Essays on Redistribution

A THESIS

SUBMITTED TO THE FACULTY OF THE GRADUATE SCHOOL

OF THE UNIVERSITY OF MINNESOTA

BY

Daniel Samano-Penaloza

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS

FOR THE DEGREE OF

Doctor Of Philosophy

Narayana R. Kocherlakota, Adviser

June, 2009

Acknowledgements

First, I would like to thank my adviser, Narayana R. Kocherlakota for his support and guidance through all stages of this dissertation. Right from the beginning I knew I was in hands of a great economist. Very soon after, I realized that I was also in hands of a great human being. This dissertation would not have been possible without the insight, effort, dedication and professionalism of Narayana. Thank you, Narayana; I hope one day I can do for you just a bit of what you have done for me.

Secondly, I would like to thank Chris Phelan and Larry Jones for being part of my dissertation committee. This work has benefited greatly from the ideas and suggestions of both of them. Thank you for being part of the project. I am grateful to Professor Songying Fang from the Political Science Department for serving as my external examiner. I would also like to acknowledge the genuine interest of Aldo Rustichini in the topic of this dissertation and the interesting conversations I had with him. All errors and omissions are mine. I am indebted to Daniel Feenberg from the NBER and the Economic and Social Data Service at the University of Essex for generously providing access to the SOI and the SPI databases, respectively.

I would like to thank Professor Simran Sahi for having assigned me interesting teaching opportunities over all these years. I am also indebted to the amazing staff of the Department of Economics for their service. I am particularly grateful to Caty Bach, Kara Kersteter and Wendy Williamson. I also would like to thank CONACYT for the financial support.

I would like to thank my friends at Minnesota. Without them, my years as a graduate student would have being much tougher. In particular, I would like to thank Santiago Bazdresch, Juan Cardenas, Mafer Cordova, Julian Diaz, Jonathan Gomez Raja, Facundo Piguillem, Andres Malcolm, Carla Manzoni, Marina Mendes-Tavares, Veronica Michel, Blanca Monter, Christoph Mueller, Belen Naranjo, Miguel Ricaurte, Anderson Schneider, Rene Schwengber, Mario Solis, Donaji Valencia, Lucia Zacchi and the Whitehouse family. Thank you guys for all the moments of joy. I am sure there will be more of them in the future. I also would like to thank Ina Simonovska for walking very close to me throughout these years, particularly this last year.

Next, I would like to thank Alberto Torres for being a great mentor and friend over all these years. Thank you for trusting me. I am sure I am joining a great team. Also, thanks to Manuel Ramos-Francia for giving me this opportunity.

Agradezco a la familia Vega Sámano, a la familia Sámano García, a la familia Sámano Delgado y a la familia Velázquez Peñaloza por su gran afecto. Gracias por siempre haber sido un gran respaldo.

Finalmente quiero agradecer a mi familia por su apoyo inmenso en todos estos años. En reciprocidad, esta tesis va dedicada a ustedes. Gracias a mis padres Alicia y Agustín por la dedicación que siempre tuvieron en mi formación como humano. A mis abuelos, Clementina y Tomás, por todo su cariño y amor. Gracias a mis hermanos, Adriana y Agustín por hacer mi vida mas feliz.

Contents

List of Tables

List of Figures

Chapter 1

Introduction

Optimal tax theory has difficulty rationalizing high marginal tax rates at the upper end of the labor income distribution. The work of Mirrlees (1971) provides a strong reason against high marginal labor income taxes at the high end of the income distribution: they reduce the labor supply of highly productive individuals. Nevertheless, in reality most countries impose high marginal tax rates on affluent individuals. Recent figures show that the highest marginal tax rates in Belgium, Japan and Sweden are close to 50%, in Australia, China, France, Germany and Italy are close to 45% while in Brazil, India and the United States are around 30%.[1] Can we reconcile theory with reality? Diamond (1998) and Saez (2001) obtain positive asymptotic marginal tax rates under the assumption that the labor earnings distribution's upper tail is distributed according to a Pareto distribution. However, if such condition fails to be satisfied, their result no

[1] See the article "Taxing the Rich" by The Economist with information by the KPMG.

longer holds.[2]

In Chapter 2, I provide an alternative explanation that does not depend crucially on the functional form describing the tail of the earnings distribution. With this purpose, I construct a model of optimal income taxation in which, by introducing a novel preference formulation, agents display jealousy toward the consumption of others. I derive a simple expression for optimal taxes that accommodates consumption externalities within the Mirrlees (1971) framework. Using this expression, I conduct a positive analysis of taxation. Assuming observed labor income taxes in the United States and the United Kingdom are optimal, I calculate the consumption externalities that rationalize these tax schedules in the light of my model. I show that only a *moderate* amount of jealousy toward the rich is sufficient to rationalize the observed labor income taxes in the aforementioned countries. This is the main contribution of this essay. My estimations suggest that the progressivity of actual tax schedules may be highly driven by corrective considerations, particularly at the top of the earnings distribution.

In the model of Chapter 2 there is a single consumption good and leisure. Thus, there is only one margin to distort for redistributive and corrective purposes; namely, the marginal rate of substitution between consumption and leisure. In Chapter 3, I construct a model that considers two types of goods, in addition to leisure, that differ in the magnitude of consumption externalities they generate. In other words, the two goods differ in their degree of positionality, where a good is said to be *positional* if it is

[2] In Appendix A.2 I show direct evidence that the upper tail of the gross income distribution in the United States is not well approximated by the aforementioned distribution at very high income levels.

valued relative to other agents' consumption of the same good. I parametrize positional considerations and analyze optimal tax policy. This is the normative analysis that I conduct in Chapter 3.

I show that optimal allocations in this environment may be implemented by a *linear* tax on the positional good in combination with a non-linear marginal income tax schedule with standard Mirrleesian properties, namely, zero taxation at extremes. Thus, the two instruments are needed for redistributive and corrective purposes. Under reasonable parameters describing the magnitude of positional considerations, my numerical calculations show that the flat tax imposed on the consumption of the positional good is by no means negligible. Moreover, while my calculations suggest that a flat tax on the positional good does almost as well as a non-linear one in terms of aggregate welfare, large distributional effects are induced by the former.

Finally, Chapter 4 concludes.

Chapter 2

Explaining Taxes on the Rich: The Role of Jealousy

> I must confess that I had expected the rigorous analysis of income-taxation in the utilitarian manner to provide an argument for high taxes. It has not done so. [Sir James A. Mirrlees, *"An Exploration in the Theory of Optimum Income Taxation"*]

2.1 Introduction

The tradeoff between redistribution and incentives has been analyzed extensively by economists. A very important lesson extracted from these studies is that if the government redistributes excessively from rich to poor, highly productive individuals are left with little or no incentives to work. In order to avoid this, consumption and income inequalities arise as a consequence of incentive problems.

Mirrlees (1971) was the first study that formulated this problem in a rigorous fashion. He finds that despite inequality aversion considerations by the government, low marginal labor income taxes on the rich are desirable. Other studies, such as Sadka (1976) and Seade (1977) refined one of Mirrlees (1971)'s assumptions and obtained the well known result that the tax rate at the top of the distribution must be zero. These results may be surprising, but the intuition behind them is quite clear: a high marginal income tax induces leisure of highly productive individuals. This is very costly for the economy as a whole in terms of forgone output. Thus, low marginal taxes at the high end of the income distribution are optimal despite redistributional considerations.

Having such a strong theoretical argument against, why is it that in most of the countries marginal taxes on affluent people are high? One explanation is provided by Diamond (1998) and Saez (2001). They obtain positive asymptotic marginal taxes when the functional form describing the labor earnings distribution's upper tail is Pareto. However, the result no longer holds if the previous condition fails to be satisfied.

In this paper, I provide an alternative explanation: taxation on the rich occurs in order to correct consumption externalities. I show that only a *moderate* amount of jealousy toward the rich is sufficient to rationalize the observed labor income taxes in the United States and the United Kingdom. This is the main contribution of this paper.

To carry out my analysis, I model an economy populated by a continuum of agents with heterogeneous privately-known productivities. Using a novel formulation of preferences, agents display jealousy over the consumption of others. Thus, agents' consumption generates an externality in the form of a consumption benchmark. This introduces an additional reason for government intervention, namely, taxing income for corrective purposes. This occurs in absence of a non-linear consumption tax. I derive a simple expression for optimal taxes that accommodates consumption externalities within the Mirrlees (1971) framework. This expression decomposes the observed tax schedule into two components: the Mirrleesian and the Pigouvian tax. I then show that only a *moderate* amount of jealousy toward the rich is sufficient to rationalize the observed labor income taxes in the United States and the United Kingdom as Pigouvian.

If actual fiscal policy is supposed to be influenced by positional concerns, we need to gather empirical evidence in this regard. In this respect, there are a few studies that have attempted to measure the degree to which relative consumption matters for people's satisfaction. Surveying individuals about their choice among hypothetical worlds they could live in is one approach.[1] In world A, the assets of the subject are higher than in world B. However, agents are worse off in world A than in world B with respect to the population average. Thus, individuals' choices end up revealing their concern for relative positions. For the sake of concreteness, most surveys focus on particular assets such as cars, houses and leisure. In some cases, they also

[1] To be more precise, most of those surveys ask individuals where they would like an imagined future relative of them to live in. According to Alpizar, Carlsson, and Johansson-Stenman (2005), this in order to liberate them from current circumstances.

include income. Using a survey applied to a representative sample of the Swedish population, Carlsson, Johansson-Stenman, and Martinsson (2007) find evidence that supports the relative consumption hypothesis for income and cars but not for leisure.[2] Alpizar, Carlsson, and Johansson-Stenman (2005) survey students from Costa Rica and obtain similar results. J. Solnick and Hemenway (1998) survey a sample of American students and find that about 50% of them would prefer a world in which they had half their absolute income as long as their relative standing was high.

An alternative and more common approach involves applying a regression analysis. Luttmer (2005) merges a database on individuals' self reported happiness to information about local (geographically speaking) average earnings and finds that self reported happiness is negatively affected by the earnings of others in their area. Using data on British workers job satisfaction, Clark and Oswald (1996) construct reference groups that comprise individuals with the same labor market characteristics such as age, education, sex, monthly earnings and hour per week worked. They find that people are less satisfied with their jobs the higher the income of their reference group is. On social comparisons among family members, Neumark and Postlewaite (1998) find that a woman outside the formal labor force is 16-25% more likely to work outside the home if her sister's husband earns more than her own husband. At an experimental level, Rustichini and Vostroknutov (2007) conduct a "burning money" game and find that individuals are willing to incur a cost in order to reduce the winnings of others. This

[2] Some of the results of this paper are striking. For instance, they find that about 50% of the utility obtained from cars and income comes from relative concerns.

occurs mostly when the winnings result from skill rather than luck.

Regarding studies of optimal taxation under consumption externalities, the work of Ljungqvist and Uhlig (2000) presents a complete markets dynamic economy driven by productivity shocks. The negative externality they model is known as external habit formation.[3] That is, when agents increase their consumption, they do not take into account their effect over the aggregate desire of other agents to catch up. Nevertheless, since the consumption distribution is degenerate in their model, the policy implications that can be extracted are to prevent consumption addiction, not jealousy itself. To the best of my knowledge, the only studies prior to mine regarding optimal income taxation à la Mirrlees under utility interdependence are Tuomala (1990) and Oswald (1983). However, both articles conduct a normative analysis of taxation under the assumption that agents value their consumption relative to the average consumption. The numerical calculations of the former show that optimal income taxes are progressive and that higher overall taxes correspond to a higher degree of jealousy in agents' preferences. The second article highlights the fact that the zero marginal taxes at the extremes of the skills distribution no longer holds under utility interdependence. In contrast to my work, neither article attempts to rationalize observed tax schedules.

The rest of this paper proceeds as follows: section 2.2 presents the model, section 2.3 describes the data and estimation procedure, section 2.4 shows and discusses the results and section 2.5 concludes.

[3] External habit formation was first introduce in the finance literature in order to explain the equity premium puzzle. See Constantinides (1990) and Heaton (1995).

2.2 The Model

Consider a static economy populated by a continuum of agents with heterogeneous productivity or skill. Let $\theta \in \Theta$, where $\Theta \equiv [\underline{\theta}, \bar{\theta}]$ and $0 < \underline{\theta} < \bar{\theta} < \infty$, be individual's productivity distributed according to the density $f : \Theta \to \mathbf{R}_{++}$. Productivity is *privately* known to each agent. An agent with productivity θ has a utility function separable in "consumption" and leisure of the form

$$U(c, y, C; \theta) = u(c, C) - v\left(\frac{y}{\theta}\right)$$

where c, y are individual's consumption and effective labor, respectively.[4] Moreover,

$$C \equiv \int_{\Theta} c(\theta)\psi(\theta)d\theta \tag{2.1}$$

is the society's consumption benchmark specified as a *weighted* average of consumption. As usual, preferences satisfy $u_c > 0$, $u_{cc} \leq 0$, $v' > 0$ and $v'' > 0$, and $u(\cdot)$ is jointly concave.[5] I also impose the condition that $u_C < 0$.

According to the previous specification, individuals value their own consumption *relative* to what the rest of all individuals consume. Hence, the so called "reference group" in this economy is the whole society itself. Since the utility of an agent decreases as the weighted average of consumption increases, we say that preferences exhibit *jealousy*.[6]

[4] As standard in this literature, I define effective labor as $y = \theta l$ where l is the amount of time worked.

[5] Utility specification such as $u(c, C) = \tilde{u}(c - \alpha C)$, $\alpha \in [0, 1]$ with $\tilde{u}' < 0$ and $\tilde{u}'' \leq 0$ satisfy those assumptions.

[6] This is in line with the terminology proposed in Dupor and Liu (2003).

It is important to remark that according to (2.1), the weighting function $\psi(\theta) : \Theta \to \mathbf{R}$ does not need to be equal to $f(\theta)$.[7] In other words, agents may contribute to the consumption externality that society faces in a magnitude different from their population size. Hereafter, I will refer to $\psi(\theta)$ as the externality weighting function. For further reference, notice that (2.1) can be reexpressed as

$$C \equiv \int_{\Theta} c(\theta) f(\theta) \frac{\psi(\theta)}{f(\theta)} d\theta \tag{2.2}$$

hence the ratio $\frac{\psi(\theta)}{f(\theta)}$ acts as a weighting variable of the consumption externality that the society faces.

An *allocation* in this economy is $\{c(\theta), y(\theta)\}_{\theta\in\Theta}$, where $c : \Theta \to \mathbf{R}_+$ and $y : \Theta \to \mathbf{R}_+$. Abstracting from government expenditure, I define an allocation $\{c(\theta), y(\theta)\}_{\theta\in\Theta}$ to be *feasible* if

$$\int_{\Theta} c(\theta) f(\theta) d\theta = \int_{\Theta} y(\theta) f(\theta) d\theta \tag{2.3}$$

Making use of the Revelation Principle, an allocation is *incentive compatible* if

$$u(c(\theta), C) - v\left(\frac{y(\theta)}{\theta}\right) \geq u(c(\theta'), C) - v\left(\frac{y(\theta')}{\theta}\right) \quad \forall \theta, \theta' \in \Theta \tag{2.4}$$

Observe that since C cannot be affected unilaterally by a single agent, it is not

[7] Most of the literature on relative consumption valuation assumes that individuals value their own consumption relative to the *average* consumption. In that case $\psi(\theta) = f(\theta)$. An example of this is Tuomala (1990). Abel (2005) is an exception since he considers a weighted geometric average in the context of an overlapping generations economy.

a function of θ. An allocation that is incentive compatible and feasible is said to be *incentive-feasible.* Finally, let $g : \Theta \to \mathbf{R}_+$ be the density according to which individuals are weighted by the benevolent planner.

Definition 1. *An optimal allocation is an allocation* $\{c^*(\theta), y^*(\theta)\}_{\theta\in\Theta}$ *that maximizes the following planner problem*

$$\int_\Theta \left[u(c(\theta), C) - v\left(\frac{y(\theta)}{\theta}\right)\right] g(\theta)d\theta \tag{2.5}$$

subject to $\{c(\theta), y(\theta)\}_{\theta\in\Theta}$ *being incentive-feasible and* C *as defined in (2.1).*

2.2.1 Characterization of Optimal Allocations

The following proposition expresses the necessary conditions that any interior optimal allocation must satisfy. Let $\epsilon^*(\theta) \equiv \frac{v'(\frac{y^*(\theta)}{\theta})}{v''(\frac{y^*(\theta)}{\theta})\frac{y^*(\theta)}{\theta}}$.

Proposition 1. *Any interior optimal allocation* $\{c^*(\theta), y^*(\theta)\}_{\theta\in\Theta}$ *must be incentive-feasible and satisfy*

$$\frac{u_c(c^*(\theta), C^*)}{v'(\frac{y^*(\theta)}{\theta})\frac{1}{\theta}} - 1 = \frac{\gamma\psi(\theta)}{\lambda f(\theta)} + \frac{u_c(c^*(\theta), C^*)}{\theta f(\theta)}\left[1 + \frac{1}{\epsilon^*(\theta)}\right] \times$$

$$\int_{\underline{\theta}}^{\theta}\left[\frac{g(t)}{\lambda} - \frac{f(t)}{u_c(c^*(t), C^*)} - \frac{\gamma}{\lambda}\frac{\psi(t)}{u_c(c^*(t), C^*)}\right] dt \tag{2.6}$$

$$\frac{\gamma}{\lambda} = \frac{-\int_\Theta \frac{u_C(c^*(\theta),C^*)}{u_c(c^*(\theta),C^*)} f(\theta)d\theta}{1 + \int_\Theta \frac{u_C(c^*(\theta),C^*)}{u_c(c^*(\theta),C^*)}\psi(\theta)d\theta}. \tag{2.7}$$

where

$$C^* \equiv \int_\Theta c^*(\theta)\psi(\theta)d\theta \tag{2.8}$$

Proof. See Appendix A. $\square$

Notice that the previous solution collapses into the solution of a Mirrlessian economy with no consumption externalities when $\gamma = 0$. This is true since in that case $u_C(c(\theta), C) = 0$. In section 2.2.4, I make assumptions that facilitate the interpretation of the previous expression.

2.2.2 Implementation

Agents in this economy trade effective labor for consumption. There is a single firm that employs all agents. It produces one unit of output for every unit of effective labor, y. Every unit of effective labor receives a payment of w. Agents are also subject to an income tax schedule $T(y(\theta))$, assumed to be twice differentiable and to induce no bunching. Without loss of generality, there are not taxes on consumption c. An agent with effective labor y pays taxes $T(y(\theta))$. Thus, taking as given $T(y(\theta))$, C and the wage w, the problem solved by the agent with productivity $\theta, \forall\theta \in \Theta$ is

$$\max_{c(\theta),y(\theta)} u(c(\theta), C) - v\left(\frac{y(\theta)}{\theta}\right) \tag{2.9}$$

s.t.

$$c(\theta) \leq wy(\theta) - T(y(\theta))$$

$$c(\theta), y(\theta) \geq 0$$

Definition 2. *Given a labor tax $T(y(\theta))$ and C, an equilibrium in this economy is an allocation $\{c^{eq}(\theta), y^{eq}(\theta)\}_{\theta \in \Theta}$ and wage w^{eq} such that*

i. $(c^{eq}(\theta), y^{eq}(\theta))$ *solve (2.9)* $\forall \theta \in \Theta$

ii. $C = \int_{\Theta} c^{eq}(\theta)\psi(\theta)d\theta$

iii. $w^{eq} = 1$

iv. $\int_{\Theta} T(y^{eq}(\theta))f(\theta)d\theta = 0$

v. $\int_{\Theta} c^{eq}(\theta)f(\theta)d\theta = \int_{\Theta} y^{eq}(\theta)f(\theta)d\theta$

An allocation $\{c(\theta), y(\theta)\}_{\theta \in \Theta}$ is *implementable* by the income tax $T(y(\theta))$ if $\{c(\theta), y(\theta)\}_{\theta \in \Theta}$ and w are an equilibrium.

2.2.3 Characterization of Optimal Income Tax

Define the following tax mechanism $T : y \to \mathbf{R}$,

$$T(y(\theta)) = \begin{cases} y^*(\theta) - c^*(\theta) & \text{if } y(\theta) = y^*(\theta) \\ y(\theta) & \text{otherwise.} \end{cases} \tag{2.10}$$

together with

$$\frac{T'(y^*(\theta))}{1 - T'(y^*(\theta))} = \frac{\gamma\psi(\theta)}{\lambda f(\theta)} + \frac{u_c(c^*(\theta), C^*)}{\theta f(\theta)}\left[1 + \frac{1}{\epsilon^*(\theta)}\right] \times$$

$$\int_{\underline{\theta}}^{\theta} \left[\frac{g(t)}{\lambda} - \frac{f(t)}{u_c(c^*(t), C^*)} - \frac{\gamma}{\lambda}\frac{\psi(t)}{u_c(c^*(t), C^*)}\right] dt \tag{2.11}$$

if $y(\theta) = y^*(\theta)$.

