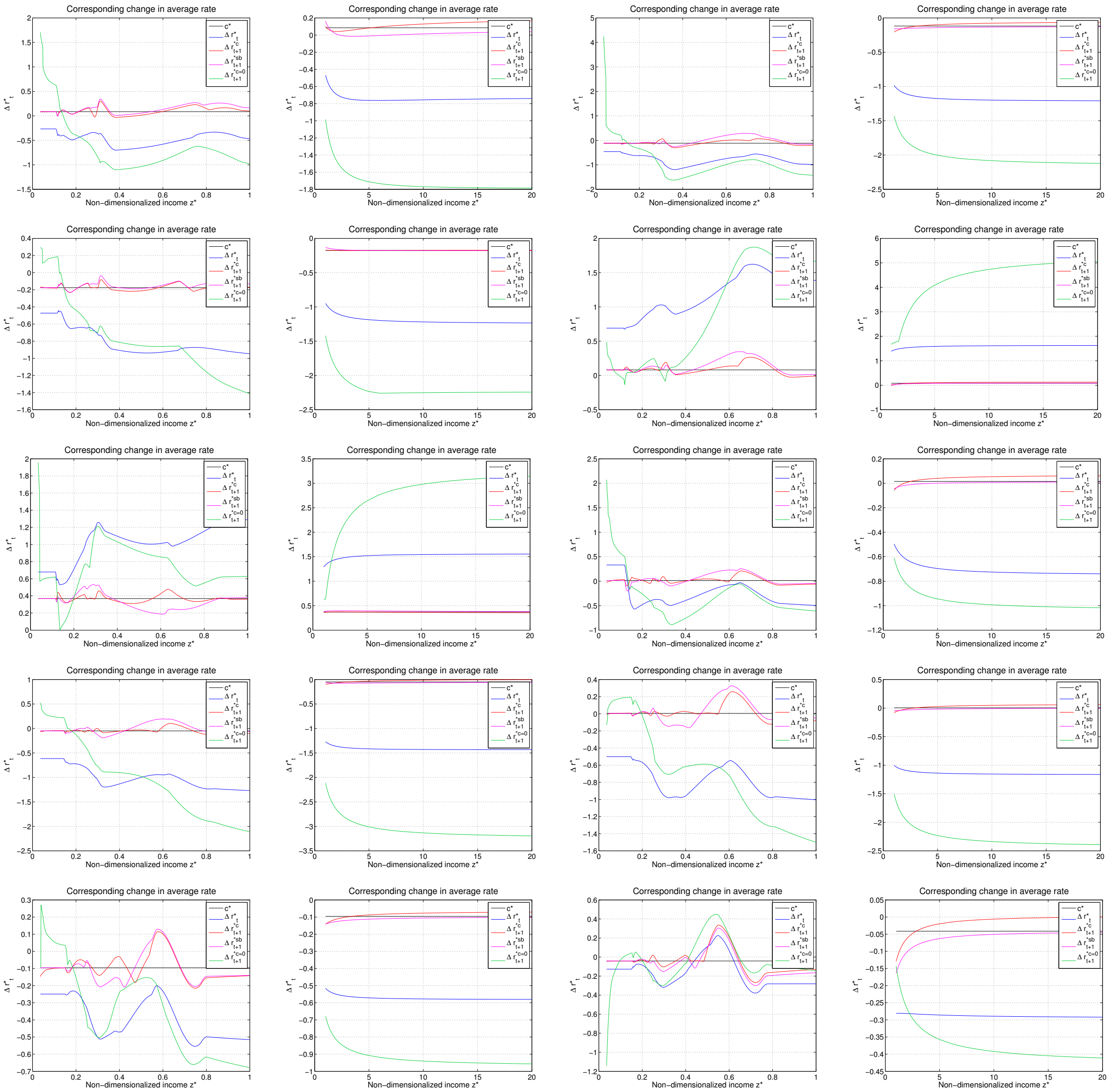



Figure C.3: *Changes in average tax rates from 1997 to 2006 when the rates and brackets in Figure C.2 are used.*

### C.1.4 Uniformity of change in average tax rates

Figure C.4 displays the changes in the average rates as a result of the control of instability with constant change in fairness on the lower income ranges only as well as the control on the upper tail only, from 1997 to 2006. It is compared with changes in the average rates from data and the control of instability with no regard to change in fairness. Legends, colors and notations are the same as in Figure 5.4, explained in section 5.3.4 of chapter 5.