Proposition 2. *Any optimal allocation* $\{c^*(\theta), y^*(\theta)\}$ *can be implemented by a tax schedule* $T(y(\theta))$ *defined by (2.10) and (2.11).*

Proof. See Appendix A. $\square$

2.2.4 The Quasi-linear Environment

In order to keep the model tractable and facilitate the interpretation of the optimal income tax, I consider the case of $u(c, C) = (1 - \alpha)c + \alpha(c - C), \alpha \in [0, 1)$, and C defined as in (2.1).[8] There are two important features to highlight about this preference specification. First, the parameter α measures positional concerns agents may have. As α approaches one from below, consumption is valued almost fully with respect to the economy's endogenous consumption benchmark. Conversely, when $\alpha = 0$, the consumption of others is completely irrelevant for a given individual's satisfaction. Second, under the former specification, the shadow price of aggregate consumption is $1 - \alpha$. That is, if all agents in an economy were to consume one unit of the good, a share α of aggregate utility would vanish due to jealousy. To see why, suppose all agents were to be given one unit of the consumption good. Such consumption would provide only $1 - \alpha$ utils to an agent after she realizes that not just her, but *all* agents are consuming an extra unit. Hereafter, I will refer to the term α as the jealousy parameter.

[8] Carlsson, Johansson-Stenman, and Martinsson (2007) also use this functional form to measure what they call *marginal degree of positionality.* Notice that this functional form is equivalent to $u(c, C) = c - \alpha C$. It is the fraction of the marginal utility in income that is due to the increase in relative income, the term, $c - C$. For this specification, the marginal degree of positionality is α.

To state the next proposition, let $G(\theta) \equiv \int_{\underline{\theta}}^{\theta} g(t)dt$, $F(\theta) \equiv \int_{\underline{\theta}}^{\theta} f(t)dt$ and $\Psi(\theta) \equiv \int_{\underline{\theta}}^{\theta} \psi(t)dt$.

Proposition 3. *Suppose $u(c, C) = c - \alpha C, \alpha \in [0, 1)$ and C defined as in (2.1), then any optimal marginal income tax satisfies*

$$\frac{T'(y^*(\theta))}{1 - T'(y^*(\theta))} = \underbrace{\frac{[1 + \epsilon^*(\theta)^{-1}]}{\theta f(\theta)}[G(\theta) - F(\theta)]}_{\textit{Mirrleesian tax}} + \underbrace{\frac{\alpha}{1-\alpha}\left[\frac{\psi(\theta)}{f(\theta)} + \frac{[1 + \epsilon^*(\theta)^{-1}]}{\theta f(\theta)}[G(\theta) - \Psi(\theta)]\right]}_{\textit{Pigouvian tax}} \tag{2.12}$$

Proof. See Appendix A ☐

From (2.12), it is easy to see that without utility interdependence ($\alpha = 0$), the optimal income tax is simply the Mirrleesian tax.[9] Moreover, observe that if $g(\theta) = f(\theta) \forall \theta \in \Theta$, the Mirrleesian tax is equal to zero $\forall \theta \in \Theta$. Atkinson (1990) refers to this case as complete distributional indifference.

Regarding the Pigouvian tax, notice that it is increasing in the jealousy parameter α. The term $\frac{\psi(\theta)}{f(\theta)}$ affects this tax component since this is precisely the magnitude that the Pigouvian tax corrects directly. But, what is the role of the term $G(\theta) - \Psi(\theta)$? If the planner redistributive taste is such that $G(\theta) \leq \Psi(\theta) \forall \theta \in \Theta$,[10] then the Pigouvian tax will be reduced over its *direct* component, the term $\frac{\alpha}{1-\alpha}\frac{\psi(\theta)}{f(\theta)}$. The intuition is as follows: if the consumption of those individuals heavily weighted by the planner does

[9]See Salanié (1997) for a direct derivation of the Mirrleesian tax for a quasi-linear model without consumption externalities. The quasi-linear case has also been analyzed by Atkinson (1990) and Diamond (1998).

[10]In this case, $G(\theta)$ first order stochastically dominates $\Psi(\theta)$.

not contribute much to the consumption externality, it will be optimal to lower the Pigouvian tax below its direct component. Conversely, if $G(\theta) \geq \Psi(\theta) \forall \theta \in \Theta$ then the Pigouvian tax is higher. That is, if the consumption externality is highly driven by those agents with a low weight assigned by the planner, then it would be optimal to increase the Pigouvian tax over its direct component. The former case may represent a situation in which the planner weights poor agents higher than rich ones while at the same time rich agents have a higher contribution to the consumption externality than poor ones.

Corollary 1 (Proposition 3)**.** *Marginal taxes at the top and the bottom of the earnings distribution satisfy* $\frac{T'(y^*(\bar{\theta}))}{1-T'(y^*(\bar{\theta}))} = \frac{\alpha}{(1-\alpha)} \frac{\psi(\bar{\theta})}{f(\bar{\theta})}$ *and* $\frac{T'(y^*(\underline{\theta}))}{1-T'(y^*\underline{\theta})} = \frac{\alpha}{(1-\alpha)} \frac{\psi(\underline{\theta})}{f(\underline{\theta})}$.

This corollary makes clear that non zero taxation at the top and the bottom of the distribution is optimal whenever $\alpha > 0$. This is due to corrective considerations. Moreover, unless $\psi(\theta) = f(\theta) \forall \theta \in \Theta$, taxes at the top and the bottom do not need to be equal. These features of the optimal income tax occur despite the Mirrleesian component at the extremes being zero. In Appendix A.2, I derive an expression for the asymptotic optimal income tax when $f(\theta)$ has an unbounded support. This is done under the assumption that $f(\theta)$ is distributed according to a Pareto distribution. A similar case (without consumption externalities) is analyzed by Saez (2001) and Diamond (1998) based on the premise that the upper tail of the earnings distribution can be well approximated by the aforementioned distribution. However, such insight cannot be confirmed after analyzing the very top of the income distribution in the U.S. for several years using non-parametric smoothing techniques. This analysis is also presented in Appendix A.2. This implies that the thickness of the income distribution may not make a case for positive taxes at the high end of the labor income distribution. Nonetheless, as seen in Corollary 1, positional concerns and corrective considerations can account for that.

2.2.5 Recovering the Externality Weighting Function

In this section I state and prove my main theoretical result. Put simply, this result states that conditional on a social planner's weighting function $g(\theta)$ and given a *marginal* income tax schedule $T'(y)$, gross income density $f_Y(y)$ and labor supply elasticity, it is always possible to find the externality weighting function $\psi(\theta)$ and the jealousy parameter α that make a case for the observed marginal income tax schedule. Given the imposed functional forms, the strength of this result is the possibility it creates to apply the model to the data.

Definition 3. *The parameter $\alpha \in \mathbf{R}_{++}$ and the externality weighting function $\psi : \Theta \to \mathbf{R}$ rationalize a marginal tax schedule $T'(y)$ if the resulting equilibrium allocation $\{c^{eq}(\theta), y^{eq}(\theta)\}_{\theta\in\Theta} = \{c^*(\theta), y^*(\theta)\}_{\theta\in\Theta}$.*

Theorem 1. *Suppose $T'(y) \in \mathcal{C}\ \forall y \in [\underline{y}, \bar{y}]$ with $\underline{y} > \max\{0, \tilde{y}\}$, $\tilde{y} = \inf\{y \mid (1 - T'(y))\phi y^{\phi-1} + T''(y)y^{\phi} > 0\}$. If $u(c, C) = c - \alpha C$, $C = \int_\Theta c(\theta)\psi(\theta)d\theta$, $v\left(\frac{y}{\theta}\right) = \frac{1}{1+\phi}\left(\frac{y}{\theta}\right)^{1+\phi}, \phi > 0$, there exists a unique mapping $\mathcal{M} : (g(\cdot); \phi, f_Y(y), T'(y)) \to (\psi(\cdot), \alpha)$ that rationalizes $T'(y)$. The externality weighting function $\psi : \Theta \to \mathbf{R}$ and the jealousy* *parameter $\alpha \in \mathbf{R}$ that rationalize $T'(y)$ satisfy*

$$\psi(\theta) = b(\theta) + (1+\phi)\theta^{\phi}\left[\int_{\underline{\theta}}^{\theta} \frac{b(t)}{t^{1+\phi}}dt\right] \tag{2.13}$$

where

$$b(\theta) \equiv \frac{(1-\alpha)f(\theta)T'(y(\theta))}{\alpha(1-T'(y(\theta)))} - \frac{(1+\phi)G(\theta)}{\alpha\theta} + \frac{(1-\alpha)(1+\phi)F(\theta)}{\alpha\theta} \tag{2.14}$$

$$\alpha = \frac{\int_{\Theta} \frac{f(\theta)}{\theta^{\phi+1}(1-T'(y(\theta)))}d\theta - \int_{\Theta} \frac{g(\theta)}{\theta^{1+\phi}}d\theta}{\int_{\Theta} \frac{f(\theta)}{\theta^{\phi+1}(1-T'(y(\theta)))}d\theta} \tag{2.15}$$

and

$$f(\theta) = f_Y(\Phi^{-1}(\theta))\frac{\partial\Phi^{-1}(\theta)}{\partial\theta} \tag{2.16}$$

where $\Phi(y) = \left[\frac{y^{\phi}}{1-T'(y)}\right]^{\frac{1}{1+\phi}}$. *If in addition,* $\int_{\Theta} \frac{f(\theta)}{\theta^{\phi+1}(1-T'(y(\theta)))}d\theta - \int_{\Theta} \frac{g(\theta)}{\theta^{1+\phi}}d\theta \geq 0$, *then* $\alpha \in [0,1)$.

Proof. If $u(c,C) = c - \alpha C$ and $v(\frac{y}{\theta}) = \frac{1}{1+\phi}(\frac{y}{\theta})^{1+\phi}$, it follows from Proposition 3 that

$$\frac{T'(y(\theta))}{1-T'(y(\theta))} = \frac{\alpha}{(1-\alpha)}\frac{\psi(\theta)}{f(\theta)} + \frac{(1+\phi)}{(1-\alpha)\theta f(\theta)}[G(\theta) - (1-\alpha)F(\theta) - \alpha\Psi(\theta)] \tag{2.17}$$

where $\psi(\theta) = \Psi'(\theta)$. Expression (2.17) is a first order ordinary differential equation of the form

$$\psi(\theta) + a(\theta)\Psi(\theta) = b(\theta)$$

where $a(\theta) \equiv -\frac{1}{\theta}[1+\phi]$ and $b(\theta) \equiv \frac{(1-\alpha)f(\theta)T'(y(\theta))}{\alpha(1-T'(y(\theta)))} - \frac{(1+\phi)G(\theta)}{\alpha\theta} + \frac{(1-\alpha)(1+\phi)F(\theta)}{\alpha\theta}$. If $a(\theta)$ and $b(\theta)$ are continuous on Θ, according to Coddington (1989), Theorem 3, Chapter 1, $\Psi(\theta)$ satisfies

$$\Psi(\theta) = e^{-\int_{\underline{\theta}}^{\theta} a(t)dt} \left[\int_{\underline{\theta}}^{\theta} e^{\int_{\underline{\theta}}^{t} a(x)dx} b(t)dt + \kappa \right] \tag{2.18}$$

The fact that $T'(y)$ is continuous $\forall y \in [\underline{y}, \bar{y}]$ with $\underline{y} > 0$ guarantee the continuity of $a(\theta)$ and $b(\theta)$. By setting $\kappa = 0$ in (2.18) we have $\Psi(\underline{\theta}) = 0$ since $\Psi(\theta)$ is a cumulative weighting function. Moreover, since α is an argument of $\Psi(\theta)$ in (2.18) through $b(t)$, setting

$$\alpha = \frac{\int_{\Theta} \frac{f(\theta)}{\theta^{\phi+1}(1-T'(y(\theta)))} d\theta - \int_{\Theta} \frac{g(\theta)}{\theta^{1+\phi}} d\theta}{\int_{\Theta} \frac{f(\theta)}{\theta^{\phi+1}(1-T'(y(\theta)))} d\theta}$$

we normalize $\Psi(\bar{\theta}) = 1$. Using $\int_{\Theta} \frac{f(\theta)}{\theta^{\phi+1}(1-T'(y(\theta)))} d\theta - \int_{\Theta} \frac{g(\theta)}{\theta^{1+\phi}} d\theta \geq 0$, we obtain $\alpha \in [0, 1)$.

The last step involves the identification of the skills distribution $f(\theta)$ from income distribution $f_Y(y)$.[11] For that, notice that from the first order condition of the consumer's problem we have $(1 - T'(y)) = \left(\frac{y}{\theta}\right)^{\phi} \left(\frac{1}{\theta}\right)$ and hence, $\theta = \Phi(y) = \left[\frac{y^{\phi}}{1-T'(y)}\right]^{\frac{1}{1+\phi}}$. Therefore, $f(\theta) = f_Y(\Phi^{-1}(\theta)) \frac{\partial \Phi^{-1}(\theta)}{\partial \theta}$. The fact that $y > \tilde{y}$ guarantees the invertibility of $\Phi(\cdot)$. $\square$

In order to gain insight into how expression (2.15) makes possible calculating the jealousy parameter, α, let us consider a very simple example. Suppose that $g(\theta) = f(\theta) \forall \theta \in \Theta$. Thus, we are in a case of distributional indifference and as established before, the observed marginal tax $T'(y(\theta))$ must be purely Pigouvian (see equation

[11] This approach is also followed by Saez (2001).

(2.12)). Further, let us assume that $T'(y(\theta)) = T' \forall y \in [\underline{y}, \bar{y}]$, i.e., society faces a flat tax. Evaluating equation (2.15), we obtain $\alpha = T'$. Thus, we recover the jealousy parameter from taxes.

2.2.6 An Upper Bound for the Jealousy Parameter

As stated in Theorem 1, it is possible to calculate $\psi(\theta)$ and α conditional on the redistributive taste of the planner represented by the density $g(\theta)$. This element, however, is non observable and thus the model I present has an identification problem. Nevertheless, it is possible to calculate an upper bound of α under the loose assumption that society favors redistribution. This statement is formally expressed in Proposition 4.

Proposition 4. *Let* $\Gamma_F \equiv \{g : \Theta \to \mathbf{R}_+ \mid G(\theta) \geq F(\theta) \quad \forall \theta \in \Theta\}$. *Then* $\alpha \leq \bar{\alpha} \quad \forall g \in \Gamma_F$, *where* α *satisfies (2.15) and*

$$\bar{\alpha} = \frac{\int_\Theta \frac{f(\theta)T'(y(\theta))}{\theta^{1+\phi}(1-T'(y(\theta)))} d\theta}{\int_\Theta \frac{f(\theta)}{\theta^{1+\phi}(1-T'(y(\theta)))} d\theta} = \frac{E_Y[y^{-\phi}T'(y)]}{E_Y[y^{-\phi}]}$$

Moreover,

$$f(\theta) = \arg\max_{g(\cdot)} \{\alpha \mid g \in \Gamma_F\}$$

Proof. See Appendix A.3. □

Proposition 5. *Suppose* $T'(y(\bar{\theta})) \geq T'(y(\theta)) \forall \theta \in \Theta$, *then* $\frac{\psi(\bar{\theta})}{f(\bar{\theta})} \geq 1 \ \forall g \in \Gamma_F$.

Proof. See Appendix A.3. □

Two things need to be highlighted about Proposition 4. First, the calculation of the upper bound of the jealousy parameter α requires only observed variables: the labor earnings distribution, the marginal tax schedule on labor income and the elasticity of labor supply. Second, the upper bound of the jealousy parameter is attained when the planner is utilitarian, i.e. $g(\theta) = f(\theta) \forall \theta \in \Theta$. The intuition for this result is straightforward: when the planner is utilitarian, there is complete distributional indifference (given the quasi-linearity in preferences) and the Mirrleesian tax is zero at any income level. Thus, all taxation must be Pigouvian. Proposition 5 finds a lower bound for the contribution to the consumption externality at the top of the income distribution, the ratio $\frac{\psi(\bar{\theta})}{f(\bar{\theta})}$. In the next section, I estimate the upper bound of the jealousy parameter and the corresponding $\psi(\cdot)$ for the U.S. and the U.K.

2.3 The Data

My numerical analysis was done for the U.S. and the U.K. This choice was based on the public availability of micro-file tax data that I describe in this section. For the U.S, I use the Statistics of Income (SOI) Public Use Tax files elaborated by the Internal Revenue Service (IRS) and distributed by the NBER.[12] The data consists of the information that U.S. citizens and residents submit to the IRS through the 1040, 1040A and 1040EZ tax forms. Cross-section samples of approximately 100,000 to 150,000 observations are available for each year from 1960 to 2004. Such samples were designed to make national

[12] See Internal-Revenue-Service (1995-2004).

level estimates by including a weighting variable to make up for the stratified nature of the sample.[13] In order to abstract from capital holdings, the definition of *gross* income that I use is salaries and wages. This is the entry of the 1040 tax forms specified as "Wages, salaries, tips, etc". In addition, the marginal income tax corresponding to different income brackets was collected from the "Tax Rate Schedule" from 1995 to 2004 published by the IRS. An additional measure of taxes that I use is total tax liabilities minus total tax credits.[14]

For the U.K, I use the Survey of Personal Incomes (SPI) Public Use File from the Economic and Social Data Service provided by the University of Essex.[15] These files are compiled by Her Majesty's Revenue and Customs: Knowledge, Analysis & Intelligence, and are based on information held by HM Revenue and Customs (HMRC) tax offices on individuals who could be liable for U.K. taxes. Cross-section samples of approximately 450,000 observations are available for each year from 1995 to 2004.[16] The data set also contains a variable that allows to obtain figures for the whole U.K. population. I use two variables from this database. The income measure is the variable defined as *pay* that stands for pay from employment net of benefits and foreign earnings and the variable

[13]The General Description Booklet for the Public Use Tax Files (several years) indicates that the sample design is a stratified probability sample and the population of tax return is classified into sub-populations (strata). According to the same source, independent samples are selected independently from each stratum. A weighting variable is obtained by dividing the population count of returns in a stratum by the number of sample returns for that stratum.

[14]The last two concepts are not extracted directly from tax forms but provided in SOI files. Total tax liabilities are derived from *total tax liability* entry whereas total tax credit are derived from *total tax credits* entry. Some minor adjustments, made by SOI data providers, are made for the derivation.

[15]See HM-Revenue-&-Customs (1998-2007).

[16]There is an extra file for the fiscal year 1985-1986. Moreover, the HM Revenue and Customs: Knowledge, Analysis & Intelligence recently released the file for 2005.

tottax that is calculated as total tax liability less tax credits.

2.3.1 Estimation of Gross Income Densities

The period of analysis is 1995-2004 since I have comparable data for the two countries during those years. Even though the model is static, taking into account the gross income distributions over several years allows me to obtain a more "robust" distribution for each country calculated as

$$f_Y(y) = \frac{1}{10}\sum_{t=1995}^{2004} f_{Y,t}(y) \tag{2.19}$$

To calculate $f_{Y,t}(y)$ for $t = 1995, ..., 2004$, I selected the domain $y > \$100$ measured in 2004 dollars. I calculated $f_{Y,t}(y)$ using a gaussian kernel over many points unequally separated, locating more points at the bottom of the domain than at the top. This was done in order to obtain more accurate estimates from numerical integration while at the same time keeping moderate the number of grid points.[17] Income observations were weighted to obtain population estimates.

Since income observations are more sparse as income is higher, I smoothed the data after transforming it into a logarithmic scale. According to Wand, Marron, and Ruppert (1991), this transformation is appropriate under the presence of a global width parameter h and data being more sparse at the top than at the bottom of its domain. Without

[17]For both countries, the number of grid points was 5,000. Moreover, I also performed my calculations using alternative kernel functions such as the Epanechnikov and Triangular and results were almost identical. The latter was not surprising given the sample size. See Silverman (1986).

this transformation, either the data at the bottom of the distribution would be over-smoothed or the tail would exhibit spurious bumps. The smoothing window or width for the U.S. was set at $h = 0.9 \times s.d. \times n^{-1/5}$, where n is the number of observations and $s.d.$ is the standard deviation of $\log(y_i)$, for $i = 1, ..., n$.[18] Silverman (1986) indicates that such a choice of width performs very well in terms of mean integrated squared error for a wide range of densities. For the U.K., I set $h = 2 \times s.d. \times n^{-1/5}$ since lower widths produced a spurious bump at the top of the distribution.[19] Figure 2.1 shows the estimated densities.

2.3.2 Estimation of Marginal Income Tax Schedules

For both countries, the marginal income tax that I use is the statutory one. For the U.S., I collect the tax rates corresponding to different income brackets as published by the IRS in the "Tax Rate Schedule" from 1995 to 2004. I use the brackets corresponding to single people. To come up with a single marginal tax rate schedule for several years, I express all tax rate schedules in 2004 dollars and for every y, I take a simple average over the ten years I collected data. To maintain the constructed tax rate as a step function, I define the boundaries of the brackets as the average boundary of yearly brackets. For the U.K., I employ exactly the same procedure. This time, I collect statutory income tax rates

[18] Indeed, the formula suggested by Silverman (1986) is $width = 0.9 \times A \times n^{-1/5}$, where $A = \min\{s.d., iqr/1.34\}$, where iqr stands for interquartile range. For both countries and all years, $s.d.$ was lower than $iqr/1.34$.

[19] For the U.K., I first used Silverman (1986) optimal bandwidth. However, the resulting income density appeared to need further smoothing at the upper tail. To correct for this, I increased the bandwidth sequentially until the income density looked smooth. This explains the use of $h = 2 \times s.d. \times n^{-1/5}$.

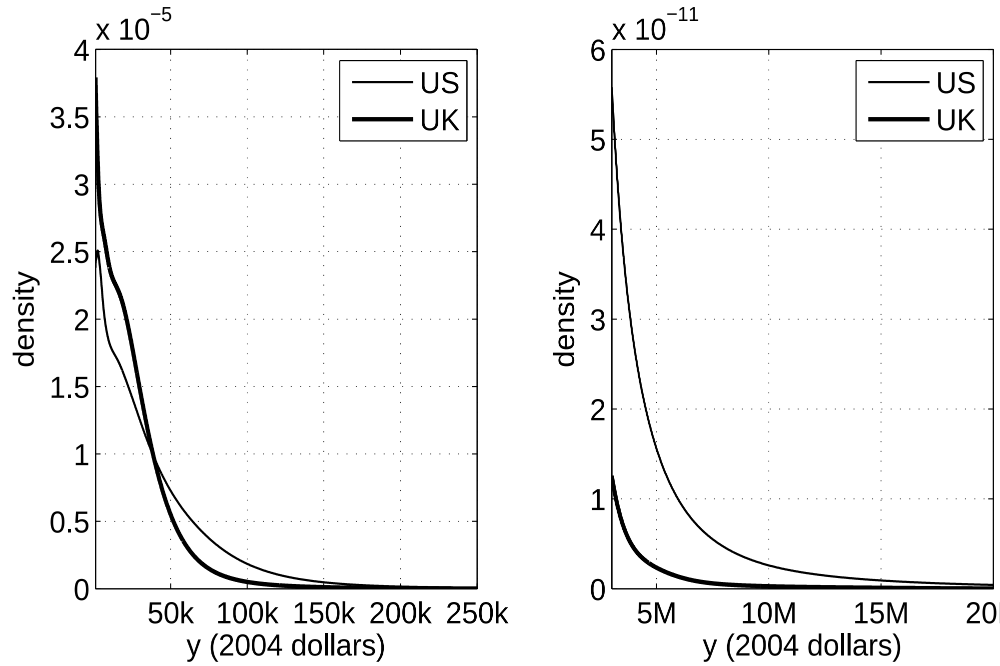

FIGURE 2.1: Gross Income Distribution in the U.S. and the U.K. 1995-2004

from the "Survey of Personal Incomes Public Use Tape Documentation: Annex D: Rates of Income Tax: 1990-91 to 2004-05" located in HM-Revenue-&-Customs (1998-2007). Figure 2.2 shows the estimated marginal income tax schedules.

2.4 Results

In this section, I present my estimates of agents' contributions to the consumption externality, the ratio $\frac{\psi(\cdot)}{f(\cdot)}$.[20] This is done for the case in which the benevolent planner is utilitarian, i.e., $g(\theta) = f(\theta) \forall \theta \in \Theta$. In this instance, all taxation is Pigouvian

[20] The estimation uses a continuous and differentiable version of the marginal tax schedule. See Appendix A.4 for details.

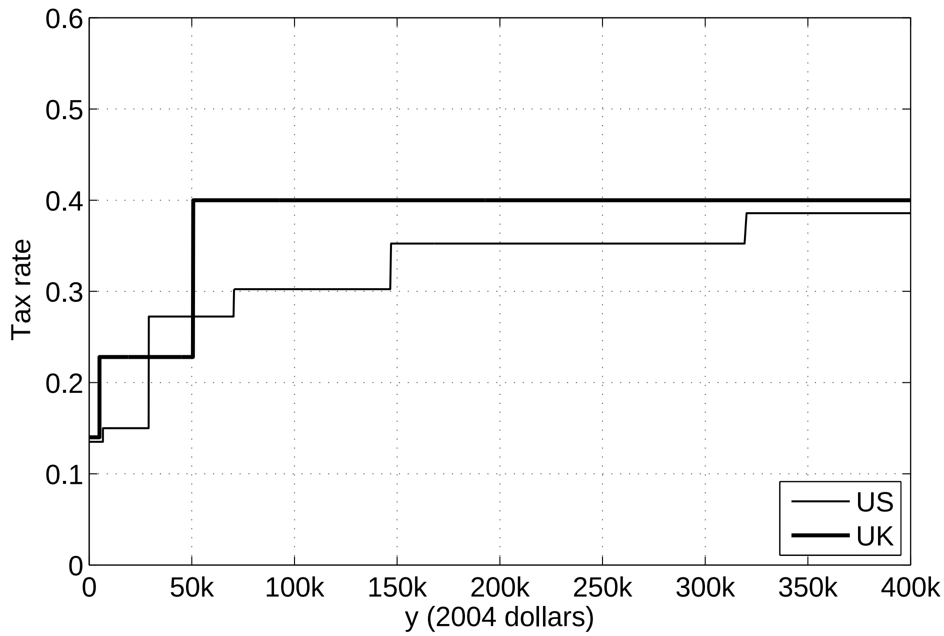

FIGURE 2.2: Statutory Marginal Income Tax in the U.S. and the U.K. 1995-2004

or corrective and the jealousy parameter attains its upper bound. That is, $\alpha = \bar{\alpha}$. Under the assumption that the American and British societies redistribute, the jealousy parameter α associated with the "actual" planner's weighting density will not be higher than $\bar{\alpha}$. The estimated parameters are surprisingly moderate. For the American society, I estimate $\bar{\alpha}^{us} = 0.135$ while for the British one, I obtain $\bar{\alpha}^{uk} = 0.14$. Thus, under the assumption that the planner is utilitarian, the British society seems to have slightly higher positional concerns than the American one. Now, the question is, who are these societies jealous of? To answer this, I plot $\frac{\psi(\cdot)}{f(\cdot)}$ against the gross income cdf, $F_Y(y)$. I present my estimations under the assumption that $\phi = 3$ for both countries.[21] That is, the elasticity of labor supply is $\frac{1}{3}$.[22]

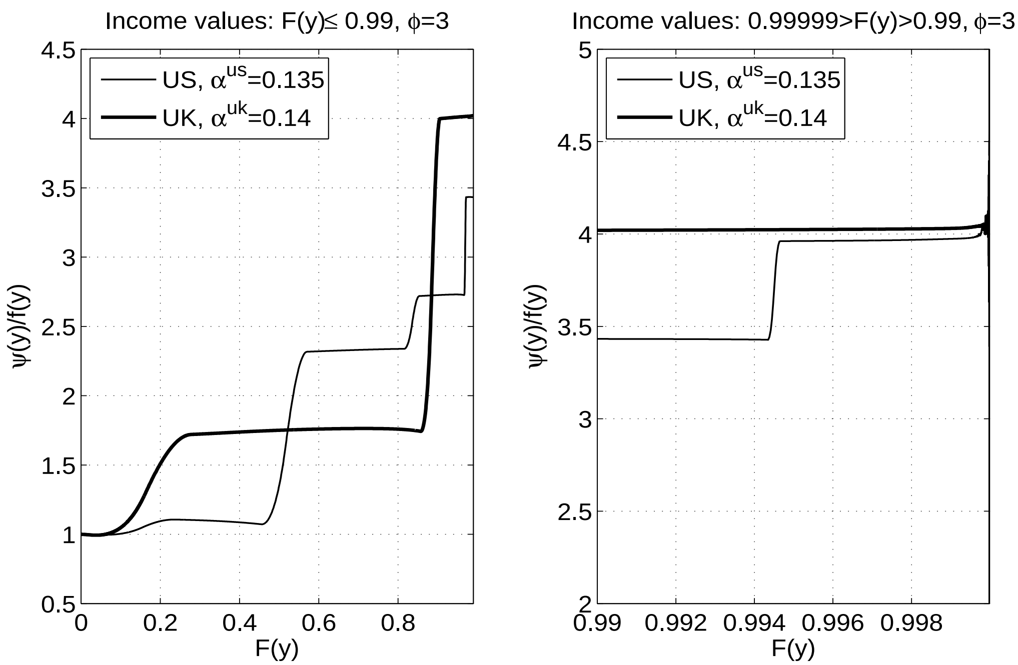

FIGURE 2.3: Contribution to Consumption Externality in the U.S. and the U.K. 1995-2004

According to Figure 2.3, the ratio $\frac{\psi(y)}{f(y)}$ is increasing in income. In words, the contribution to the consumption externality is higher the more affluent individuals are. The ratio $\frac{\psi(y)}{f(y)}$ in the United Kingdom is almost flat and close to 1.75 up to the 85th percentile of the gross income distribution. From there on, it increases sharply reaching a level a bit less than 4. For the United States, this ratio increases sharply around the

[21] Evers, Mooij, and Vuuren (2005) find that differences in estimates of labor supply elasticities across countries appear to be small. Both, U.S. and U.K. are included in their sample of countries.

[22] This choice is in line with the work of Diamond (1998) who chooses $\phi = \{2, 5\}$ for a model with no income effects and constant elasticity of labor supply. This is based on the work of Pencavel (1986).

median of the gross income distribution reaching a level of around 2.5. This variable exhibits another sharp increase at the very upper tail of the gross distribution where it hits a level close to 4.[23]

2.4.1 American Tax System Under British Positional Preferences

In this section, I simulate the American income tax schedule imposing on the United States the positional preferences of the United Kingdom. Those positional preferences are represented by the parameter α, the ratio $\psi(\theta)/f(\theta)$ and $\Psi(\theta)$. Since I am assuming that the planner in both the U.S. and the U.K. is utilitarian, we have that $G(\theta) = F(\theta)$ and the tax is purely Pigouvian. Thus, I evaluate the following expression

$$\frac{\hat{T}'(y^*(\theta))}{1-\hat{T}'(y^*(\theta))} = \frac{\alpha^{uk}}{1-\alpha^{uk}}\left[\left(\frac{\psi(\theta)}{f(\theta)}\right)^{uk} + \frac{(1+\phi)}{\theta f(\theta)}\Big(F(\theta) - \Psi^{uk}(\theta)\Big)\right] \tag{2.20}$$

which is plotted in Figure 2.4. Observe that under the British positional considerations, the tax schedule in the United States is slightly increasing up to an income level of \$150,000 (2004 dollars) with a mean of around 23%. After that income level, it increases sharply up to a level of approximately 43%.

[23] Appendix A.5 presents an estimation for the ratio $\psi(\cdot)/f(\cdot)$ using *effective* marginal income taxes as estimated in Gouveia and Strauss (1994) instead of *statutory* marginal income taxes. Qualitatively, both estimations are similar, yet the choice of the minimum income level is crucial for the estimations using effective taxes.

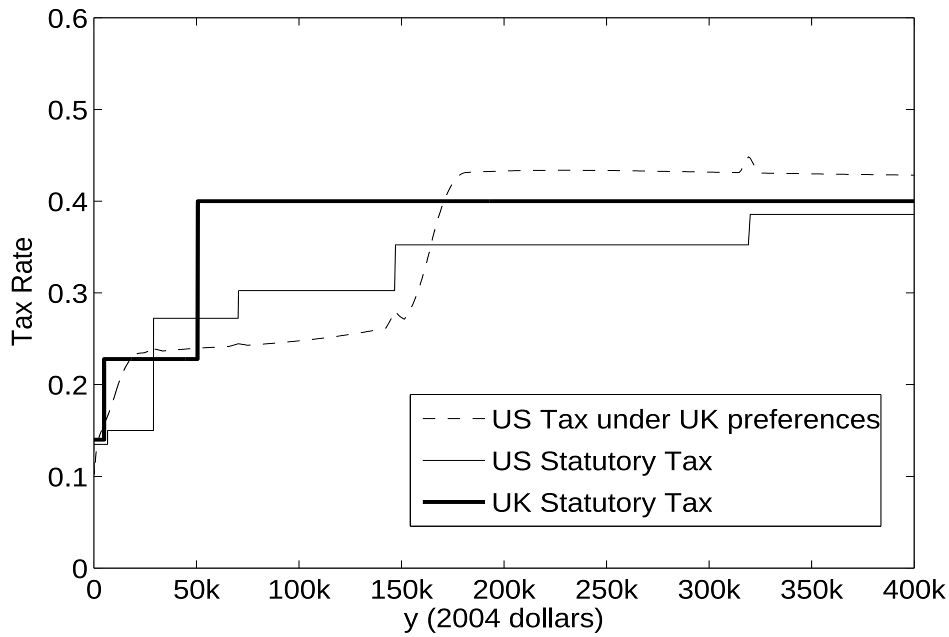

FIGURE 2.4: Marginal Income Tax Schedule in the U.S. under British Positional Preferences

1995-2004

2.5 Conclusions

In this chapter I have presented a model that rationalizes high labor income taxes on affluent individuals: taxation on the rich occurs in order to correct consumption externalities. This happens in the absence of a non-linear consumption tax schedule. Surprisingly, the estimated jealousy parameters for the U.S. and the U.K. are moderate, yet produce quantitatively high effects over labor income tax rates.

Rationalizing observed labor income taxes as Pigouvian requires that the jealousy provoked by individuals be increasing in income. In other words, more affluent individuals must generate a higher consumption externality than poorer ones. Despite productive efficiency arguments, this may be one of the reasons policy-makers in practice face impediments to reducing labor income taxes at high level.

In this light, it is important to analyze to what extent the presumed higher contribution of rich consumers to the consumption externality is a result of affluent individuals having access to consumption goods with higher positional effects as suggested by Frank (2008). This line of research is explored in Chapter 3. Further research is also necessary to understand the underlying preferences for which positional concerns are instruments.

Chapter 3

Optimal Linear Taxation of Positional Goods

3.1 Introduction

More that a hundred years ago, Veblen (1899) coined the term "conspicuous consumption" to refer to the consumption incurred by individuals with the primary goal of attaining status or social position. In modern capitalist societies, at least some goods possess this characteristic. Why is it that people are willing to pay fortunes to own a mansion in Beverly Hills, drive a brand new German convertible, have access to exclusive country clubs, etc? No one can deny the intrinsic value derived from the consumption of these goods; however, these purchases may also be motivated, at least partially, by positional considerations. In other words, the consumption of these goods is partially motivated by the status that they confer to the buyer.

By definition, status is a social ranking or standing. To the extent that agents' preferences are sensitive to a ranking based on the consumption of a particular set of goods, the consumption of the "Joneses" may be harmful. Government intervention to correct this type of externality, known as a positional externality, may be desirable but is highly controversial. Some articles such as Frank (2005) have argued in favor of policy targeting this type of externality. The author claims that "tax cuts for the wealthy are spent largely on positional goods. Dollars that could have been used to pay for additional non-positional goods have been spent instead on larger houses and more expensive cars".

The goal of this article is to conduct a normative analysis of taxation in the presence of consumption goods that generate positional externalities. From now on I will refer to these goods as positional goods. I conduct my analysis in a framework similar to the one pioneered by Mirrlees (1971). Agents in my model are endowed with heterogeneous privately-known productivities. As is well known, these models capture a conflict between redistribution and incentives which results in an endogenous non-degenerate consumption distribution. As individuals' preferences are sensitive to positional considerations, consumption inequality becomes harmful and government intervention à la Pigou is desirable. In the analysis below, I parameterize the strength of positional considerations agents may have and analyze optimal tax policy.

Not surprisingly, constrained efficient allocations in this environment exhibit a wedge between non-positional goods (for instance, necessities) and positional goods (for instance, luxuries). The consensus in the literature, however, is that taxing luxuries is not efficient. Using the Ramsey approach to optimal taxation, Atkinson and Stiglitz (1972) shows that it is optimal to tax goods with low rather than high income elasticities. Thus, necessities must be taxed at a higher rate than luxuries. A uniform commodity taxation result is obtained in Atkinson and Stiglitz (1976) in a framework with heterogeneous agents like the one analyzed in Mirrlees (1971). Remarkably, only the assumption of separability between consumption goods and leisure in preferences is needed to derive the latter result. Why is it that in this economic environment taxing positional goods such as luxuries is optimal? In the model that I present, taxation of the positional good occurs due to Pigouvian considerations. Thus, this instrument corrects over-consumption of a good that generates positional externalities. As preferences in this model display no utility interdependence, the uniform commodity taxation as in Atkinson and Stiglitz (1976) holds and all taxation happens due to redistributive purposes and may be carried out through the labor income tax.[1]

Under no restrictions on the class of taxes that can be used to implement optimal allocations, the presence of positional considerations implies that constrained efficient allocations may be implemented through a *non-linear* tax on the positional good in combination with a non-linear labor income tax with standard Mirrleesian properties, namely without distortions at the extremes. The non-linearity in the consumption of the

[1]Consumption taxes may also be used for redistribution, however, the uniform commodity taxation must hold.

positional good or "luxury tax" is driven by the fact that I allow agents to contribute to the positional externality in an arbitrary way. This is in line with the findings of Samano (2009) which finds that a progressive labor income tax may be partially rationalized as a Pigouvian one whose role is to correct consumption externalities, particularly at high income levels. The estimations of this paper suggest that such an externality may be increasing in income. The previous implementation, however, is subject to arbitrage opportunities across consumption goods, since agents face a non-linear consumption tax on positional goods. In order to prevent this, a non-arbitrage constraint is imposed by equalizing the marginal rate of substitution between the positional and the non-positional good across agents. I show that the resulting *double* constrained efficient allocations can be implemented through a *linear* positional tax together with a non-linear marginal labor income tax. The non-linear income tax that implements double constrained efficient allocations differs from the one implementing constrained efficient ones as the former must offset income effects produced by the flattening in the "luxury tax".

Numerical calculations indicate that while the aggregate welfare losses due to the prevention of arbitrage are very small, large distributional effects occur. When the positional externality is increasing in income, individuals at the high end of the income distribution experience large *gains*, since for them a flat tax effectively reduces the after-tax price of positional goods. This generates a positive income effect that cannot be fully offset by increases in the labor income tax as optimality requires no distortions at the top. Thus, the consumption of highly skilled individuals increases due to this income effect. Conversely, individuals at the bottom of the skills distribution experience losses since they effectively face a higher after-tax price of the positional good and, consequently, a negative income effect. In this case, a marginal income tax reduction cannot offset such an income effect due to incentive problems. Both effects are reduced when preferences over positional goods are non-homothetic, as price changes can be offset by small adjustments in the marginal income tax. This is true since price changes cause income-dependent income effects. My results suggest that the effectiveness of a linear "luxury tax" in correcting positional externalities would crucially depend on the degree of non-homotheticity in preferences over positional goods. Moreover, under reasonable parameters describing the magnitude of positional considerations, my numerical calculations show that the flat tax imposed on the consumption of positional goods is by no means negligible, though it is highly dependent on the structural parameters of the economy.

The rest of the paper proceeds as follows. Section 3.2 presents the model and shows the characterization of constrained efficient allocations, section 3.3 presents the characterization and one implementation of double constrained efficient allocations, section 3.4 presents calculations of the endogenous distributions and optimal taxes for a parameterized version of the model and section 3.5 concludes.

3.2 The Model

Consider a static economy populated by a continuum of agents with heterogeneous productivity or skill. Let $\theta \in \Theta$, where $\Theta \equiv [\underline{\theta}, \bar{\theta}]$ and $0 < \underline{\theta} < \bar{\theta} < \infty$, be an individual's productivity distributed according to the density $f : \Theta \to \mathbf{R}_{++}$. There are two consumption goods and leisure. Since the idea is to model goods with different degrees of positionality, I will assume that one of the consumption goods is a necessity while the other is a luxury. Productivity is *privately* known to each agent. An agent with productivity θ has a utility function of the form

$$U(c_n, c_l, y, C; \theta) = u(c_n, c_l) - \alpha C - v\left(\frac{y}{\theta}\right), \quad \alpha \in [0, 1)$$

where c_n is a necessity, c_l is a luxury good and y is effective output.[2] Moreover, let

$$C \equiv \int_{\Theta} [\omega c_n(\theta) + (1 - \omega) c_l(\theta)] \psi(\theta) d\theta, \quad \omega \in [0, 1/2) \tag{3.1}$$

be society's endogenous consumption benchmark specified as a *weighted* average of necessities and luxuries. As usual, preferences satisfy $u_{c_n} > 0$, $u_{c_l} > 0$, $u(\cdot)$ is jointly strictly concave and $v(\cdot)$ is a convex function.[3] Also, observe that according to the previous utility specification, $u_C = -\alpha$. Thus, following the terminology of Dupor and Liu

[2] As standard in this literature, I define effective labor as $y = \theta l$ where l is the amount of time worked.

[3] Later in the text I will introduce a functional form that properly captures good c_l being a luxury in terms of the income elasticity. For the moment, it is enough to think that a luxury is simply more positional than a necessity. In other words, the consumption externalities generated by the consumption of the luxury are of greater magnitude than those of the necessity.

(2003), agents exhibit jealousy. Notice that with the assumption that $\omega < 1/2$, I capture the notion that luxuries provoke more jealousy than necessities as claimed by Frank (2008).[4] In other words, luxuries are more *positional* than necessities. Obviously, when $\omega = 0$, only luxuries are positional.