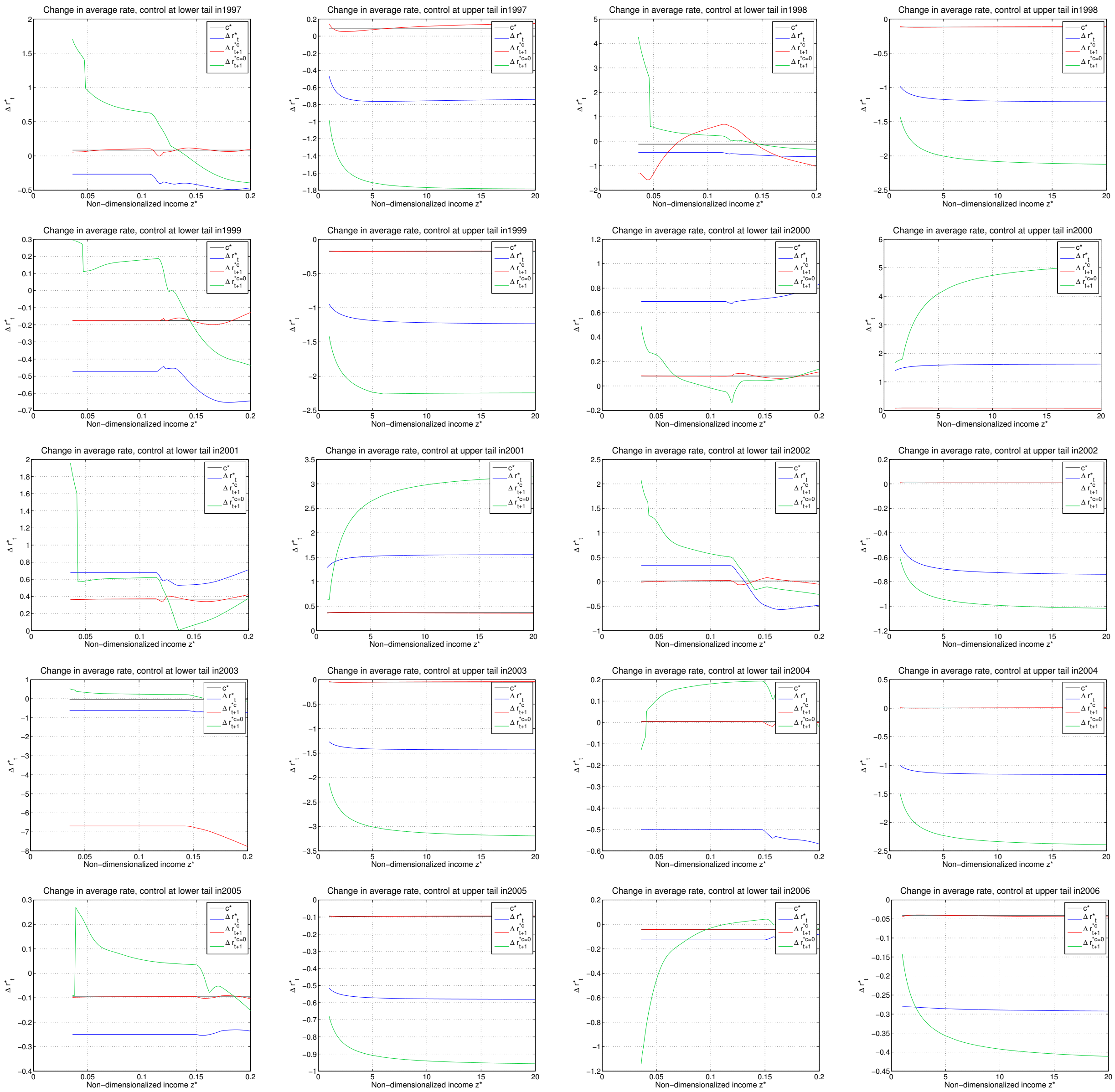



Figure C.4: *Change in average tax rates when controlling for the instability with constant change in fairness on the lower tail only and upper tail only, from 1997 to 2006.*

### C.1.5 Corresponding change in average tax rates with control on both tails

Figure C.5 Changes in average tax rates from 1997 to 2006 when we control of instability with constant change in fairness on both lower and upper tails. It is compared with changes in the average rates from data and the control of instability with no regard to change in fairness. Legends, colors and notations are the same as in Figure 5.4, explained in section 5.3.4 of chapter 5.