An *allocation* in this economy is $\{c_n(\theta), c_l(\theta), y(\theta)\}_{\theta\in\Theta}$, where $c_n : \Theta \to \mathbf{R}_+$, $c_l : \Theta \to \mathbf{R}_+$ and $y : \Theta \to \mathbf{R}_+$. Abstracting from government expenditure, I define an allocation $\{c_n(\theta), c_l(\theta), y(\theta)\}_{\theta\in\Theta}$ to be *feasible* if

$$\int_\Theta c_n(\theta) f(\theta) d\theta + \int_\Theta c_l(\theta) f(\theta) d\theta = \int_\Theta y(\theta) f(\theta) d\theta \tag{3.2}$$

Observe that in the previous definition I am assuming that both consumption goods are substitutes in production. This assumption is made for simplicity. A *reporting strategy* is a mapping $\sigma : \Theta \to \Theta$, where $\sigma(\theta)$ represents the skill announced by an agents with skill θ in a direct revelation game. An agents reports her type truthfully if $\sigma(\theta) = \theta$. Thus, making use of the Revelation Principle, an allocation is *incentive compatible* if

$$u(c_n(\theta), c_l(\theta)) - \alpha C - v\left(\frac{y(\theta)}{\theta}\right) \geq u(c_n(\sigma(\theta)), c_l(\sigma(\theta))) - \alpha C - v\left(\frac{y(\sigma(\theta))}{\theta}\right) \forall \theta, \sigma(\theta) \in \Theta \tag{3.3}$$

Observe that since C cannot be affected unilaterally by a single agent, it is not a function of θ. An allocation that is incentive compatible and feasible is said to be

[4]Empirical evidence of this fact is also presented in Carlsson, Johansson-Stenman, and Martinsson (2007) and Solnick and Hemenway (2005).

incentive-feasible. Finally, let $g : \Theta \to \mathbf{R}_+$ be the density according to which individuals are weighted by the benevolent planner.

Definition 4. *A constrained efficient allocation is any allocation* $\{c_n^{sp}(\theta), c_l^{sp}(\theta), y^{sp}(\theta)\}_{\theta\in\Theta}$ *that maximizes the following planner problem*

$$\int_\Theta \left[u(c_n(\theta), c_l(\theta)) - \alpha C - v\left(\frac{y(\theta)}{\theta}\right)\right] g(\theta) d\theta \tag{3.4}$$

subject to $\{c_n(\theta), c_l(\theta), y(\theta)\}_{\theta\in\Theta}$ *being incentive-feasible,* C *as defined in (3.1) and* $c_n(\theta), c_l(\theta), y(\theta) \geq 0 \quad \forall\theta \in \Theta$.

3.2.1 Characterization of Constrained Efficient Allocations

The following proposition expresses the necessary conditions that any *interior* constrained efficient allocation must satisfy. Let $\epsilon^{sp}(\theta) \equiv \frac{v'(\frac{y^{sp}(\theta)}{\theta})}{v''(\frac{y^{sp}(\theta)}{\theta})\frac{y^{sp}(\theta)}{\theta}}$.

Proposition 6. *Any interior constrained efficient allocation* $\{c_n^{sp}(\theta), c_l^{sp}(\theta), y^{sp}(\theta)\}_{\theta\in\Theta}$ *must be incentive-feasible and satisfy*

$$\frac{u_{c_n}(c_n^{sp}(\theta), c_l^{sp}(\theta))}{v'(\frac{y^{sp}(\theta)}{\theta})\frac{1}{\theta}} - 1 = \frac{\alpha}{\lambda}\frac{\omega\psi(\theta)}{f(\theta)} + \frac{u_{c_n}(c_n^{sp}(\theta), c_l^{sp}(\theta))}{\theta f(\theta)}\left[1 + \frac{1}{\epsilon^{sp}(\theta)}\right] I^{sp}(\theta) \quad \forall\theta \in \Theta \tag{3.5}$$

$$\frac{u_{c_l}(c_n^{sp}(\theta), c_l^{sp}(\theta))}{u_{c_n}(c_n^{sp}(\theta), c_l^{sp}(\theta))} = \frac{1 + \frac{\alpha(1-\omega)}{\lambda}\frac{\psi(\theta)}{f(\theta)}}{1 + \frac{\alpha\omega}{\lambda}\frac{\psi(\theta)}{f(\theta)}} \quad \forall\theta \in \Theta \tag{3.6}$$

where

$$\lambda = \frac{1 - \omega\alpha \int_{\Theta} \frac{\psi(\theta)}{u_{c_n}(c_n^{sp}(\theta), c_l^{sp}(\theta))} d\theta}{\int_{\Theta} \frac{f(\theta)}{u_{c_n}(c_n^{sp}(\theta), c_l^{sp}(\theta))} d\theta} \tag{3.7}$$

$$I^{sp}(\theta) \equiv \int_{\underline{\theta}}^{\theta} \left[\frac{g(t)}{\lambda} - \frac{f(t)}{u_{c_n}(c_n^{sp}(t), c_l^{sp}(t))} - \frac{\alpha}{\lambda} \frac{\omega\psi(t)}{u_{c_n}(c_n^{sp}(t), c_l^{sp}(t))} \right] dt \quad \forall \theta \in \Theta \tag{3.8}$$

Proof. See Appendix B. □

According to Proposition 6, the marginal rate of substitution (MRS) between the necessity and the luxury varies across agents if $\psi(\theta)$ is different from $f(\theta)$ as observed in expression 3.6. For the sake of exposition, consider the case where $\omega = 0$, that is, only the luxury good generates positional externalities. Moreover, assume that the ratio $\frac{\psi(\theta)}{f(\theta)}$ is strictly increasing, that is, the consumption of more affluent individuals is more harmful to the society as a whole. In this case, it is optimal to have an increasing MRS between the luxury and the necessity as income goes up. The previous argument breaks down in two cases: either when $\psi(\theta)$ equals $f(\theta)$ and when $\alpha = 0$. In both cases, the MRS across consumption goods is constant across agents. In the second case, under no utility interdependence, the MRS across goods equals one (marginal rate of transformation given the assumed technology) so the uniform commodity taxation result holds.

From a theoretical point of view, the case where the MRS is non-constant across agents is more challenging to analyze. The reason for this is that if the planner cannot observe agents' consumptions, individuals could meet in a re-trading market after being assigned their consumption bundles and exchange consumption goods at a given price. In this re-trading market, they all would end up equalizing their MRS between the luxury and the necessity to an equilibrium relative price. To better illustrate this, suppose that the non-constant wedge were to be implemented by a non-linear tax on the consumption of the positional good or "luxury tax". Then, individuals could enter into "non-exclusive" arrangements to exchange luxuries for necessities until all agents equalized their respective MRS's.[5] In order to take into account the previous fact, we need to refine the notion of constrained efficiency in this environment. Before formally stating this, we need a few definitions.

3.3 Double Constrained Efficient Allocations

I start by posing the agent's problem in the re-trading market contingent on having announced being of $\sigma(\theta)$ type.

3.3.1 Agent's Problem

Given allocations $\{c_n(\theta), c_l(\theta), y(\theta)\}_{\theta\in\Theta}$, externality X and price q, an agent who decides to report $\sigma(\theta)$ attains utility

$$V(\{c_n(\theta), c_l(\theta), y(\theta)\}_{\theta\in\Theta}, X, q \mid \sigma(\theta)) =$$

[5]The term "non-exclusivity" is used to emphasize that agents are not constrained to trade with one single partner.

$$\max_{x_n(\sigma(\theta)),x_l(\sigma(\theta))} u(x_n(\sigma(\theta)), x_l(\sigma(\theta))) - \alpha X - v\left(\frac{y(\sigma(\theta))}{\theta}\right) \tag{3.9}$$

s.t.

$$x_n(\sigma(\theta)) + qx_l(\sigma(\theta)) \leq c_n(\sigma(\theta)) + qc_l(\sigma(\theta))$$

$$x_n(\sigma(\theta)), x_l(\sigma(\theta)) \geq 0$$

where $x_n(\sigma(\theta))$ and $x_l(\sigma(\theta))$ represent the private consumption of the necessity and the luxury respectively, q is the relative price of the luxury good and $X \equiv \int_\Theta [\omega x_n(\sigma(\theta)) + (1-\omega)x_l(\sigma(\theta))]\psi(\theta)d\theta$. Moreover, I define

$$V(\{c_n(\theta), c_l(\theta), y(\theta)\}_{\theta\in\Theta}, X, q) \equiv \max_{\sigma(\theta)\in\Theta} V(\{c_n(\theta), c_l(\theta), y(\theta)\}_{\theta\in\Theta}, X, q \mid \sigma(\theta)) \tag{3.10}$$

which represents the utility level attained by optimally announcing $\sigma(\theta)$, given q, the externality X and $\{c_n(\theta), c_l(\theta), y(\theta)\}_{\theta\in\Theta}$.

3.3.2 Equilibrium in the Re-Trading Market

An *equilibrium* in the re-trading market is allocations $\{x_n(\sigma(\theta)), x_l(\sigma(\theta))\}_{\theta\in\Theta}$, strategies $\{\sigma(\theta)\}_{\theta\in\Theta}$ and a price q such that

i) Taking as given $\{c_n(\theta), c_l(\theta), y(\theta)\}_{\theta\in\Theta}$, X and q, agents solve (3.9) and (3.10),

ii) Re-trading market clears

$$\int_\Theta x_n(\sigma(\theta))f(\theta)d\theta = \int_\Theta c_n(\sigma(\theta))f(\theta)d\theta$$

$$\int_{\Theta} x_l(\sigma(\theta))f(\theta)d\theta = \int_{\Theta} c_l(\sigma(\theta))f(\theta)d\theta \tag{3.11}$$

Let $\hat{V}(\{c_n(\theta), c_l(\theta), y(\theta)\}_{\theta\in\Theta})$ equals $V(\{c_n(\theta), c_l(\theta), y(\theta)\}_{\theta\in\Theta}, \hat{X}, \hat{q})$ where $\hat{q}$ is the equilibrium price and $\hat{X}$ is the value of the externality at that price. Given the previous definitions, we are in a position to define efficiency in this environment.

Definition 5. *Any double constrained efficient allocation is any allocation $\{c_n^*(\theta), c_l^*(\theta), y^*(\theta)\}_{\theta\in\Theta}$ that maximizes the following planner problem*

$$\int_{\Theta} \left[u(c_n(\theta), c_l(\theta)) - \alpha C - v\left(\frac{y(\theta)}{\theta}\right)\right] g(\theta)d\theta \tag{3.12}$$

s. t.

$$u(c_n(\theta), c_l(\theta)) - \alpha C - v\left(\frac{y(\theta)}{\theta}\right) \geq \hat{V}(\{c_n(\theta), c_l(\theta), y(\theta)\}_{\theta\in\Theta}) \tag{3.13}$$

$$\int_{\Theta} c_n(\theta)f(\theta)d\theta + \int_{\Theta} c_l(\theta)f(\theta)d\theta = \int_{\Theta} y(\theta)f(\theta)d\theta \tag{3.14}$$

where C is defined as in (3.1) and $c_n(\theta), c_l(\theta), y(\theta) \geq 0 \quad \forall\theta \in \Theta$.

Notice that the previous notion of constrained efficiency takes explicitly into account the re-trading market for consumption goods across agents as a constraint. Lemma 1 establishes an equivalence statement of the problem stated in Definition 5.

Lemma 1. *Any double constrained efficient allocation $\{c_n^*(\theta), c_l^*(\theta), y^*(\theta)\}_{\theta\in\Theta}$ together* *with the equilibrium relative price of luxuries q is a solution to the planner problem*

$$\max_{c_n(\cdot),c_l(\cdot),y(\cdot),q} \int_{\Theta} \left[u(c_n(\theta), c_l(\theta)) - \alpha C - v\left(\frac{y(\theta)}{\theta}\right) \right] g(\theta) d\theta \tag{3.15}$$

s. t.

$$\frac{u_{c_l}(c_n(\theta), c_l(\theta))}{u_{c_n}(c_n(\theta), c_l(\theta))} = q \quad \forall \theta \in \Theta \tag{3.16}$$

$$u(c_n(\theta), c_l(\theta)) - \alpha C - v\left(\frac{y(\theta)}{\theta}\right) \geq V(\{c_n(\theta), c_l(\theta), y(\theta)\}_{\theta \in \Theta}, X, q \mid \sigma(\theta)) \quad \forall \sigma(\theta) \neq \theta \tag{3.17}$$

$$\int_{\Theta} c_n(\theta) f(\theta) d\theta + \int_{\Theta} c_l(\theta) f(\theta) d\theta = \int_{\Theta} y(\theta) f(\theta) d\theta \tag{3.18}$$

where C is defined as in (3.1) and $c_n(\theta), c_l(\theta), y(\theta) \geq 0 \quad \forall \theta \in \Theta$.

Proof. The proof follows closely da Costa (2009). Let q be the equilibrium relative price associated with the allocation $\{c_n(\theta), c_l(\theta), y(\theta)\}_{\theta \in \Theta}$ and assume that (3.13) is satisfied. If (3.16) is violated, there is an alternative consumption choice that increases the agents utility holding strategy $\sigma(\theta) = \theta$ fixed. Because q is an equilibrium price, this violates (3.13). To see that (3.17) holds, observe that $\hat{V}(\{c_n(\theta), c_l(\theta), y(\theta)\}_{\theta \in \Theta})$ is simply (3.10) at equilibrium price. Hence, there is no strategy $\sigma(\theta)$ that yields higher utility that truth-telling.

Now assume (3.16) and (3.17) are satisfied. When the price is q, agents find it in their best interest to reveal their types truthfully, according to (3.17). This means that there is no strategy $\sigma(\theta)$ combined with optimal re-trading that increases utility for the agents at that price. Equation (3.16) implies that no-trade is optimal for the agent at the very same price q. No-trade trivially satisfies (3.11) which guarantees that q is indeed an equilibrium price. Therefore, (3.17) implies (3.13). $\square$

Lemma 1 states that if individual allocations are such that the MRS between consumption goods is *constant* across agents, in equilibrium, re-trading does not occur. Observe that the optimal MRS between luxuries and necessities is yet to be determined.

Theorem 2. *Any double constrained efficient allocation* $\{c_n^*(\theta), c_l^*(\theta), y^*(\theta)\}_{\theta\in\Theta}$ *together with the equilibrium relative price of luxuries* q *solve the following planner problem*

$$\max_{c_n(\cdot),c_l(\cdot),y(\cdot),q} \int_\Theta \left[u(c_n(\theta), c_l(\theta)) - \alpha C - v\left(\frac{y(\theta)}{\theta}\right)\right] g(\theta)d\theta \tag{3.19}$$

s. t.

$$\frac{u_{c_l}(c_n(\theta), c_l(\theta))}{u_{c_n}(c_n(\theta), c_l(\theta))} = q \quad \forall\theta \in \Theta, \tag{3.20}$$

$$u(c_n(\theta), c_l(\theta)) - \alpha C - v\left(\frac{y(\theta)}{\theta}\right) \geq u(c_n(\sigma(\theta)), c_l(\sigma(\theta))) - \alpha C - v\left(\frac{y(\sigma(\theta))}{\theta}\right) \forall\theta, \sigma(\theta) \in \Theta \tag{3.21}$$

$$\int_\Theta c_n(\theta)f(\theta)d\theta + \int_\Theta c_l(\theta)f(\theta)d\theta = \int_\Theta y(\theta)f(\theta)d\theta \tag{3.22}$$

where C *is defined as in (3.1) and* $c_n(\theta), c_l(\theta), y(\theta) \geq 0 \quad \forall\theta \in \Theta.$

Proof. Suppose that a deviating agent announces $\sigma(\theta) = \theta'$. Hence, she receives $c_n(\theta'), c_l(\theta')$. Suppose she re-trades, thus $[x_n(\theta'), x_l(\theta')] \neq [c_n(\theta'), c_l(\theta')]$. However, truthful agent θ' and the deviating agent impersonating θ' who reported $\sigma(\theta) = \theta'$ have the same income in the re-trading market. By strict concavity of preferences it must be the case that $[x_n(\theta'), x_l(\theta')] = [c_n(\theta'), c_l(\theta')]$, otherwise there would exist another bundle strictly preferred to $[c_n(\theta'), c_l(\theta')]$, implying that the previous bundle is not optimal for truthful agent θ'. Thus, we arrive at a contradiction. This result implies that agents do not re-trade in equilibrium and thus $X = C$. Having established this fact, it follows that $V(\{c_n(\theta), c_l(\theta), y(\theta)\}, X, q \mid \sigma(\theta)) = u(c_n(\sigma(\theta)), c_l(\sigma(\theta))) - \alpha C - v\left(\frac{y(\sigma(\theta))}{\theta}\right)$ which together with (3.21) implies that (3.17) is satisfied. The other side of the proof is trivially satisfied given that agents do not re-trade. $\square$

3.3.3 Characterization of Double Constrained Efficient Allocations

The following proposition expresses the necessary conditions that any *interior double constrained efficient allocation* must satisfy. Let $\epsilon^*(\theta) \equiv \frac{v'(\frac{y^*(\theta)}{\theta})}{v''(\frac{y^*(\theta)}{\theta})\frac{y^*(\theta)}{\theta}}$.

Proposition 7. *Any interior double constrained efficient allocation* $\{c_n^*(\theta), c_l^*(\theta), y^*(\theta)\}_{\theta\in\Theta}$ *must be incentive-feasible and satisfy*

$$\frac{u_{c_n}(c_n^*(\theta), c_l^*(\theta))}{v'(\frac{y^*(\theta)}{\theta})\frac{1}{\theta}} - 1 = \frac{\alpha}{\lambda}\frac{\omega\psi(\theta)}{f(\theta)} + \frac{u_{c_n}(c_n^*(\theta), c_l^*(\theta))}{\theta f(\theta)}\left[1 + \frac{1}{\epsilon^*(\theta)}\right] I^*(\theta) \quad \forall\theta \in \Theta \quad (3.23)$$

$$\frac{u_{c_l}(c_n^*(\theta), c_l^*(\theta))}{u_{c_n}(c_n^*(\theta), c_l^*(\theta))} = \frac{\int_\Theta \frac{f(\theta)}{B^*(\theta)}d\theta + \frac{\alpha(1-\omega)}{\lambda}\int_\Theta \frac{\psi(\theta)}{B^*(\theta)}d\theta}{\int_\Theta \frac{f(\theta)}{B^*(\theta)}d\theta + \frac{\alpha\omega}{\lambda}\int_\Theta \frac{\psi(\theta)}{B^*(\theta)}d\theta} \quad \forall\theta \in \Theta \quad (3.24)$$

where

$$\lambda = \frac{1 - \omega\alpha \int_{\Theta} \frac{\psi(\theta)}{u_{cn}(c_n^*(\theta), c_l^*(\theta))} d\theta}{\int_{\Theta} \frac{f(\theta)}{u_{cn}(c_n^*(\theta), c_l^*(\theta))} d\theta} \tag{3.25}$$

$$B^*(\theta) \equiv \frac{u_{c_n c_l}(c_n^*(\theta), c_l^*(\theta))}{u_{c_n}(c_n^*(\theta), c_l^*(\theta))} - \frac{u_{c_l c_l}(c_n^*(\theta), c_l^*(\theta))}{u_{c_l}(c_n^*(\theta), c_l^*(\theta))} \quad \forall \theta \in \Theta \tag{3.26}$$

$$I^*(\theta) \equiv \int_{\underline{\theta}}^{\theta} \left[\frac{g(t)}{\lambda} - \frac{f(t)}{u_{c_n}(c_n^*(t), c_l^*(t))} - \frac{\alpha}{\lambda} \frac{\omega\psi(t)}{u_{c_n}(c_n^*(t), c_l^*(t))} \right] dt \quad \forall \theta \in \Theta \tag{3.27}$$

Proof. See Appendix B.1. □

3.3.4 Implementation of Double Constrained Efficient Allocations

Agents in this economy trade effective labor for consumption of the necessity and the luxury. There is a single firm that employs all agents. It produces one unit of output for every unit of effective labor, y. Necessities and luxuries are perfect substitutes in production. Every unit of effective labor receives a payment of w. Agents are also subject to an income tax schedule $T(y(\theta))$, assumed to be twice differentiable and to induce no bunching and a *linear* tax on the luxury good τ. Without loss of generality, there are no taxes on the consumption of the necessity c_n. An agent with effective labor y pays $T(y(\theta))$ of taxes.

Thus, taking as given $T(y(\theta))$, τ, C and the wage w, the problem solved by the agent with productivity $\theta, \forall \theta \in \Theta$ is

$$\max_{c_n(\theta), c_l(\theta), y(\theta)} u(c_n(\theta), c_l(\theta)) - \alpha C - v\left(\frac{y(\theta)}{\theta}\right) \tag{3.28}$$

s.t.

$$c_n(\theta) + (1+\tau)c_l(\theta) \leq wy(\theta) - T(y(\theta))$$

$$c_n(\theta), c_l(\theta), y(\theta) \geq 0$$

Definition 6. *Given a labor tax $T(y(\theta))$, luxury tax τ and C, an equilibrium in this economy is an allocation $\{c_n^{eq}(\theta), c_l^{eq}(\theta), y^{eq}(\theta)\}_{\theta \in \Theta}$ and wage w^{eq} such that*

i. $\left(c_n^{eq}(\theta), c_l^{eq}(\theta), y^{eq}(\theta)\right)$ *solve (3.28)* $\forall \theta \in \Theta$

ii. $C = \int_\Theta [\omega c_n^{eq}(\theta) + (1-\omega) c_l^{eq}(\theta)] \psi(\theta) d\theta$

iii. $w^{eq} = 1$

iv. Government balances its budget

$$\int_\Theta [T(y^{eq}(\theta)) + \tau c_l^{eq}(\theta)] f(\theta) d\theta = 0$$

v. $\int_\Theta c_n^{eq}(\theta) f(\theta) d\theta + \int_\Theta c_l^{eq}(\theta) f(\theta) d\theta = \int_\Theta y^{eq}(\theta) f(\theta) d\theta$

An allocation $\{c_n(\theta), c_l(\theta), y(\theta)\}_{\theta \in \Theta}$ is *implementable* by the income tax $T(y(\theta))$ and the luxury tax τ if $\{c_n(\theta), c_l(\theta), y(\theta)\}_{\theta \in \Theta}$ and w are an equilibrium.

3.3.5 Characterization of Optimal Income Tax and Linear Positional or "Luxury" Tax

Define the following tax mechanism $T : y \to \mathbf{R}$,

$$T(y(\theta)) = \begin{cases} y^*(\theta) - c_n^*(\theta) - (1+\tau)c_l^*(\theta) & \text{if } y(\theta) = y^*(\theta) \\ y(\theta) & \text{otherwise.} \end{cases} \tag{3.29}$$

where

$$\tau = \frac{\frac{\alpha(1-2\omega)}{\lambda}\int_\Theta \frac{\psi(\theta)}{B^*(\theta)}d\theta}{\int_\Theta \frac{f(\theta)}{B^*(\theta)}d\theta + \frac{\alpha\omega}{\lambda}\int_\Theta \frac{\psi(\theta)}{B^*(\theta)}d\theta} \tag{3.30}$$

together with

$$\frac{T'(y^*(\theta))}{1-T'(y^*(\theta))} = \frac{\alpha}{\lambda}\frac{\omega\psi(\theta)}{f(\theta)} + \frac{u_{c_n}(c_n^*(\theta), c_l^*(\theta))}{\theta f(\theta)}\left[1 + \frac{1}{\epsilon^*(\theta)}\right] I^*(\theta) \tag{3.31}$$

if $y(\theta) = y^*(\theta)$.

Proposition 8. *Any interior double constrained efficient allocation* $\{c_n^*(\theta), c_l^*(\theta), y^*(\theta)\}$ *can be implemented by an income tax schedule* $T(y(\theta))$ *defined by (3.29) and (3.31) and a flat tax* τ *on the positional good satisfying (3.30).*

Proof. See Appendix B.1. □

Corollary 2 (Proposition 8)**.** *Suppose* $\omega = 0$, $u(c_n, c_l) = \left[\eta c_n^{1-\frac{1}{\sigma}} + (1-\eta)c_l^{1-\frac{\rho}{\sigma}}\right]^{\frac{\sigma}{\sigma-1}}$, $\sigma < \eta, \eta \leq 1$, *then the optimal marginal income tax and luxury tax satisfy*

$$\frac{T'(y^*(\theta))}{1-T'(y^*(\theta))}=\frac{u_{c_n}(c_n^*(\theta),c_l^*(\theta))}{\theta f(\theta)}\left[1+\frac{1}{\epsilon^*(\theta)}\right]\int_{\underline{\theta}}^{\theta}\left[\frac{g(t)}{\lambda}-\frac{f(t)}{u_{c_n}(c_n^*(\theta),c_l^*(\theta))}\right]dt \quad (3.32)$$

$$\tau=\frac{\alpha}{\lambda}\frac{\int_{\Theta}c_l^*(\theta)\psi(\theta)d\theta}{\int_{\Theta}c_l^*(\theta)f(\theta)d\theta} \quad (3.33)$$

where $\lambda=\frac{1}{\int_{\Theta}\frac{f(\theta)}{u_{c_n}(c_n^*(\theta),c_l^*(\theta))}d\theta}$ *and*

$$u_{c_n}(c_n^*(\theta),c_l^*(\theta))\equiv\left[\eta c_n^*(\theta)^{1-\frac{1}{\sigma}}+(1-\eta)c_l^*(\theta)^{1-\frac{\rho}{\sigma}}\right]^{\frac{1}{\sigma-1}}\eta c_n^*(\theta)^{-\frac{1}{\sigma}}.$$

It is important to highlight that the results of Corollary 2 hold with or without homothetic preferences since the utility function (known as *generalized* elasticity of substitution and introduced in Pakos (2006)) becomes homothetic when $\rho=1$.[6] In section 3.4, I compute constrained and double constrained efficient allocations and the taxes that implement them using the previous functional form. Corollaries 3 and 4 show expressions for optimal taxes under another class of preferences standard in the literature.

[6]To observe this, notice that

$$\frac{\partial c_l}{\partial y}\frac{y}{c_l}=\frac{\left[\frac{(\sigma-\rho)}{(\sigma-1)}\frac{(1-\eta)}{\eta}\right]^{-\sigma}c_l^{\rho}+qc_l}{\rho\left[\frac{(\sigma-\rho)}{(\sigma-1)}\frac{(1-\eta)}{\eta}\right]^{-\sigma}c_l^{\rho}+qc_l}$$

where q is the relative price of the luxury good in terms of the necessity. Thus, $\frac{\partial c_l}{\partial y}\frac{y}{c_l}=1$ if $\rho=1$ and $\frac{\partial c_l}{\partial y}\frac{y}{c_l}>1$ if $\rho<1$. In the latter case, these preferences properly represent the good c_l as a luxury.

Corollary 3 (Proposition 8). *Suppose* $\omega = 0$*,* $u(c_n, c_l) = \frac{1}{1-\sigma}{c_n}^{1-\sigma} + \frac{1}{1-\rho}{c_l}^{1-\rho}$*, where* $\sigma, \rho > 0$ *then the optimal marginal income tax and luxury tax satisfy*

$$\frac{T'(y^*(\theta))}{1 - T'(y^*(\theta))} = \frac{c_n^*(\theta)^{-\sigma}}{\theta f(\theta)} \left[1 + \frac{1}{\epsilon^*(\theta)}\right] \int_{\underline{\theta}}^{\theta} \left[\frac{g(t)}{\lambda} - \frac{f(t)}{c_n^*(t)^{-\sigma}}\right] dt \tag{3.34}$$

$$\tau = \frac{\alpha}{\lambda} \frac{\int_\Theta c_l^*(\theta)\psi(\theta)d\theta}{\int_\Theta c_l^*(\theta) f(\theta) d\theta} \tag{3.35}$$

where $\lambda = \frac{1}{\int_\Theta c_n^*(\theta)^\sigma f(\theta)d\theta}$.

Corollary 4 (Proposition 8). *Suppose* $\omega = 0$*,* $u(c_n, c_l) = \frac{1}{1-\sigma}{c_n}^{1-\sigma} - e^{-\rho c_l}$*, where* $\sigma, \rho > 0$ *then the optimal marginal income tax and luxury tax satisfy*

$$\frac{T'(y^*(\theta))}{1 - T'(y^*(\theta))} = \frac{c_n^*(\theta)^{-\sigma}}{\theta f(\theta)} \left[1 + \frac{1}{\epsilon^*(\theta)}\right] \int_{\underline{\theta}}^{\theta} \left[\frac{g(t)}{\lambda} - \frac{f(t)}{c_n^*(t)^{-\sigma}}\right] dt \tag{3.36}$$

$$\tau = \frac{\alpha}{\lambda} \tag{3.37}$$

where $\lambda = \frac{1}{\int_\Theta c_n^*(\theta)^\sigma f(\theta)d\theta}$.

3.4 Welfare Losses and Gains Due to Linear Taxation on Positional Goods

In this section I compute the model constrained efficient and double constrained efficient allocations and evaluate the welfare in both environments. I also compute the taxes that implement these allocations. For this quantitative exercise, I assume that preferences are represented by $U(c_n, c_l, C, y; \theta) \equiv u(c_n, c_l) - \alpha C - \frac{1}{1+\phi}\left(\frac{y}{\theta}\right)^{1+\phi}$, $\phi > 0$, where

$$u(c_n, c_l) = \left[\eta c_n^{1-\frac{1}{\sigma}} + (1-\eta) c_l^{1-\frac{\rho}{\sigma}}\right]^{\frac{\sigma}{\sigma-1}},$$

with $\rho \leq 1$ and $\sigma < 1$ with $\sigma < \rho$. For $\rho < 1$, this specification of preferences exhibits *non-homotheticity*. That is, as income of individuals goes up, the share of disposable income spent on the luxury good increases.[7] As closely estimated by Pakos (2006), I set $\sigma = 0.5$ and vary the parameter that measures the ratio of income elasticities between necessities and luxuries, ρ, to take values in $\{0.80, 0.91, 1\}$.[8] Recall that when $\rho = 1$, preferences exhibit homotheticity. I set $\eta = 0.75$ and $\phi \in \{0.5, 1.5, 3\}$.[9] The latter implies an uncompensated elasticity of labor supply of 2, 1/2 and 1/3, respectively. I assume that the support of the distribution is $\Theta = [1, \infty)$, although for computational purposes I use the domain $[1, 6]$. I consider two cases for the exogenous distributions of

[7] Ait-Sahalia, Parker, and Yogo (2004) also model consumption as a composite of necessities and luxuries using non-homothetic preferences.

[8] The value of ρ is also in line with the work of Costa (2001) who estimates income elasticities for food and recreation in the U.S. up to 1994.

[9] In Appendix B.4 I also present an exercise where $\phi = 0.2$ which implies a uncompensated labor supply elasticity of 5!

the model. In the first one, skills density $f(\theta)$ is distributed according to a truncated *exponential* distribution with parameter $\lambda_f = 1$, whereas $g(\theta)$ and $\psi(\theta)$ follow the same distribution with parameters $\lambda_g = 1$ and $\lambda_\psi = 1.25$, respectively.[10] Notice that I assume that the planner is utilitarian since $\lambda_g = \lambda_f$. In the second case, skills are distributed according to a Pareto distribution with parameters k_f, k_g and k_ψ, respectively. I set $k_f = k_g = 1.8$ and $k_\psi = 1.08$. As before, the planner is utilitarian. These distributions are plotted in Figure 3.1.

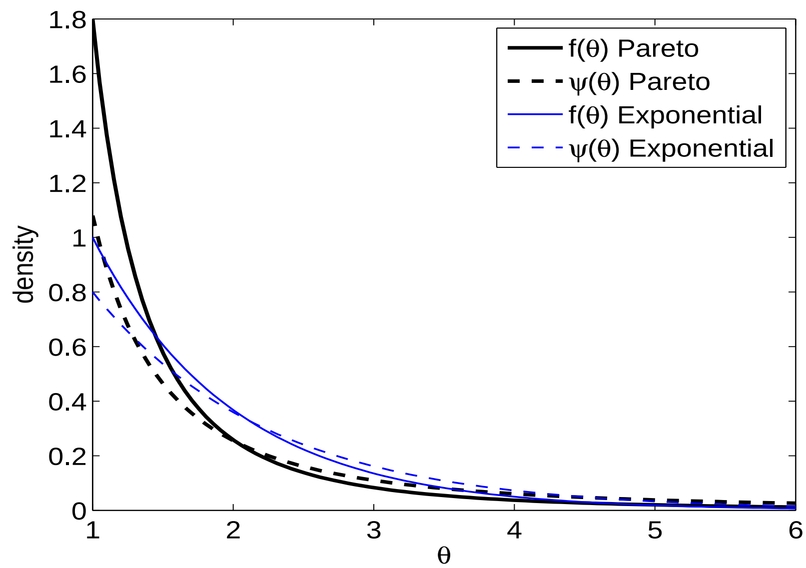

FIGURE 3.1: Skills and Externality Weighting Distributions

I present my calculations for $\omega = 0$. That is, in this economy, only luxuries generate positional externalities. I let the parameter $\alpha \in \{0.05, 0.15\}$.[11] Welfare losses are

[10] Recall that the exponential distribution is $f(\theta) = \lambda e^{-\lambda\theta}, \theta \in [0, \infty)$. Thus, a lower tail truncated exponential distribution at $\theta = \underline{\theta}$ is $f^t(\theta) = \frac{1}{1-F(\underline{\theta})} f(\theta)$ where $F(\theta) = 1 - e^{-\lambda\theta}$.

[11] Samano (2009) obtains a value of α for aggregate consumption close to 0.15 using U.S. and U.K.

measured as in Golosov and Tsyvinski (2007). That is, let

$$U_{\Theta}^{sp} \equiv \int_{\Theta} \left[u(c_n^{sp}(\theta), c_l^{sp}(\theta)) - \alpha C^{sp} - \frac{1}{1+\phi} \left(\frac{y^{sp}(\theta)}{\theta} \right)^{1+\phi} \right] g(\theta) d\theta,$$

thus I find the value λ_{Θ} such that

$$\int_{\Theta} \left[u((1+\lambda_{\Theta}) c_n^{*}(\theta), (1+\lambda_{\Theta}) c_l^{*}(\theta)) - \alpha C^{*} - \frac{1}{1+\phi} \left(\frac{y^{*}(\theta)}{\theta} \right)^{1+\phi} \right] g(\theta) d\theta = U_{\Theta}^{sp}.$$

In other words, I find the percentage increase in aggregate consumption λ_{Θ} that would deliver the same aggregate welfare in the double constrained efficient allocations as in the constrained efficient ones. The subindex Θ in the parameter λ makes explicit that welfare comparisons are made for *all* agents in the economy. To quantify *non-aggregate* welfare effects, I also calculate the value of $\lambda_L \equiv \lambda_{\{[\underline{\theta},\tilde{\theta}] \mid F(\tilde{\theta})=0.90\}}$ and $\lambda_H \equiv \lambda_{\{[\tilde{\theta},\bar{\theta}] \mid F(\tilde{\theta}=0.90)\}}$.[12]

Tables 3.1 and 3.2 show the optimal linear positional or "luxury tax" rate and welfare losses (and gains, whenever the variable λ is negative) for several parameters when skills are distributed according to an exponential distribution. Tables 3.3 and 3.4 show the same information when skills are Pareto distributed. As expected, the double

income data. Dynan and Ravina (2007) obtain values close to 0.10.

[12] To be precise, λ_L is calculated as

$$\int_{\underline{\theta}}^{\tilde{\theta}} \left[u((1+\lambda_L) c_n^{*}(\theta), (1+\lambda_L) c_l^{*}(\theta)) - \alpha C^{*} - \frac{1}{1+\phi} \left(\frac{y^{*}(\theta)}{\theta} \right)^{1+\phi} \right] g(\theta) d\theta = U_L^{sp},$$

where $U_L^{sp} \equiv \int_{\underline{\theta}}^{\tilde{\theta}} \left[u(c_n^{sp}(\theta), c_l^{sp}(\theta)) - \alpha C^{sp} - \frac{1}{1+\phi} \left(\frac{y^{sp}(\theta)}{\theta} \right)^{1+\phi} \right] g(\theta) d\theta$ and $F(\tilde{\theta}) = 0.90$. λ_H is calculated similarly.

constrained environment delivers lower aggregate welfare given that an extra constraint is imposed, namely that MRS between goods is constant across agents.

My calculations indicate that the *aggregate* welfare losses suffered by taxing the positional good in a linear fashion are very low, particularly when the labor supply is reasonably inelastic. Thus a linear "luxury tax" does almost as well as a non-linear one in terms of aggregate welfare. Nevertheless, this policy has considerable distributional effects. Assuming that the positional externality is increasing in income, individuals at the high end of the income distribution experience large *gains* under linear taxation of the positional good. The reason for this is as follows: a flat luxury tax effectively reduces the price of luxuries for rich individuals. The drop in the price generates a positive income effect that cannot be offset by an increase in the marginal income tax as optimality requires no distortion at the top. Consequently, the consumption of agents at the top increases. When preferences are non-homothetic, the gains of highly skilled individuals are reduced, but to negligible levels as the income effect generated by the drop in the price of luxuries is smaller. In this case, a small adjustment in the labor income tax is necessary to extract output from individuals at the high end of the income distribution. The opposite is true for individuals at the bottom of the skills distribution. They experience considerable welfare losses since a flat luxury tax increases the after-tax price of luxuries. This generates a negative income effect that is offset by a reduction in the marginal income tax. Such a reduction cannot fully offset the income effect since this would violate incentives. If preferences are non-homothetic, the adjustment in the income tax is smaller than in the homothetic case.

Observe that for $\alpha = 0.15$, the luxury tax is around 35% when the labor supply is inelastic ($\phi = 3, 1.5$) and reaches levels of around 40% when the labor supply is more elastic ($\phi = 0.50$). When positional considerations are weaker ($\alpha = 0.05$) this tax is around 11%. The more elastic the labor supply, the higher the optimal luxury tax. Also, notice that when the labor supply is inelastic, the *aggregate* welfare losses, represented by the parameter λ_Θ, are no higher than 0.16%. When labor supply becomes more elastic, this loss increases to up to 0.24%. Regarding *non-aggregate* welfare measures, when the labor supply is inelastic ($\phi = 3, 1.5$), individuals at the top decile of the skills distribution experience gains between 0.61% and 6.67%. These gains, however, diminish when preferences are highly non-homothetic, the labor supply is elastic and positional concerns are weaker. Regarding welfare changes of individuals below the tenth decile of the skills distribution, the welfare losses due to linear taxation of positional goods are very high when preferences are homothetic, the labor supply is inelastic and the positional considerations are high. These losses decrease considerably when preferences exhibit non-homotheticity.

Figures 3.2 to 3.4 show comparative statics of the endogenous distributions of both consumption goods, effective output, utility levels and the optimal non-linear income tax and luxury tax that implement constrained efficient (CE) and double constrained efficient (DCE) allocations when skills are exponential distributed. Figures 3.5 to 3.7

show the same information when skills are distributed according to a Pareto distribution.[13] Figures 3.2 and 3.5 show the aforementioned distributions and taxes for different income elasticities, one of them corresponding to homothetic preferences. Observe that when preferences exhibit homotheticity ($\rho = 1$), then as the positional tax becomes flat, the labor marginal income tax schedule drops to zero. This is due to the fact that (homothetic) constant elasticity of substitution preferences together with a constant MRS between luxuries and necessities imply that the marginal utility of necessities is equalized across agents. This is proved formally in Appendix B.3. An equal marginal utility of necessities combined with a utilitarian planner delivers the zero tax schedule result. If the planner is not utilitarian (for instance, if she assigns higher weight to low skilled individuals), the result no longer holds. This is one of the cases analyzed in Appendix B.4.

Furthermore, under non-homothetic preferences, the flattening of the positional tax only reduces the marginal labor income tax regardless of whether or not the planner is utilitarian. Figures 3.3 and 3.6 show CE and DCE allocations and taxes in both environments assuming two different labor supply elasticities and homothetic preferences in both cases. Clearly, allocations are very sensitive to this parameter. Notice, however, that even when the labor supply elasticity is low, the effective output distribution under the DCE is not too different to the CE one. This is because changes in the after-tax

[13] Distributions were calculated using the Epanechnikov kernel with a bandwidth of $0.4 \times std(x) \times n^{-1/5}$ where x is the smoothed variable. The optimal bandwidth of Silverman (1986) over-smoothed the upper tail.

price of positional goods are offset by reductions in the marginal labor income tax to the extent that incentives are not violated. Finally, Figures 3.4 and 3.7 present the same set of results considering a very elastic labor supply ($\phi = 0.50$) and non-homothetic preferences.[14] These figures confirm the direction of adjustments in the marginal income tax upon imposing linearity in the positional tax. The labor income tax decreases to offset income effects due to changes in the after-tax price of positional goods. Nevertheless, since preferences are non-homothetic, changes in the labor tax are much more moderate than in the homothetic case. Appendix B.4 presents additional comparative statics exercises.

TABLE 3.1: Summary of Variables when Distribution is Exponential, $\alpha = 0.15$

Variable		$\rho = 1$			$\rho = 0.91$			$\rho = 0.80$	
	$\phi = 3$	$\phi = 1.5$	$\phi = 0.50$	$\phi = 3$	$\phi = 1.5$	$\phi = 0.50$	$\phi = 3$	$\phi = 1.5$	$\phi = 0.50$
τ (%)	34.58	35.94	41.35	35.77	36.87	41.08	37.24	37.91	40.03
λ_Θ (%)	0.06	0.06	0.13	0.05	0.08	0.17	0.06	0.06	0.10
λ_L (%)	0.94	0.89	1.49	0.71	0.69	0.92	0.52	0.45	0.43
λ_H (%)	-3.03	-2.23	-1.46	-2.62	-1.94	-1.17	-2.17	-1.58	-0.85

τ: Linear positional tax, λ_Θ: Aggregate welfare loss, λ_L: Welfare Loss Bottom λ_H: Welfare Loss Top. All numbers are reported in percentage terms. λ_f and λ_ψ imply that at the top of the distribution $\psi/f = 2.17$.

[14]For presentation purposes, I left out of the plot the bottom tail of the distributions in order to see clearly changes in the upper tail.

Table 3.2: Summary of Variables when Distribution is Exponential, $\alpha = 0.05$

Variable	$\rho = 1$			$\rho = 0.91$			$\rho = 0.80$		
	$\phi = 3$	$\phi = 1.5$	$\phi = 0.50$	$\phi = 3$	$\phi = 1.5$	$\phi = 0.50$	$\phi = 3$	$\phi = 1.5$	$\phi = 0.50$
τ (%)	11.07	11.17	12.93	11.16	11.52	12.88	11.68	11.94	12.75
λ_Θ (%)	0.01	0.01	0.01	0.01	0.01	0.02	0.02	0.01	0.04
λ_L (%)	0.42	0.35	0.52	0.29	0.25	0.30	0.20	0.16	0.14
λ_H (%)	-1.34	-0.89	-0.51	-1.06	-0.73	-0.45	-0.80	-0.61	-0.23

τ: Linear positional tax, λ_Θ: Aggregate welfare loss, λ_L: Welfare Loss Bottom λ_H: Welfare Loss Top. All numbers are reported in percentage terms. k_f and k_ψ imply that at the top of the distribution $\psi/f = 2.17$.