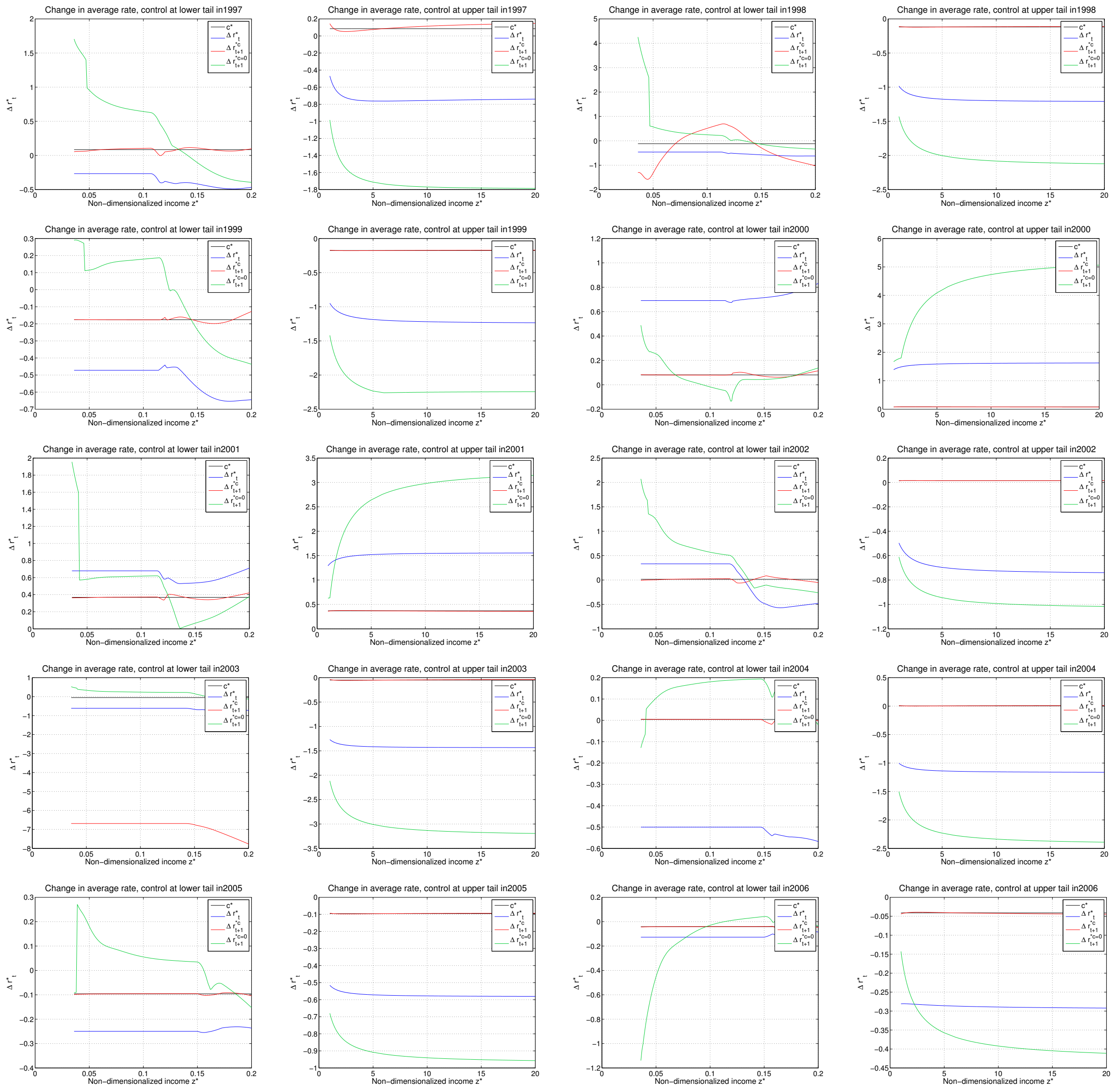



Figure C.5: *Change in average tax rates when controlling for the instability with constant change in fairness on both the lower and upper tail from 1997 to 2006.*

## C.2 Results on control with a model on $g_{t+1}$

### C.2.1 Optimal tax parameters

Figure C.6 displays the optimal tax brackets and marginal rates required to control for the instability with constant change in fairness when the distribution of number of returns depends on the tax liability itself, year to year from 1997 to 2006. It is compared with actual brackets and rates from data, with optimal rates and brackets to control of the instability with no regard to fairness and with the suboptimal benchmark rates and brackets. The year to year corresponding normalized revenue is displayed as well. Legends, colors and notations are the same as in Figure 5.5, explained in section 5.4.1 of chapter 5.