Table 3.3: Summary of Variables when Distribution is Pareto, $\alpha = 0.15$

Variable	$\rho = 1$			$\rho = 0.91$			$\rho = 0.80$		
	$\phi = 3$	$\phi = 1.5$	$\phi = 0.50$	$\phi = 3$	$\phi = 1.5$	$\phi = 0.50$	$\phi = 3$	$\phi = 1.5$	$\phi = 0.50$
τ (%)	32.93	35.6	43.87	33.51	35.29	41.96	34.21	35.52	39.32
λ_Θ (%)	0.14	0.16	0.21	0.15	0.15	0.24	0.12	0.12	0.14
λ_L (%)	1.79	1.79	3.15	1.33	1.23	1.85	0.88	0.86	0.67
λ_H (%)	-6.67	-4.94	-3.17	-5.71	-4.13	-2.78	-4.61	-3.73	-1.60

τ: Linear positional tax, λ_Θ: Aggregate welfare loss, λ_L: Welfare Loss Bottom λ_H: Welfare Loss Top. All numbers are reported in percentage terms. k_f and k_ψ imply that at the top of the distribution $\psi/f = 2.17$.

Table 3.4: Summary of Variables when Distribution is Pareto, $\alpha = 0.05$

Variable	$\rho = 1$			$\rho = 0.91$			$\rho = 0.80$		
	$\phi = 3$	$\phi = 1.5$	$\phi = 0.50$	$\phi = 3$	$\phi = 1.5$	$\phi = 0.50$	$\phi = 3$	$\phi = 1.5$	$\phi = 0.50$
τ (%)	10.35	11.03	13.92	10.55	11.14	13.26	10.82	11.28	12.64
λ_Θ (%)	0.02	0.02	0.03	0.02	0.03	0.04	0.01	0.04	0.01
λ_L (%)	0.71	0.72	1.51	0.52	0.52	0.69	0.30	0.27	0.26
λ_H (%)	-2.68	-2.02	-1.43	-2.35	-1.78	-1.07	-1.75	-1.07	-0.72

τ: Linear positional tax, λ_Θ: Aggregate welfare loss, λ_L: Welfare Loss Bottom λ_H: Welfare Loss Top. All numbers are reported in percentage terms. k_f and k_ψ imply that at the top of the distribution $\psi/f = 2.17$.

3.5 Conclusions

In this paper I have introduced positional consumption goods within the Mirrlees (1971) framework. Positional goods are those whose valuation depends on an endogenous consumption benchmark. This consumption benchmark is a weighted average of all agents' consumption of the positional good. As the contribution of individuals to the endogenous consumption benchmark differs from their population size, constrained efficient allocations exhibit a non-linear wedge between positional and non-positional goods. Constrained efficient allocations can be implemented through a non-linear positional tax together with a non-linear income tax with standard properties, namely, without distortions at the extremes. The previous implementation, however, is subject to arbitrage opportunities across consumption goods. Thus, an extra constraint is imposed: the marginal rate of substitution between the positional and the positional good must be the same across agents. I have shown that the resulting *double* constrained efficient allocations can be implemented through a *linear* positional tax together with a non-linear labor income tax.

While aggregate welfare losses in the double constrained environment with respect to the constrained efficient environment are very low and in some cases negligible, large distributional effects arise. Assuming that the positional externality is increasing in income, individuals at the high end of the income distribution experience large gains, since for them a flat tax effectively reduces the after-tax price of positional goods and higher consumption thus occurs. This is true since the drop in the price generates a positive income effect that cannot be offset by an increase in the marginal income tax, as optimality requires no distortion at the top. The opposite is true for individuals at the bottom of the skills distribution; they experience considerable welfare losses since a flat luxury tax increases the after-tax price of luxuries. This generates a negative income effect that is offset by a reduction in the marginal income tax. Such a reduction cannot fully offset the income effect since this would violate incentives. When preferences are non-homothetic, small adjustments in the income tax are more effective in offsetting income effects derived from changes in the "luxury tax". My results suggest that the effectiveness of a linear consumption tax in correcting positional externalities would crucially depend on the degree of non-homotheticity in preferences over positional and non-positional goods. My results also show that under reasonable parameters describing the magnitude of positional considerations a society may have, the flat tax imposed on the consumption of the positional good is by no means negligible.

An important extension of the current model is to incorporate a production technology whose marginal rate of transformation between luxuries and necessities is not constant. Such a specification would allow us to incorporate in our quantitative analysis potential sharper changes in the output of the economy as a result of a good specific tax. It is also important to note that the results presented here assume that agents cannot buy positional goods in markets with different tax regimes. This consideration would impose an upper bound on this tax that is not considered in this model.

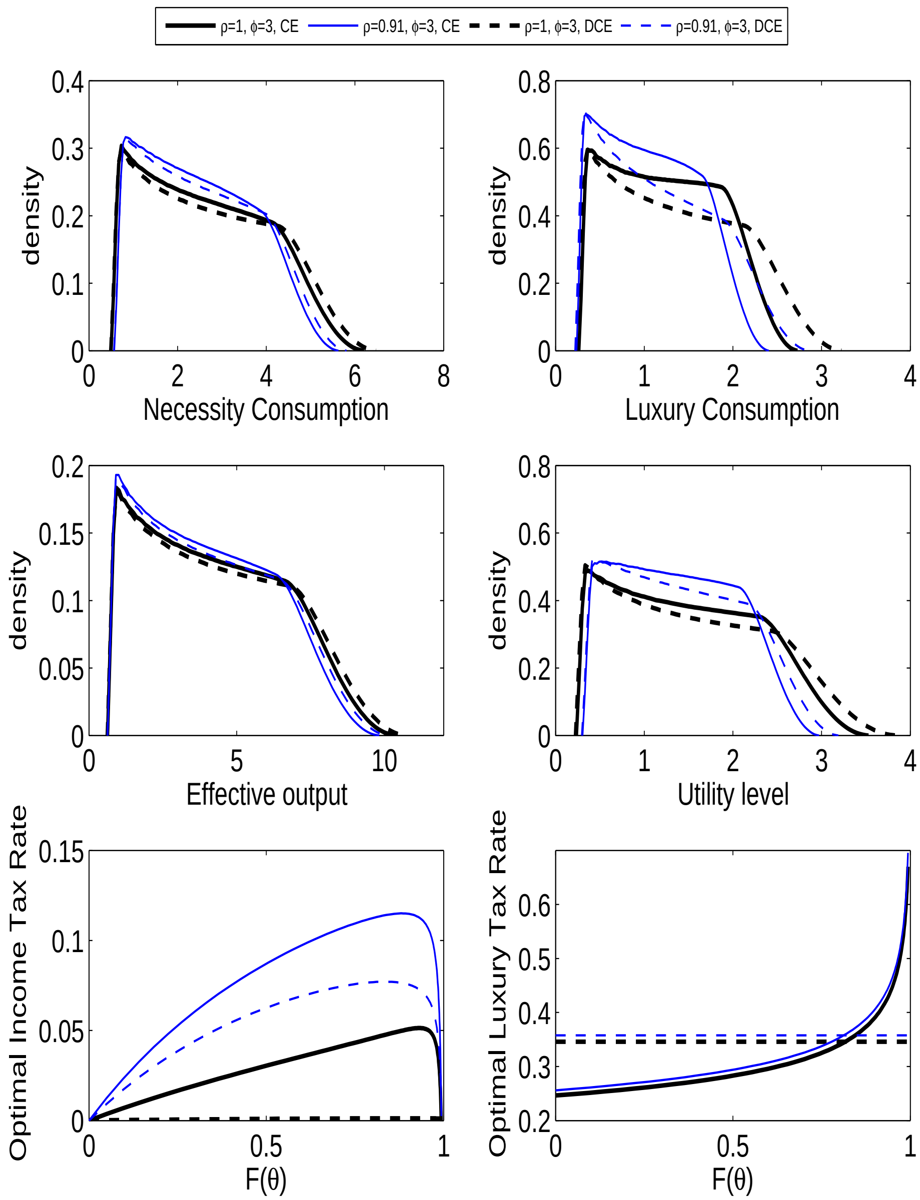

FIGURE 3.2: Endogenous distributions and optimal taxes when $\alpha = 0.15$ and distribution of skills is exponential. Changes in the income elasticity.

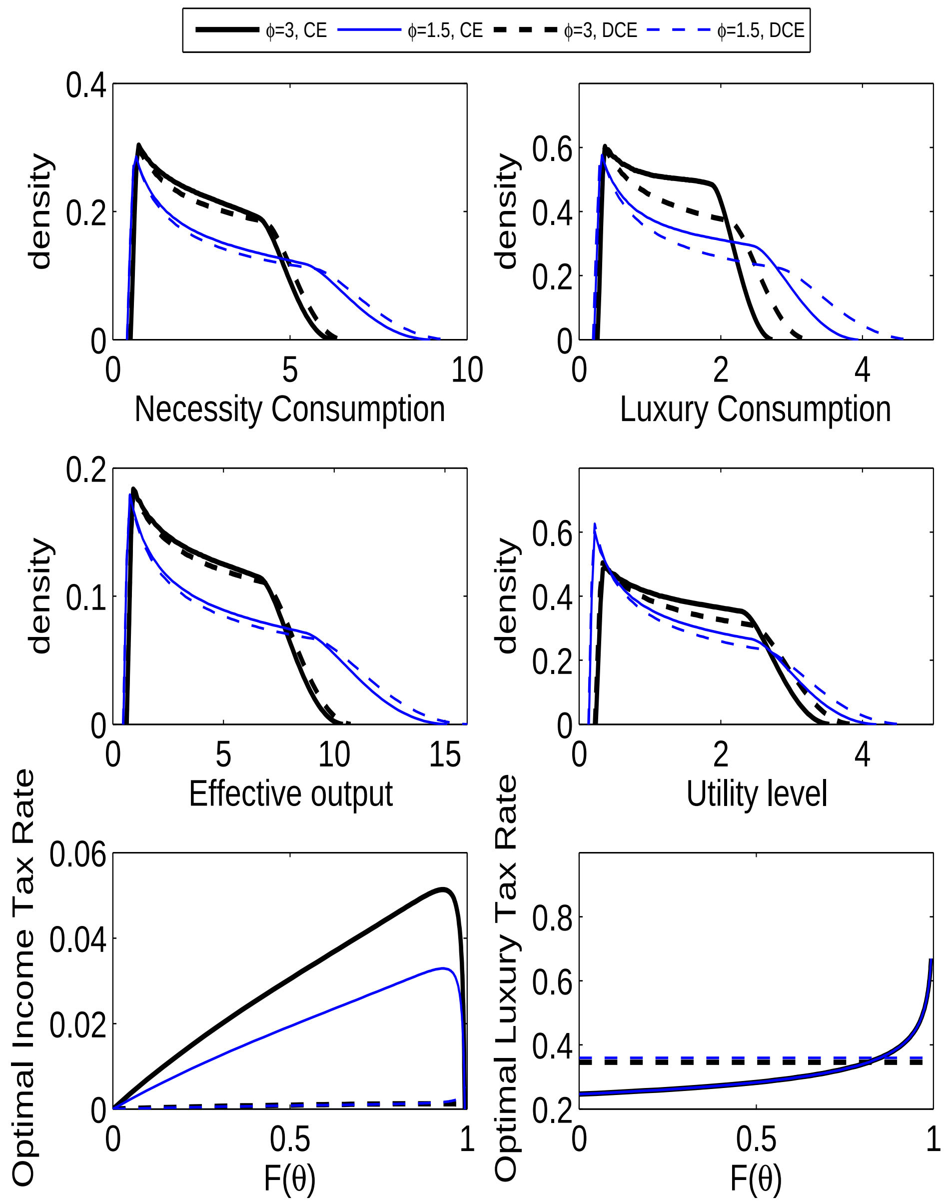

FIGURE 3.3: Endogenous distributions and optimal taxes when $\alpha = 0.15$, preferences are homothetic $\rho = 1$ and distribution of skills is exponential. Changes in the labor supply elasticity.

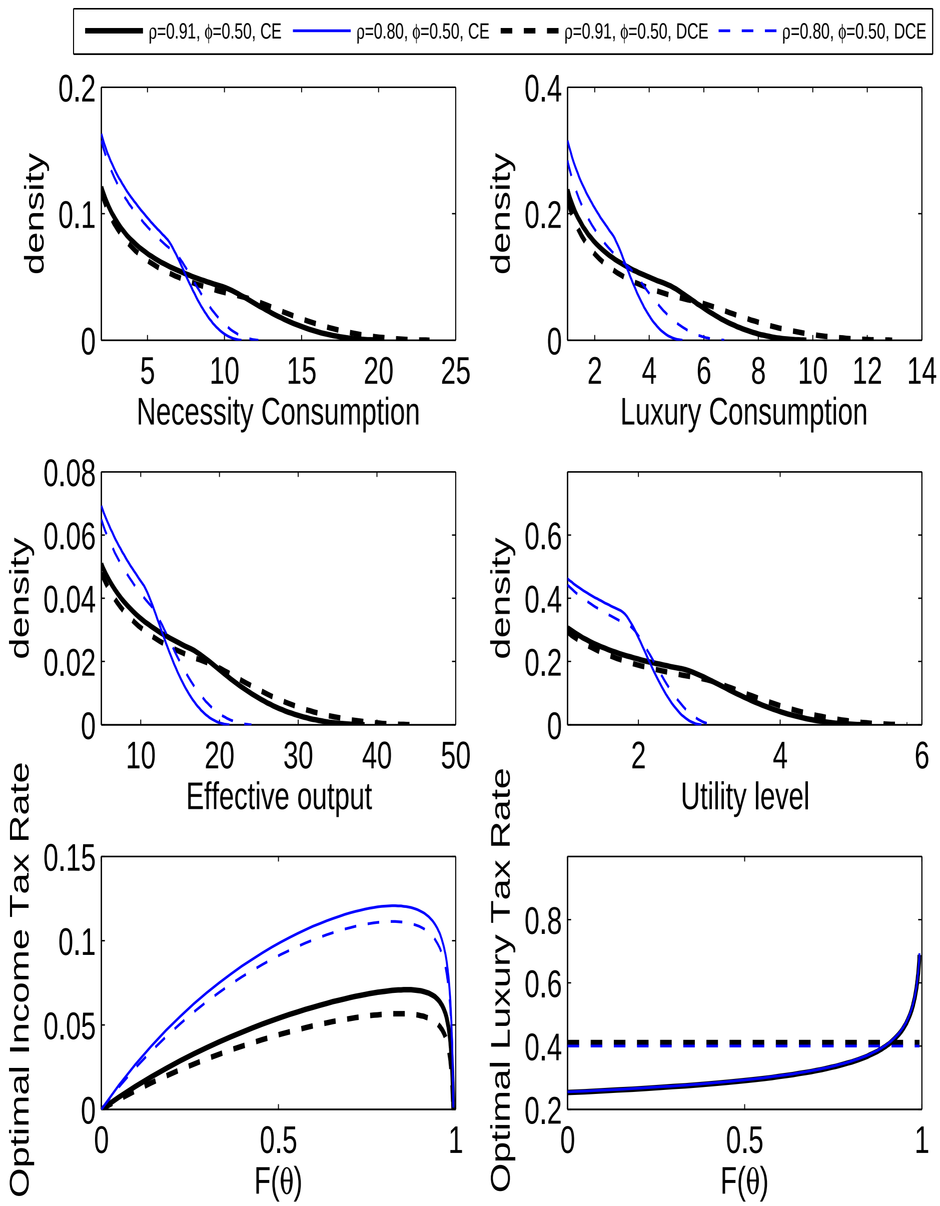

FIGURE 3.4: Upper tail of endogenous distributions and optimal taxes when $\alpha = 0.15$ and distribution of skills is exponential. Changes in the income elasticity.

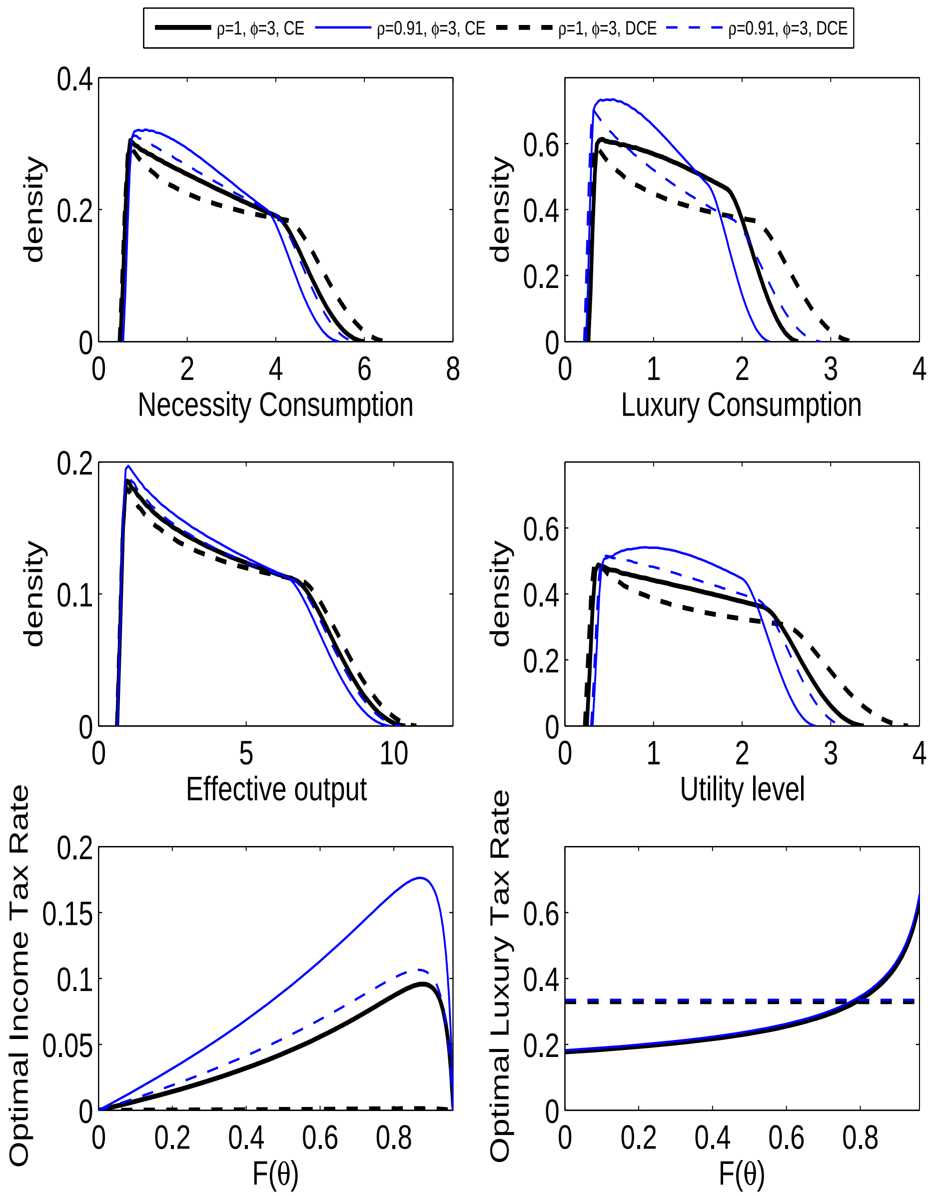

Figure 3.5: Endogenous distributions and optimal taxes when $\alpha = 0.15$ and distribution of skills is Pareto. Changes in the income elasticity.

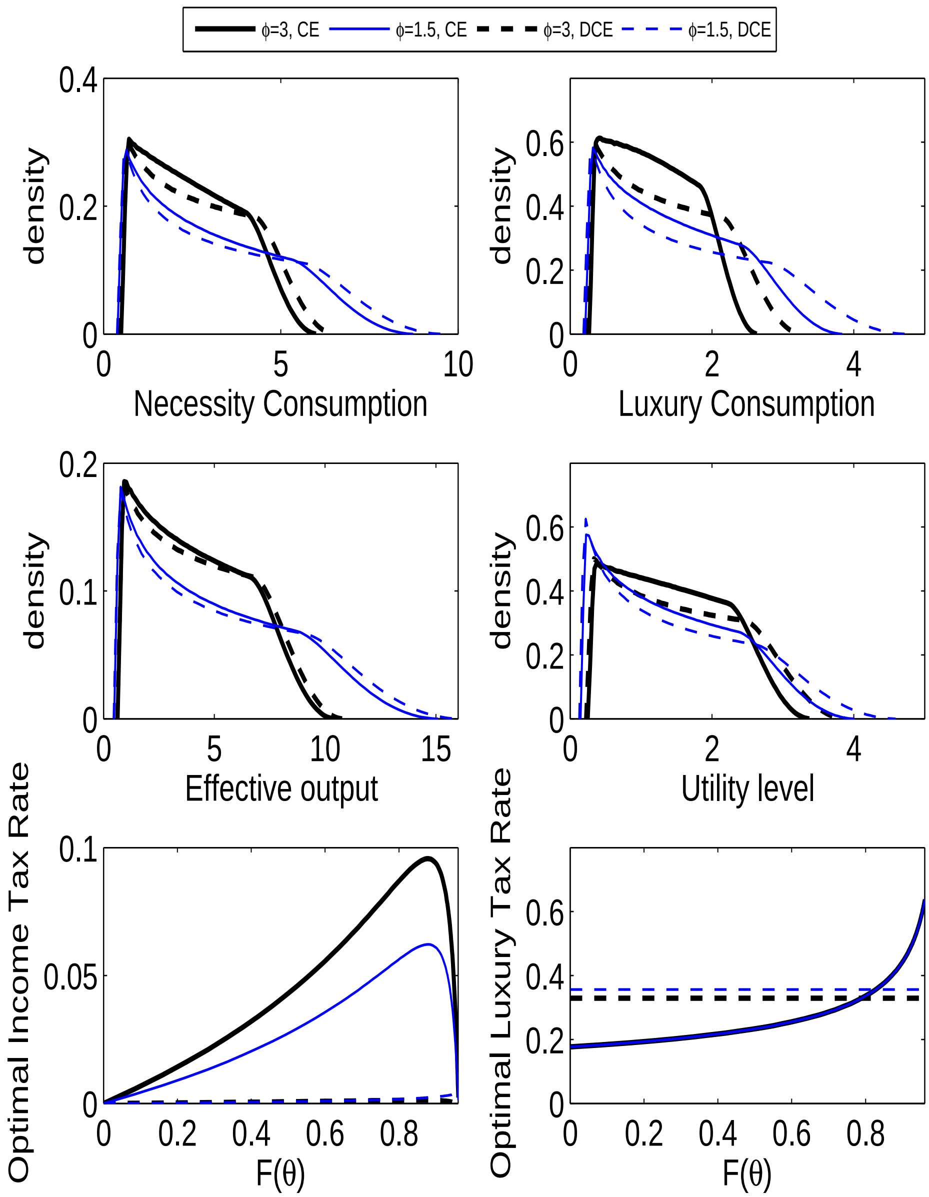

FIGURE 3.6: Endogenous distributions and optimal taxes when $\alpha = 0.15$, preferences are homothetic $\rho = 1$ and distribution of skills is Pareto. Changes in the labor supply elasticity.