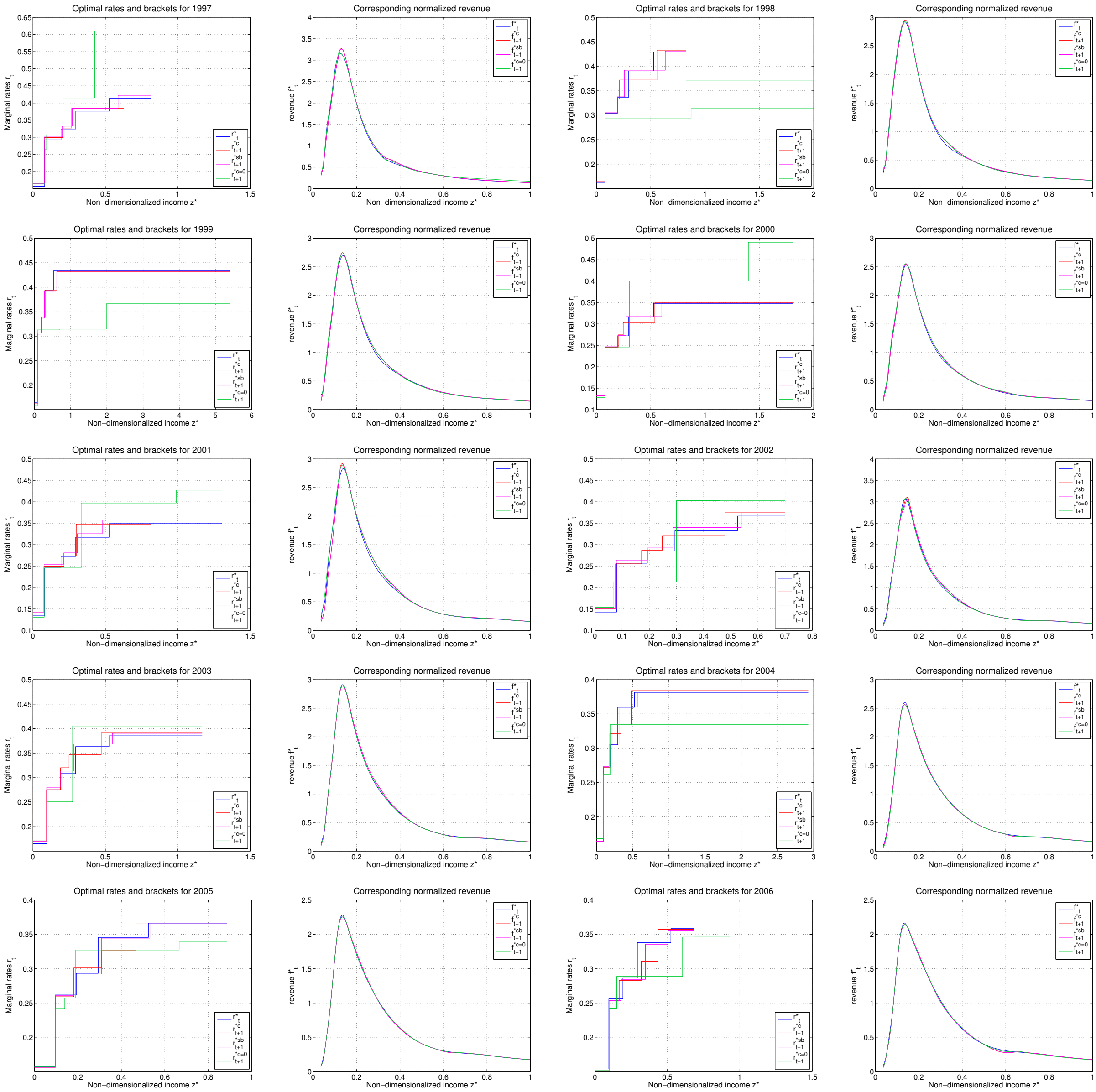



Figure C.6: *Optimal brackets and marginal rates from 1997 to 2006 when we use a model on* $g_{t+1}$ *and the corresponding normalized revenue.*

### C.2.2 Corresponding year to year change in average tax rates

Figure C.7 Changes in average tax rates from 1997 to 2006 when the rates and brackets in Figure C.6 are used. That is, changes in average rates for the control of instability with constant change in fairness where the number of returns now depends on the tax liability itself. In addition, the corresponding suboptimal parameters, the optimal rates and brackets that control instability with no regard to fairness but with a model on $g_{t+1}$ and the brackets and rates from data are plotted on both the low and high income ranges. Legends, colors and notations are the same as in Figure 5.5, explained in section 5.4.1 of chapter 5.