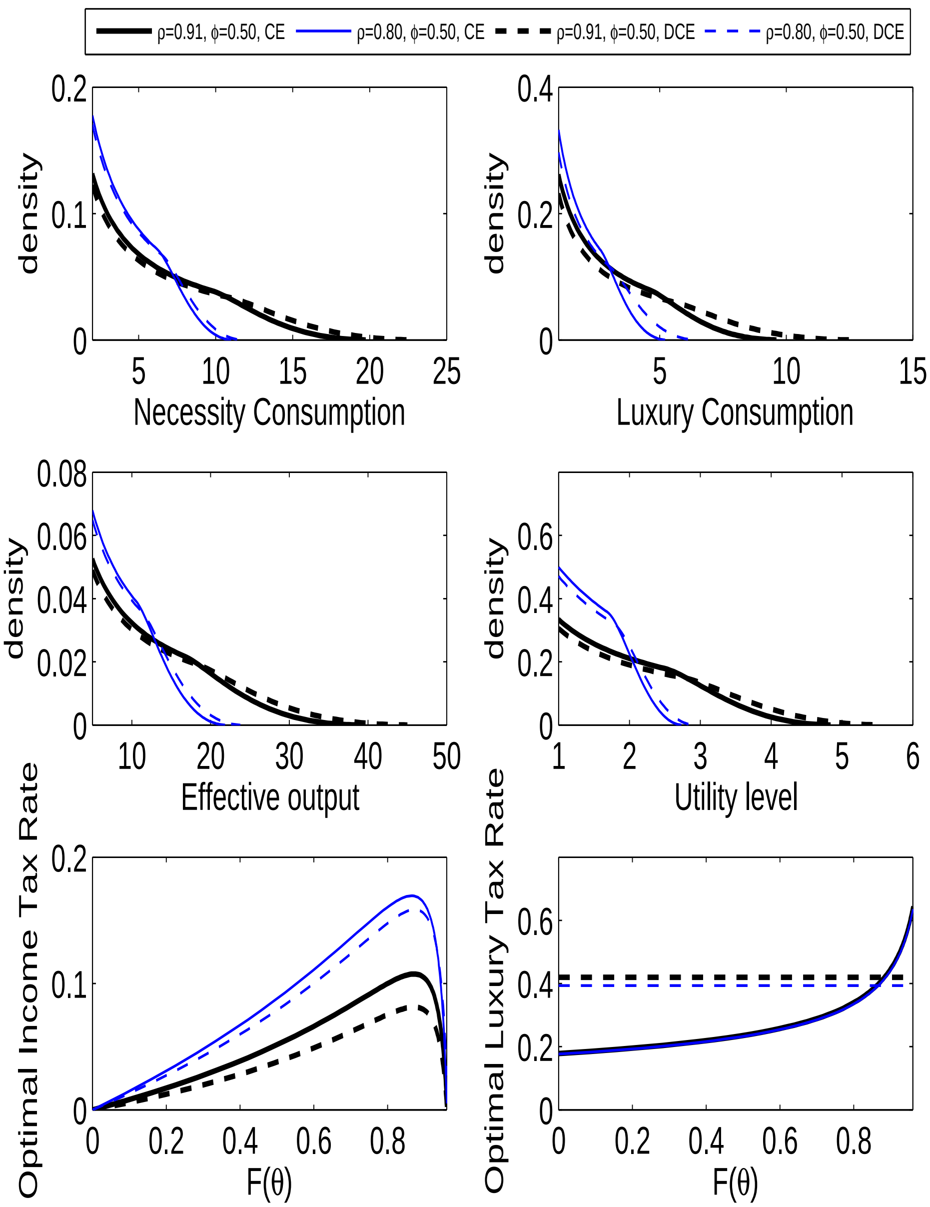

FIGURE 3.7: Upper tail of endogenous distributions and optimal taxes when $\alpha = 0.15$ and distribution of skills is Pareto. Changes in the income elasticity.

Chapter 4

Conclusion and Discussion

Optimal tax theory has difficulty rationalizing high marginal tax rates at the upper end of the labor income distribution. In Chapter 2 of this dissertation, I presented a model that rationalizes high labor income taxes on affluent individuals: taxation on the rich occurs in order to correct consumption externalities. This happens in the absence of a non-linear consumption tax schedule. Surprisingly, the estimated jealousy parameters for the U.S. and the U.K. are moderate, yet produce quantitatively high effects over labor income tax rates.

Rationalizing observed marginal labor income taxes as Pigouvian requires that affluent individuals generate higher consumption externalities than poorer ones. Despite productive efficiency arguments, this may be one of the reasons policy-makers in practice face impediments to reducing labor income taxes at high income levesl.

In Chapter 3, I presented a normative analysis of taxation in an economy where only some of the goods are positional, that is, whose consumption generates consumption externalities. I showed that optimal allocations in this environment may be implemented by a *linear* tax on the positional good in combination with a non-linear marginal income tax schedule with standard Mirrleesian properties, namely without taxation at extremes. Thus, the two instruments are used for redistributive and corrective purposes. The former tax corrects over-consumption of the positional good by making agents internalize the externality they impose to the society whereas the income tax tailors redistribution subject to incentive problems. However, the latter is not independent of the former as adjustments in the marginal labor income tax are required to offset income effects due to changes in the after-tax price of positional goods.

Under reasonable parameters describing the magnitude of positional considerations a society may have, my numerical calculations show that the flat tax imposed on the consumption of the positional good is by no means negligible. Moreover, my calculations suggest that while a flat tax on the positional good performs almost as well as a non-linear one in terms of aggregate welfare, large distributional effects are induced by the former.

Bibliography

ABEL, A. (2005): "Optimal Taxation when Consumers Have Endogenous Benchmark Levels of Consumption," *Review of Economic Studies*, 72(1), 21–42.

AIT-SAHALIA, Y., J. PARKER, AND M. YOGO (2004): "Luxury goods and the equity premium," *Journal of Finance*, pp. 2959–3004.

ALPIZAR, F., F. CARLSSON, AND O. JOHANSSON-STENMAN (2005): "How much do we care about absolute versus relative income and consumption?," *Journal of Economic Behavior and Organization*, 56(3), 405–421.

ATKINSON, A. (1990): "Public economics and the economic public," *European Economic Review*, 34(2-3), 225–248.

ATKINSON, A., AND J. STIGLITZ (1972): "The Structure of Indirect Taxation and Economic Efficiency," *Journal of Public Economics*, 1(1), 97–119.

——— (1976): "The Design of Tax Structure: Direct Versus Indirect Taxation," *Journal of Public Economics*, 6(1-2).

Carlsson, F., O. Johansson-Stenman, and P. Martinsson (2007): "Do You Enjoy Having More than Others? Survey Evidence of Positional Goods," *Economica*, 74(4).

Clark, A., and A. Oswald (1996): "Satisfaction and comparison income," *Journal of Public Economics*, 61(3), 359–381.

Coddington, E. (1989): *An Introduction to Ordinary Differential Equations*. Courier Dover Publications.

Constantinides, G. (1990): "Habit Formation: A Resolution of the Equity Premium Puzzle," *Journal of Political Economy*, 98(3), 519.

Costa, D. (2001): "Estimating real income in the United States from 1888 to 1994: Correcting CPI bias using Engel curves," *Journal of Political Economy*, 109(6), 1288–1310.

da Costa, C. E. (2009): "Yet Another Reason to Tax Goods," *Review of Economic Dynamics*, pp. 363–376.

Diamond, P. (1998): "Optimal Income Taxation: An Example with a U-Shaped Pattern of Optimal Marginal Tax Rates," *American Economic Review*, 88(1), 83–95.

Dupor, B., and W. Liu (2003): "Jealousy and Equilibrium Overconsumption," *American Economic Review*, 93(1), 423–428.

Dynan, K., and E. Ravina (2007): "Increasing income inequality, external habits, and self-reported happiness," *American Economic Review*, 97(2), 226–231.

Evers, M., R. Mooij, and D. Vuuren (2005): "What Explains the Variation in Estimates of Labour Supply Elasticities?," *CESifo Working Paper 1633*.

Frank, R. (2005): "Positional Externalities Cause Large and Preventable Welfare Losses," *American Economic Review*, 95(2), 137–141.

——— (2008): "Should public policy respond to positional externalities?," *Journal of Public Economics*, 92(8), 1777–1786.

Golosov, M., and A. Tsyvinski (2007): "Optimal Taxation With Endogenous Insurance Markets," *The Quarterly Journal of Economics*, 122(2), 487–534.

Gouveia, M., and R. Strauss (1994): "Effective Federal Individual Income Tax Functions: An Exploratory Empirical Analysis," *National Tax Journal*, 47(2), 317–339.

Heaton, J. (1995): "An empirical investigation of asset pricing with temporally dependent preference specifications," *Econometrica*, 63(3), 681–717.

HM-Revenue-&-Customs (1998-2007): *Survey of Personal Incomes, Several Years: Public Use Tape [computer file]*. Colchester, Essex: UK Data Archive [distributor].

Internal-Revenue-Service (1995-2004): *Statistics of Income: Individual Tax Returns*. Washington, D.C: U.S. Government Printing Office.

J. Solnick, S., and D. Hemenway (1998): “Is more always better?: A survey on positional concerns,” *Journal of Economic Behavior and Organization*, 37(3), 373–383.

Ljungqvist, L., and H. Uhlig (2000): “Tax Policy and Aggregate Demand Management Under Catching Up with the Joneses,” *American Economic Review*, 90(3), 356–366.

Luttmer, E. (2005): “Neighbors as Negatives: Relative Earnings and Well-Being,” *The Quarterly Journal of Economics*, 120(3), 963–1002.

Mirrlees, J. (1971): “An Exploration in the Theory of Optimum Income Taxation,” *Review of Economic Studies*, 38(114), 175–208.

Neumark, D., and A. Postlewaite (1998): “Relative income concerns and the rise in married women’s employment,” *Journal of Public Economics*, 70(1), 157–183.

Oswald, A. (1983): “Altruism, jealousy and the theory of optimal non-linear taxation,” *Journal of Public Economics*, 20(1), 77–87.

Pakos, M. (2006): “Measuring Intra-Temporal and Inter-Temporal Substitutions: The Role of Consumer Durables,” *Carnegie-Mellon University, Working paper.*

Pencavel, J. (1986): “Labor Supply of Men: A Survey,” *Handbook of Labor Economics*, 1, 3–102.

Rustichini, A., and A. Vostroknutov (2007): "Competition with Skill and Luck," Mimeo, University of Minnesota.

Sadka, E. (1976): "On income distribution, incentive effects and optimal income taxation," *Review of Economic Studies*, 43(2), 261–8.

Saez, E. (2001): "Using Elasticities to Derive Optimal Income Tax Rates," *Review of Economic Studies*, 68(1), 205–229.

Salanié, B. (1997): *The Economics of Contracts: A Primer*. MIT Press.

Samano, D. (2009): "Explaining Taxes on the Rich: The Role of Jealousy," Mimeo, University of Minnesota.

Seade, J. (1977): "On the Shape of Optimal Tax Schedules," *Journal of Public Economics*, 7(2), 203–235.

Silverman, B. (1986): *Density Estimation for Statistics and Data Analysis*. Chapman & Hall/CRC.

Solnick, S., and D. Hemenway (2005): "Are positional concerns stronger in some domains than in others?," *American Economic Review*, 95(2), 147–151.

Stokey, N. (2008): *Economics of Inaction: Stochastic Control Models with Fixed Costs*. Princeton University Press.

Tuomala, M. (1990): *Optimal Income Tax and Redistribution*. Oxford University Press, USA.

Veblen, T. (1899): *The Theory of the Leisure Class, 1934 edition*.

Wand, M., J. Marron, and D. Ruppert (1991): "Transformations in density estimation," *Journal of the American Statistical Association*, 86(414), 343–353.

Werning, I. (2007): "Pareto Efficient Income Taxation," Mimeo, MIT.

Appendix A

Appendix Chapter 2

A.1 Proofs

Proof of Proposition 1

Proof. The first step is to transform the continuum of incentive compatibility constraints (2.4) into a first order condition. Let

$$W(\theta, \theta') \equiv u(c(\theta'), C) - v\left(\frac{y(\theta')}{\theta}\right) \tag{A.1}$$

A necessary condition for truthful revelation of type is $\frac{\partial W(\theta,\theta')}{\partial \theta'}|_{\theta'=\theta} = 0$, therefore it follows that

$$u_c(c(\theta), C)c'(\theta) = v'\left(\frac{y(\theta)}{\theta}\right)\frac{y'(\theta)}{\theta} \quad \forall \theta \in \Theta \tag{A.2}$$

Moreover, under truthful revelation $W(\theta) = u(c(\theta), C) - v\left(\frac{y(\theta)}{\theta}\right)$ and hence, $W'(\theta) = u_c(c(\theta), C)c'(\theta) - v'\left(\frac{y(\theta)}{\theta}\right)\frac{y'(\theta)}{\theta} + v'\left(\frac{y(\theta)}{\theta}\right)\frac{y(\theta)}{\theta^2}$, which together with (A.2) becomes

$$W'(\theta) = v'\left(\frac{y(\theta)}{\theta}\right)\frac{y(\theta)}{\theta^2} \quad \forall\theta \in \Theta. \tag{A.3}$$

Following Werning (2007), I define the expenditure function $e(W(\theta), y(\theta), C; \theta)$ to satisfy $W(\theta) = u(e, C) - v\left(\frac{y(\theta)}{\theta}\right)$. The planner problem can be restated as

$$\max_{W(\cdot), y(\cdot), C} \int_{\Theta} W(\theta)g(\theta)d\theta \tag{A.4}$$

s.t

$$\int_{\Theta} e(W(\theta), y(\theta), C; \theta)f(\theta)d\theta = \int_{\Theta} y(\theta)f(\theta)d\theta \tag{A.5}$$

$$W'(\theta) = v'\left(\frac{y(\theta)}{\theta}\right)\frac{y(\theta)}{\theta^2} \quad \forall\theta \in \Theta \tag{A.6}$$

$$C = \int_{\Theta} e(W(\theta), y(\theta), C; \theta)\psi(\theta)d\theta \tag{A.7}$$

The corresponding Lagrangian is

$$\mathfrak{L}(W(\theta), y(\theta), C, \lambda, \mu(\theta), \gamma) = \int_{\Theta} W(\theta)g(\theta)d\theta - \lambda \int_{\Theta} \left[\left(e(W(\theta), y(\theta), C; \theta) - y(\theta)\right) f(\theta)\right] d\theta$$

$$+\int_{\Theta}\mu(\theta)\left[W'(\theta)-v'\left(\frac{y(\theta)}{\theta}\right)\frac{y(\theta)}{\theta^2}\right]d\theta+\gamma\left[C-\int_{\Theta}e(W(\theta),y(\theta),C;\theta)\psi(\theta)d\theta\right] \quad (A.8)$$

Using integration by parts, it follows that

$$\int_{\Theta}\mu(\theta)W'(\theta)d\theta=\mu(\bar{\theta})W(\bar{\theta})-\mu(\underline{\theta})W(\underline{\theta})-\int_{\Theta}\mu'(\theta)W(\theta)d\theta \quad (A.9)$$

thus, we can reexpress the above Lagrangian as

$$\mathfrak{L}(W(\theta),y(\theta),C,\lambda,\mu(\theta),\gamma)=\int_{\Theta}W(\theta)g(\theta)d\theta-\lambda\int_{\Theta}[(e(W(\theta),y(\theta),C;\theta)-y(\theta))f(\theta)]d\theta$$

$$+\mu(\bar{\theta})W(\bar{\theta})-\mu(\underline{\theta})W(\underline{\theta})-\int_{\Theta}\mu'(\theta)W(\theta)d\theta-\int_{\Theta}\mu(\theta)v'\left(\frac{y(\theta)}{\theta}\right)\frac{y(\theta)}{\theta^2}d\theta$$

$$+\gamma\left[C-\int_{\Theta}e(W(\theta),y(\theta),C;\theta)\psi(\theta)d\theta\right] \quad (A.10)$$

Assuming interior solution, it follows from first order conditions that

$W(\theta)$:

$$g(\theta)-\lambda f(\theta)e_W(W(\theta),y(\theta),C;\theta)-\mu'(\theta)-\gamma\psi(\theta)e_W(W(\theta),y(\theta),C;\theta)=0 \quad (A.11)$$

$y(\theta)$:

$$-\lambda e_y(W(\theta), y(\theta), C; \theta) f(\theta) + \lambda f(\theta) - \frac{\mu(\theta)}{\theta^2} v'\left(\frac{y(\theta)}{\theta}\right)\left[1 + \frac{1}{\epsilon(\theta)}\right]$$

$$- \gamma e_y(W(\theta), y(\theta), C; \theta)\psi(\theta) = 0 \tag{A.12}$$

C :

$$-\lambda \int_{\Theta} e_C(W(\theta), y(\theta), C; \theta) f(\theta) d\theta + \gamma - \gamma \int_{\Theta} e_C(W(\theta), y(\theta), C; \theta)\psi(\theta) d\theta = 0 \tag{A.13}$$

together with the boundary condition $\mu(\bar{\theta}) = 0$ and where $\epsilon(\theta) \equiv \frac{v'(\frac{y(\theta)}{\theta})}{v''(\frac{y(\theta)}{\theta})\frac{y(\theta)}{\theta}}$. Moreover, implicitly differentiating $W(\theta) = u(e, C) - v\left(\frac{y(\theta)}{\theta}\right)$ we have that $e_W(W(\theta), y(\theta), C; \theta) = \frac{1}{u_c(c(\theta), C)}$, $e_y(W(\theta), y(\theta), C; \theta) = \frac{v'\left(\frac{y(\theta)}{\theta}\right)\frac{1}{\theta}}{u_c(c(\theta), C)}$ and $e_C(W(\theta), y(\theta), C; \theta) = -\frac{u_C(c(\theta), C)}{u_c(c(\theta), C)}$. The result follows after manipulating (A.11)-(A.13). $\square$

Proof of Proposition 2

Proof. Taking first order conditions in agent's problem we have

$$\frac{T'(y(\theta))}{1 - T'(y(\theta))} = \frac{u_c(c^{eq}(\theta), C^{eq})}{v'\left(\frac{y^{eq}(\theta)}{\theta}\right)\frac{1}{\theta}} - 1 \tag{A.14}$$

where $C^{eq} = \int_{\Theta} c^{eq}(\theta)\psi(\theta) d\theta$. Substituting (2.11) into (A.14) it follows that

$$\frac{u_c(c^{eq}(\theta)), C^{eq})}{v'(\frac{y^{eq}(\theta)}{\theta})\frac{1}{\theta}} - 1 = \frac{\gamma\psi(\theta)}{\lambda f(\theta)} +$$

$$\frac{u_c(c^*(\theta), C^*)}{\theta f(\theta)}\left[1 + \frac{1}{\epsilon^*(\theta)}\right]\int_{\underline{\theta}}^{\theta}\left[\frac{g(t)}{\lambda} - \frac{f(t)}{u_c(c^*(t), C^*)} - \frac{\gamma}{\lambda}\frac{\psi(t)}{u_c(c^*(t), C^*)}\right]dt \tag{A.15}$$

Since the government balances its budget, we must have that

$$\int_{\Theta} c^{eq}(\theta)f(\theta)d\theta = \int_{\Theta} y^{eq}(\theta)f(\theta)d\theta \tag{A.16}$$

thus from (A.15)-(A.16) we conclude that $\{c^{eq}(\theta), y^{eq}(\theta)\}_{\theta\in\Theta} = \{c^*(\theta), y^*(\theta)\}_{\theta\in\Theta}$. □