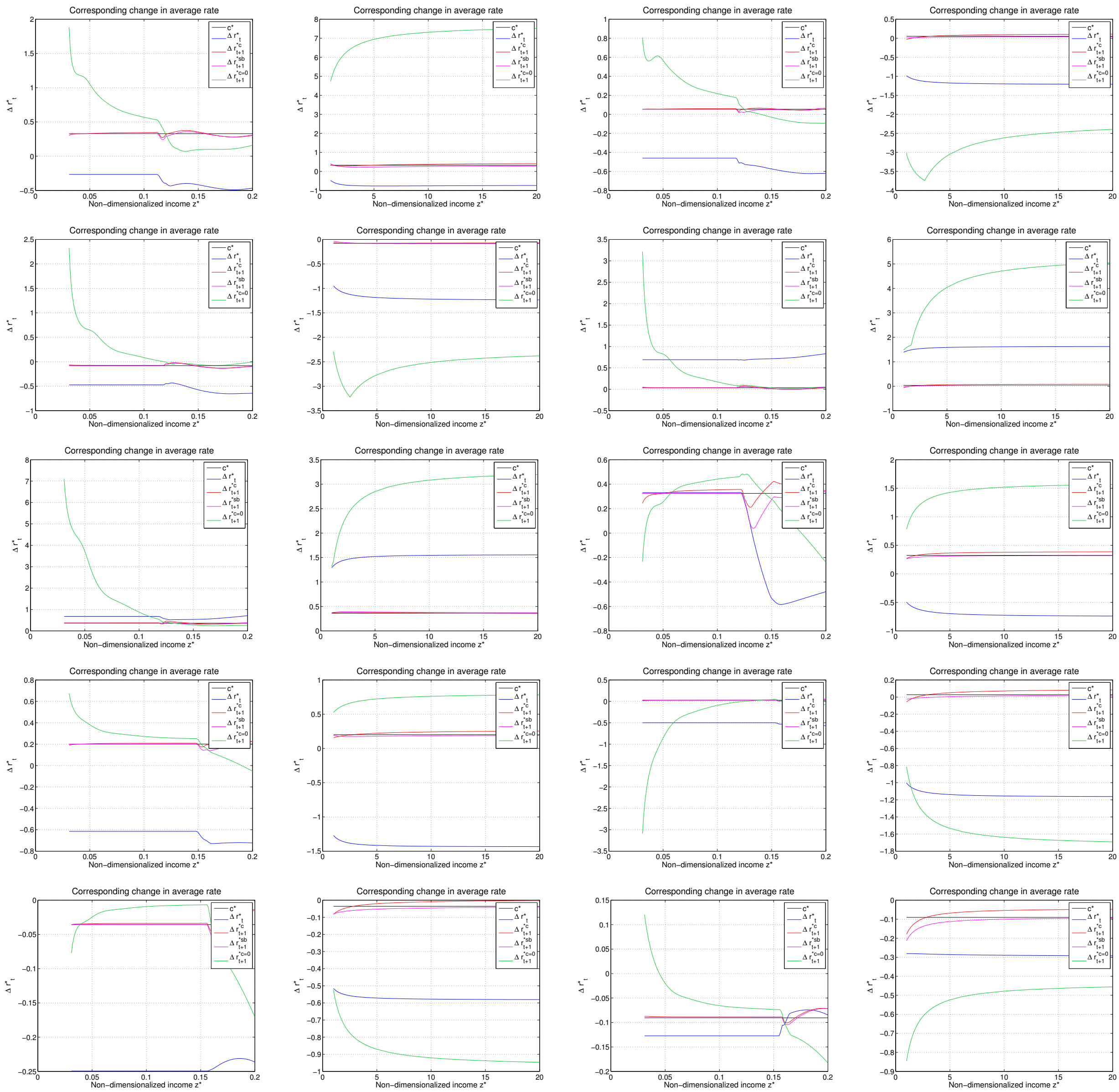


Figure C.7: *Changes in average tax rates from 1997 to 2006 when the rates and brackets in Figure C.6 are used.*

# Bibliography


G. Auten and R. Carroll. The effect of income taxes on household behavior. *Review of Economics and Statistics*, 81(4):681–693, November 1999.

D. Bradford, editor. *Distributional analysis of tax policy*. AEI Press, Washington, D.C, 1995.

H. Buddelmeyer, J. Creedy, and G. Kalb. *Tax policy design and behavioural microsimulation modelling*. Edward Elgar, 2007. ISBN 9781845429140.

E. Castillo and A. S. Haldi. A method for estimating parameters and quantiles of distributions of continuous random variables. *Computational Statistics and Data Analysis*, 20:421–439, 1995.

CBO. Historical effective tax rates: 1979-2005. Congressional Budget Office Publication, 2007a. URL http://www.cbo.gov/publication/41654.

CBO. The effect of tax changes on labor supply in cboâĂŹs microsimulation tax model. Congressional Budget Office Background Paper, Pub. No 2756, 2007b. URL http://www.cbo.gov/sites/default/files/cbofiles/ftpdocs/79xx/doc7996/04-12-laborsup

E. K. P. Chong and S. H. Zak. *Introduction to optimization*. John Wiley & Sons, 2001.

J. Creedy and A. Duncan. Behavioural microsimulation model with labour supply responses. *JOURNAL OF ECONOMIC SURVEYS*, 16(1), 2002.

J. Creedy and N. Gemmell. *Modelling Tax Revenue Growth*. Edward Elgar, 2006.

J. Creedy and N. Herault. Optimal marginal income tax reforms: A microsimulation analysis. Melbourne Institute Working Paper Series wp2009n23, Melbourne Institute of Applied Economic and Social Research, The University of Melbourne, Aug. 2009.

EGTRRA Law. Economic growth and tax relief reconciliation act of 2001,. $107^{th}$ Congress, Public Law 107 - 16, 2001. URL http://www.gpo.gov/fdsys/pkg/PLAW-107publ16/content-detail.html.

M. Feldstein. The effect of marginal tax rates on taxable income: A panel study of the 1986 tax reform act. *Journal of Political Economy*, 103(3):551–572, June 1995.

M. Feldstein. Tax avoidance and the deadweight loss of the income tax. *The Review of Economics and Statistics*, 81(4):674–680, 1999.

GAO. Understanding the tax reform debate. Government Accountability Office Publication GAO-05-1009SP, 2005. URL http://www.gao.gov/new.items/d051009sp.pdf.

S. Giertz. Panel data techniques and the elasticity of taxable income. CBO Working Paper 2008-20, 2008. Congressional Budget Office, Washington D.C.

A. Goolsbee. What happens when you tax the rich? evidence from executive compensation. *Journal of Political Economy*, 108(2):352–378, 2000.

J. Gruber and E. Saez. The elasticity of the taxable income: Evidence and implications. *Journal of Public Economics*, 84(1):1–32, April 2002.

A. C. Harberger. Principles of efficiency: The measurement of waste. *American Economic Review*, 54(3):58–76, 1964.

S. Howison. *Practical Applied Mathematics: Modelling, Analysis, Approximation*, chapter 2,3. Cambridge University Press, 2005.

JCT. Discussion of revenue estimation methodology and process. Joint Committee on Taxation, JCS-14-92, 1992. URL http://www.jct.gov/publications.html?func=download&id=1185&chk=1185&no_html=1.

R. Kanbur, M. Keen, and M. Tuomala. Optimal non-linear income taxation for the alleviation of income-poverty. *European Economic Review*, 38:1613 – 1632, 1994.

R. Kanbur, J. Pirttila, and Ì. Tuomala. Non-welfarist optimal taxation and behavioural public economics. *JOURNAL OF ECONOMIC SURVEYS Vol. 20, No. 5*, 20(5):849–868, 2006.

W. Kopczuk. Tax bases, tax rates and the elasticity of reported income. Working paper 10044, National Bureau of Economic Research, October 2003. URL http://www.nber.org/papers/w10044.

A. B. Laffer. The ellipse: An explication of the laffer curve in a two factor model. *A. B. Laffer Associates*, July 1980.

A. B. Laffer. The laffer curve: Past, present and future. *The Heritage Foundation, Backgrounder, Executive Summary*, (1765), June 2004.

P. M. Lee. *Bayesian Statistics: An Introduction*. Wiley, 3rd edition, 2004.

L. Lindsey. Individual taxpayer response to tax cuts:1982-1984, with implications for the revenue maximizing tax rate. *Journal of Public Economics*, 33(2):173–206, 1987.

A. M. Lyapunov. *Stability of Motion*. Academic Press, New York, 1892.

J. A. Mirrlees. An exploration in the theory of optimal income taxation. *Review of Economic Studies*, 38:175–208, 1971.

R. A. Musgrave. Et, ot and sbt. *Journal of Public Economics*, 6:5–18, 1976.

J. Navratil. The economic recovery tax act of 1981: evidence on individual taxpayer behavior from panel tax return data. PhD. thesis Harvard University, 1995.

OTA. Us treasury distributional analysis methodology. Office of Tax Analysis, OTA Paper 85, 1999. URL http://www.treasury.gov/resource-center/tax-policy/tax-analysis/Documents/

OTA. Treasury's panel model for tax analysis. Office of Tax Analysis, OTA Technical Working Paper 3, 2008. URL http://www.treasury.gov/resource-center/tax-policy/tax-analysis/Documents/ota85.pdf

R. Peck, C. Olsen, and J. Devore. *Introduction to statistics and data analysis*. Brooks/Cole, 2001.

J. Pencavel. Labor supply of men. *Handbook of Labor Economics*, pages 3–102, 1986.

J. Pirttila and M. Tuomala. Poverty alleviation and tax policy. *European Economic Review*, 48: 1075–1090, 2004.

R. Plato. *Concise Numerical Mathematics*, chapter 2, pages 23–33. American Mathematical Society, 2003.

R. E. Quandt. Old and new methods of estimation and the pareto distribution. *Metrika*, 10: 55–82, 1966.

M. Rahman and L. M. Pearson. A note on estimating parameters in two-parameter pareto distributions. *International Journal of Mathematics Education in Science and Technology*, 34(2): 298–306, 2003.

E. Sadka. On income distribution, incentive effects and optimal income taxation. *Review of Economic Studies*, 43(2):261–268, 1976.

E. Saez. Using elasticities to drive optimal income tax rates. *Review of Economic Studies*, 68: 205–229, 2001.

E. Saez. The effect of marginal tax rates on income: A panel study of "bracket creep". *Journal of Public Economics*, 87:1231–1258, 2003.

E. Saez. Reported incomes and marginal tax rates, 1960-2000: Evidence and policy implications. Working paper 10273, National Bureau of Economic Research, January 2004. URL http://www.nber.org/papers/w10273.

F. Sammartino and D. Weiner. Recent evidence on taxpayer responses to the rate increases in the 1990s. *National Tax Journal*, 50(3):683–705, 1997.

J. Seade. Optimal non-linear policies for non-utilitarian motives. *Income Distribution: The Limits to Redistribution. Bristol: Scientechnica.*, 1980.

J. K. Seade. On the shape of optimal tax schedules. *Journal of Public Economics*, 7(2):203–236, 1977.

J. K. Seade. On the sign of the optimum marginal income tax. *Review of Economic Studies*, 49 (4):637–643, 1982.

E. Sheshinski. Note on the shape of the optimum income tax schedule. *Journal of Public Economics*, 40:201–215, 1989.

J. Slemrod. The optimal two-bracket linear income tax. *Journal of Public Economics*, 53:269–290, 1994.

Statistics of Income. Individual income tax returns. Publication 1304 (Rev. 07-2010), Internal Revenue Service, Washington DC, 2008. URL http://www.irs.gov/taxstats/indtaxstats/article/0,,id=134951,00.html.

N. H. Stern. On the specification of models of optimum income taxation. *Journal of Public Economics*, 6:123–162, 1976.

M. Tuomala. Simplified formulae for optimal linear income taxation. *Scandinavian Journal of Economics*, 87(4):668–672, 1985.

M. Tuomala. *Optimal Income tax and Redistribution.* Oxford: Clarendon Press, 1990.

Y. Wang. *Smoothing Splines: Methods and Applications.* Monographs on statistics and applied probability. Taylor and Francis, 2010.

C. Webber. Obtaining a consistent estimate of the elasticity of taxable income using difference-in-differences. PhD Thesis, University of Michigan, 2011.

M. Weber. The statistics of income 1979-2002 continuous work history sample individual income tax return panel. Statistics of Income Division, Internal Revenue Service, P.O. Box 2608, Washington, DC 20013-2608, 2002.

H. P. Young. Progressive taxation and equal sacrifice. *American Economic Review*, 80(1):253–66, 1990.