Proof of Proposition 3

Proof. By quasi-linearity of preferences $u_c(c(\theta), C) = 1$. Thus optimal income tax satisfies

$$\frac{T'(y(\theta))}{1 - T'(y(\theta))} = \frac{\alpha}{(1-\alpha)}\frac{\psi(\theta)}{f(\theta)} + \frac{1}{\theta f(\theta)}\left[1 + \frac{1}{\epsilon(\theta)}\right]\frac{\mu(\theta)}{\lambda} \tag{A.17}$$

with

$$\mu(\theta) = \int_{\underline{\theta}}^{\theta}[g(t) - \lambda f(t) - \gamma\psi(t)]dt. \tag{A.18}$$

By boundary conditions we have $\mu(\bar{\theta}) = 0$, hence

$$\int_{\Theta}[g(t) - \lambda f(t) - \gamma\psi(t)]dt = 0 \quad \Rightarrow \quad \lambda = 1 - \gamma \tag{A.19}$$

which together with the fact that $\frac{\gamma}{\lambda} = \frac{\alpha}{1-\alpha}$ implies that $\lambda = 1 - \alpha$ and $\gamma = \alpha$. Plugging the previous values into (A.17) and substituting into (A.18) delivers the result after algebraic manipulations. □

A.2 Asymptotic Tax in a Quasi-Linear Environment

Under the assumption that $\bar{\theta} < \infty$ it can be seen from Corollary 1 that $\frac{T'(y^*(\bar{\theta}))}{1-T'(y^*(\bar{\theta}))} = \frac{\alpha}{(1-\alpha)}\frac{\psi(\bar{\theta})}{f(\bar{\theta})}$. Hence, under a bounded distribution of skills I obtain a non-zero taxation at the top due to corrective considerations. Proposition 9 exhibits the formula for the optimal marginal labor income tax as $\bar{\theta}$ goes to infinity. I consider the case of quasi-linear preferences, a constant elasticity of labor supply and $f(\theta)$ Pareto-distributed. The last assumption is used based on Diamond (1998) and Saez (2001). Both studies obtain positive asymptotic marginal tax rates, however those results depend critically on $f(\theta)$ being Pareto.[1]

Proposition 9. *Suppose $f(\theta)$ is Pareto distributed with parameter $k > 0$ and that $L_1 \equiv \lim_{\theta\to\infty} \frac{\psi(\theta)}{f(\theta)}$ and $L_2 \equiv \lim_{\theta\to\infty} \frac{g(\theta)}{f(\theta)}$ exist. If $u(c(\theta), C) = c(\theta) - \alpha C$, $\alpha \in [0, 1)$, $v\left(\frac{y}{\theta}\right) = \frac{1}{1+\phi}\left(\frac{y}{\theta}\right)^{1+\phi}, \phi > 0$ then $\frac{T'_\infty}{1-T'_\infty} = \frac{\alpha}{(1-\alpha)}L_1\left[\frac{1+\phi+k}{k}\right] + \frac{1+\phi}{k}\left[1 - \frac{1}{(1-\alpha)}L_2\right]$ where $T'_\infty \equiv \lim_{\theta\to\infty} T'(y(\theta))$.*

Proof. From Proposition 3 we have that

$$\frac{T'(y(\theta))}{1-T'(y(\theta))} = \frac{\alpha}{(1-\alpha)}\frac{\psi(\theta)}{f(\theta)} + \frac{1-F(\theta)}{(1-\alpha)\theta f(\theta)}\left[1+\frac{1}{\epsilon(\theta)}\right]\frac{[G(\theta) - (1-\alpha)F(\theta) - \alpha\Psi(\theta)]}{1-F(\theta)} \tag{A.20}$$

[1]For a Pareto distribution $f(\theta) = \frac{k\underline{\theta}^k}{\theta^{k+1}}, \theta \in [\underline{\theta}, \infty), \underline{\theta} > 0, k > 0$ and $F(\theta) = 1 - \left(\frac{\underline{\theta}}{\theta}\right)^k$.

Since $f(\theta)$ is Pareto it follows that $\frac{1-F(\theta)}{\theta f(\theta)} = \frac{1}{k}$ and since $v(\frac{y}{\theta}) = \frac{1}{1+\phi}\left(\frac{y}{\theta}\right)^{1+\phi}$ we have that $\left[1 + \frac{1}{\epsilon(\theta)}\right] = 1 + \phi$. Using the fact that $L_1 \equiv \lim_{\theta\to\infty} \frac{\psi(\theta)}{f(\theta)} < \infty$ it follows that

$$\lim_{\theta\to\infty} \frac{T'(y(\theta))}{1 - T'(y(\theta))} = \frac{\alpha}{1-\alpha} L_1 + \frac{(1+\phi)}{k(1-\alpha)} \lim_{\theta\to\infty} \frac{[G(\theta) - (1-\alpha)F(\theta) - \alpha\Psi(\theta)]}{1 - F(\theta)} \tag{A.21}$$

Using L'Hôpital's rule, $\lim_{\theta\to\infty} \frac{[G(\theta)-(1-\alpha)F(\theta)-\alpha\Psi(\theta)]}{1-F(\theta)} = -L_2 + (1-\alpha) + \alpha L_1$ and substituting into (A.21) delivers the result. □

Corollary 5 (Proposition 3)**.** *If $L_1 \geq 1$ and $L_2 \leq 1$, then $T'_\infty \geq \alpha$.*

Proof. Trivial. □

Figure A.1 shows the ratio $\frac{1-F_Y(y)}{yf_Y(y)}$ of the income distribution in the U.S. for 1992, 1993 and from 1995-2004. I include the years 1992 and 1993 since Saez (2001) estimates a Pareto distribution parameter for labor earnings based on those years. This ratio was constructed after smoothing the upper tail of $f_Y(y)$ using a gaussian kernel. The width was set at $h = 1.36 \times s.d. \times n^{-1/5}$, where the standard deviation ($s.d.$) and n were calculated for observations exceeding 13.5 million dollars (expressed in 2004 dollars). Table A.1 reports the number of observations, n, in the sample exceeding that threshold.[2]

It can be seen that at least for some years the ratio $\frac{1-F_Y(y)}{yf_Y(y)}$ at the very top of the

[2]This amount is approximately equivalent to 10 million expressed in 1992 dollars. Saez (2001) reports that starting at this income level the number of taxpayers in the database is very small.

Table A.1: In Sample Number of Gross Income Observations Exceeding $13,500,000 (2004 dollars)

Year	Observations	Year	Observations
1992	54	1999	337
1993	27	2000	498
1995	33	2001	330
1996	66	2002	228
1997	100	2003	252
1998	207	2004	391

distribution is decreasing. This fact indicates that the very top of the income distribution of the United States may not be accurately represented by a Pareto distribution. Moreover, a decreasing $\frac{1-F_Y(y)}{yf_Y(y)}$ would imply a decreasing pattern of optimal taxes in the canonical Mirrleesian model. This is indeed what Mirrlees (1971) finds since he assumes a log-normal distribution of skills for which the previous ratio is decreasing. The model with consumption externalities would deliver asymptotic non zero optimal taxes even if the hazard ratio is decreasing.

A.3 Proof of Proposition 4

Proof. Using integration by parts, we can reexpress (2.15) as

$$\alpha = \frac{\int_\Theta \frac{f(\theta)T'(y(\theta))}{\theta^{1+\phi}(1-T'(y(\theta)))}d\theta + (1+\phi)\int_\Theta \frac{(F(\theta)-G(\theta))}{\theta^{2+\phi}}d\theta}{\int_\Theta \frac{f(\theta)T'(y(\theta))}{\theta^{1+\phi}(1-T'(y(\theta)))}d\theta + (1+\phi)\int_\Theta \frac{F(\theta)}{\theta^{2+\phi}}d\theta + \frac{1}{\bar{\theta}^{1+\phi}}} \tag{A.22}$$

Thus,

$$\bar{\alpha} = \max_{G(\cdot)}\{\alpha \mid G(\theta) \geq F(\theta)\ \forall\theta \in \Theta, G'(\theta) \geq 0, G(\underline{\theta}) = 0, G(\bar{\theta}) = 1\}$$

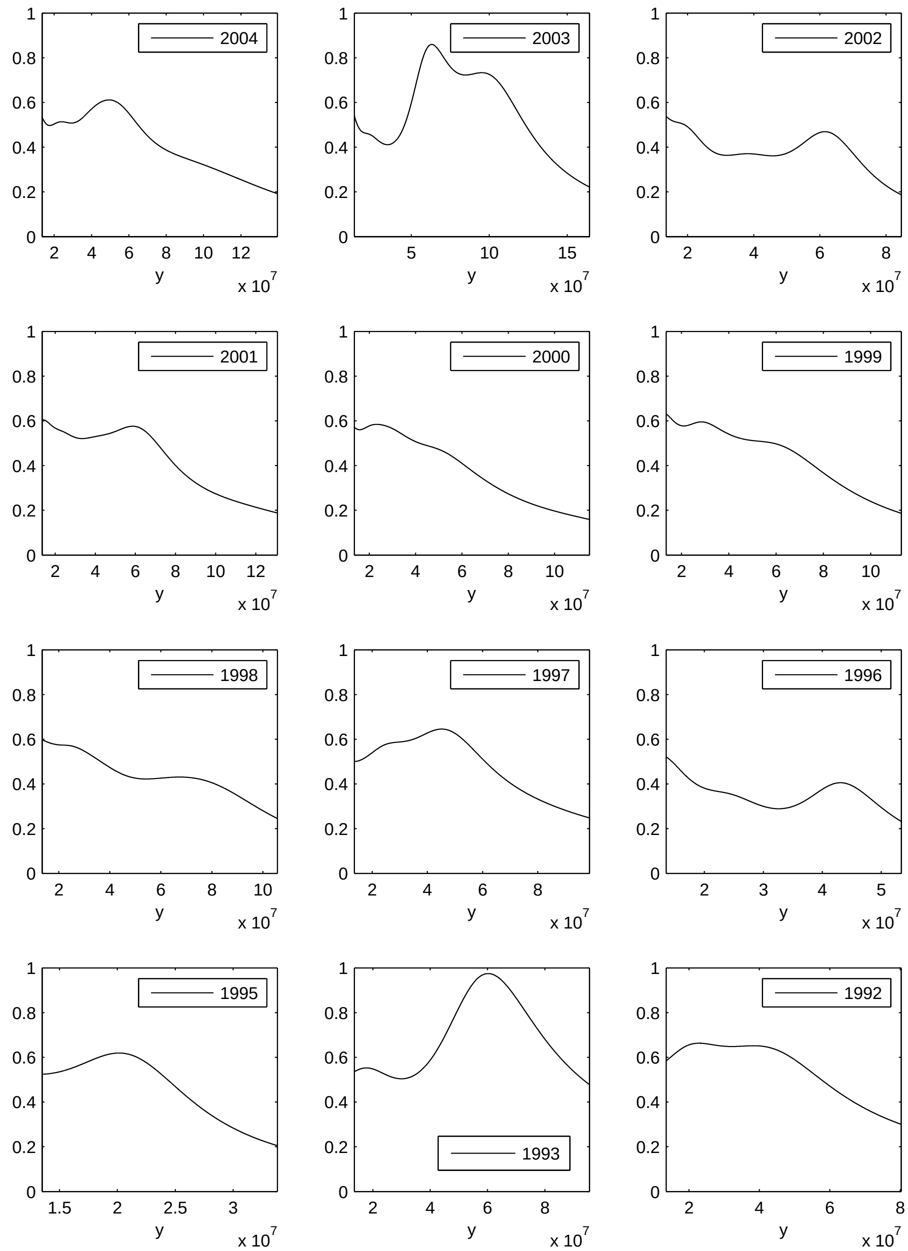

FIGURE A.1: Kernel smoothed $\frac{1-F_Y(y)}{yf_Y(y)}$ ratio in the U.S. $y \geq \$13,500,000$ (2004 dollars)

To obtain $G(\cdot)$, I solve the following maximization problem

$$\max_{G(\cdot)} \int_{\Theta} \frac{(F(\theta) - G(\theta))}{\theta^{2+\phi}} d\theta$$

s.t.

$$G(\theta) \geq F(\theta) \forall \theta \in \Theta \tag{A.23}$$

$$G'(\theta) \geq 0, \quad G(\underline{\theta}) = 0, \quad G(\bar{\theta}) = 1 \tag{A.24}$$

Let me consider the optimization problem ignoring constraint (A.24). The corresponding Lagrangian is

$$\mathfrak{L}(G(\theta), \gamma(\theta); F(\theta), \phi) = \int_{\Theta} \frac{(F(\theta) - G(\theta))}{\theta^{2+\phi}} d\theta + \int_{\Theta} \gamma(\theta)[G(\theta) - F(\theta)] d\theta$$

The first order condition with respect to $G(\theta)$ is

$$\frac{1}{\theta^{2+\phi}} = \gamma(\theta) \quad \forall \theta \in \Theta \tag{A.25}$$

Clearly, $\gamma(\theta) > 0 \forall \theta \in \Theta$. By slackness, we must have

$$\gamma(\theta)[G(\theta) - F(\theta)] = 0 \quad \forall \theta \in \Theta \tag{A.26}$$

thus, $G(\theta) = F(\theta) \forall \theta \in \Theta$. Clearly this solution satisfies (A.24). Finally, substituting

$g(\theta) = f(\theta)$ into (2.15) we obtain

$$\bar{\alpha} = \frac{\int_{\Theta} \frac{f(\theta)T'(y(\theta))}{\theta^{1+\phi}(1-T'(y(\theta)))} d\theta}{\int_{\Theta} \frac{f(\theta)}{\theta^{1+\phi}(1-T'(y(\theta)))} d\theta}.$$

Making a change of variable it can be shown that

$$\bar{\alpha} = \frac{\int_{\Phi^{-1}(\underline{\theta})}^{\Phi^{-1}(\bar{\theta})} \frac{f_Y(y)T'(y)}{y^{\phi}} dy}{\int_{\Phi^{-1}(\underline{\theta})}^{\Phi^{-1}(\bar{\theta})} \frac{f_Y(y)}{y^{\phi}} dy} = \frac{E_Y[y^{-\phi}T'(y)]}{E_Y[y^{-\phi}]} \tag{A.27}$$

□

Proof of Proposition 5

Proof. Since $T'(y(\bar{\theta})) \geq T'(y(\theta)) \forall \theta \in \Theta$ we have that $T'(y(\bar{\theta})) \geq \bar{\alpha}$. Since, $\frac{T'(y(\bar{\theta}))}{1-T'(y(\bar{\theta}))} = \frac{\alpha}{1-\alpha}\frac{\psi(\bar{\theta})}{f(\bar{\theta})}$ it follows that $\frac{\bar{\alpha}}{1-\bar{\alpha}} \leq \frac{\alpha}{1-\alpha}\frac{\psi(\bar{\theta})}{f(\bar{\theta})}$, thus $\frac{\psi(\bar{\theta})}{f(\bar{\theta})} \geq 1$. □

A.4 Constructing a Continuous and Differentiable Marginal Tax Schedule

In this Appendix, I construct a continuous and differentiable version of the statutory marginal tax schedule that I employ for my estimation of $\psi(\theta)$. I based my procedure in an insight from Stokey (2008). Without loss of generality, I will focus on a continuous and differentiable version of a statutory tax schedule with two income brackets of the form

$$T'(y) = \begin{cases} c_1 & \text{if } y \leq \tilde{y} \\ c_2 & \text{if } y > \tilde{y} \end{cases} \tag{A.28}$$

For $\epsilon > 0$, a continuous and differentiable version of the previous "step" tax is given by

$$T'(y) = \begin{cases} c_1 & \text{if } y \leq \tilde{y} - \frac{\epsilon}{2} \\ c_1 + \frac{2(c_2-c_1)}{\epsilon^2}(y - \tilde{y} + \frac{\epsilon}{2})^2 & \text{if } \tilde{y} - \frac{\epsilon}{2} < y < \tilde{y} \\ c_1 + \frac{(c_2-c_1)}{2} + \frac{2(c_2-c_1)}{\epsilon^2}(2(\tilde{y} + \frac{\epsilon}{2})y - y^2 - \tilde{y}^2 - \tilde{y}\epsilon) & \text{if } \tilde{y} \leq y < \tilde{y} + \frac{\epsilon}{2} \\ c_2 & \text{if } y \geq \tilde{y} + \frac{\epsilon}{2} \end{cases} \tag{A.29}$$

Figure A.2 shows the continuous and differentiable versions of the statutory taxes for the U.S. and the U.K. I use $\epsilon = 9,000$.

A.5 Estimation of Effective Average Labor Income Taxes

An alternative approach to the one followed in the main body of the paper is to use *effective* taxes as opposed to *statutory* ones. In this section, I describe the procedure to calculate effective marginal tax rates. For each country, I pooled the data on gross income and tax payments from 1995 to 2004, expressed in 2004 dollars. I construct a measure of effective *average* income tax rate as,

$$\bar{T}_i = \frac{t_i}{y_i}, \quad i = 1, ..., n,$$

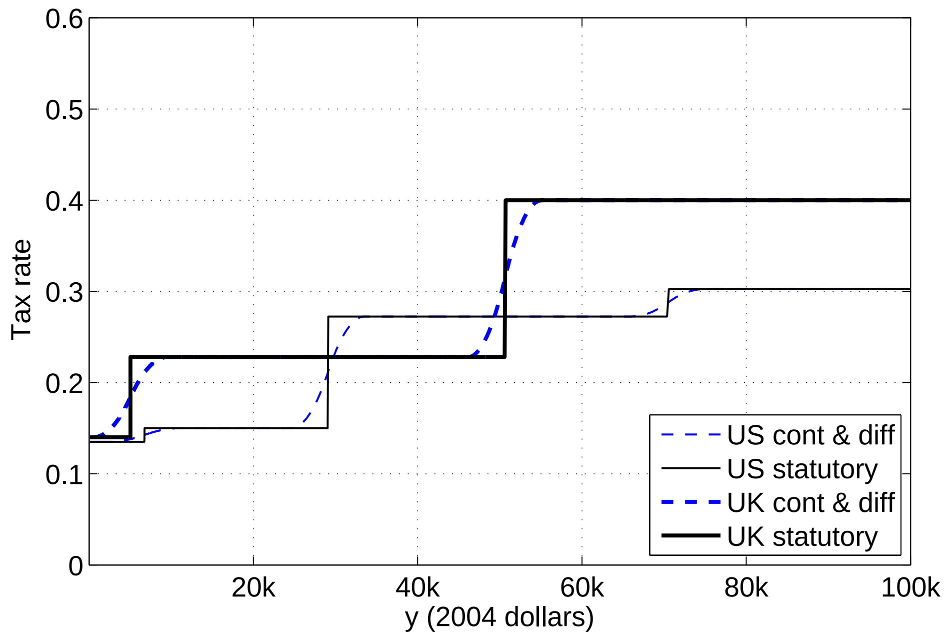

FIGURE A.2: Statutory and Smoothed Marginal Income Tax in the U.S. and the U.K. 1995-2004

where t_i is tax liabilities minus tax credits and y_i is annual gross income. Following Gouveia and Strauss (1994), I restrict the tax code to the functional form

$$T(y) = \alpha_0(y - (y^{-\alpha_1} + \alpha_2)^{-\frac{1}{\alpha_1}})$$

where $(\alpha_0, \alpha_1, \alpha_2)$ are parameters. It is easy to see that the average income tax, $\bar{T}(y)$, and the marginal tax, $T'(y)$, take the form

$$\bar{T}(y) = \alpha_0(1 - (1 + \alpha_2 y^{\alpha_1})^{-\frac{1}{\alpha_1}}) \tag{A.30}$$

$$T'(y) = \alpha_0(1 - (1 + \alpha_2 y^{\alpha_1})^{-\frac{1}{\alpha_1} - 1}). \qquad \text{(A.31)}$$

Notice that $\lim_{y\to\infty} \bar{T} = \lim_{y\to\infty} T'(y) = \alpha_0$. I estimate the parameters $(\alpha_0, \alpha_1, \alpha_2)$ using equation (A.30) by non linear least squares. I considered observations satisfying $y_i > 3,000$ and $0 < \bar{T}_i < 0.50$. The estimated effective marginal income tax schedules are shown in Figure A.3 while the results of the regressions can be seen in Table A.2. Finally, Figure A.4 shows the estimates of the contribution to the consumption externality, the ratio $\frac{\psi(\cdot)}{f(\cdot)}$ using effective marginal tax rates. The caveat of this estimation is that the functional form of marginal taxes, expression (A.31) specifies a marginal tax rate of zero at null income level. Thus, the estimated α -under the assumption that the planner is utilitarian- is very close to zero which makes the expression for $\psi(\theta)$ practically undetermined.[3] To go around thus difficulty, I estimate $\psi(\theta)$ assuming that the minimum gross income level is $3,000 dollars expressed in 2004 purchasing power. This drives the parameter α up and makes the evaluation of $\psi(\theta)$ possible. Nevertheless, the resulting $\psi(\theta)$ is very sensitive to the choice of the minimum gross income level. This difficulty does not arise if one uses the statutory marginal tax schedule as it is flat at the bottom of the gross income distribution.

[3]To see this more clearly, recall from Proposition 4 that when the planner is utilitarian, $\alpha = \frac{E_Y[y^{-\phi}T'(y)]}{E_Y[y^{-\phi}]}$ as it attains its upper bound. If follows that the marginal tax at low income levels are weighted more heavily than higher income levels as $y^{-\phi}$ decreases rapidly making the contribution of $T'(y)$ after certain income level effectively null.

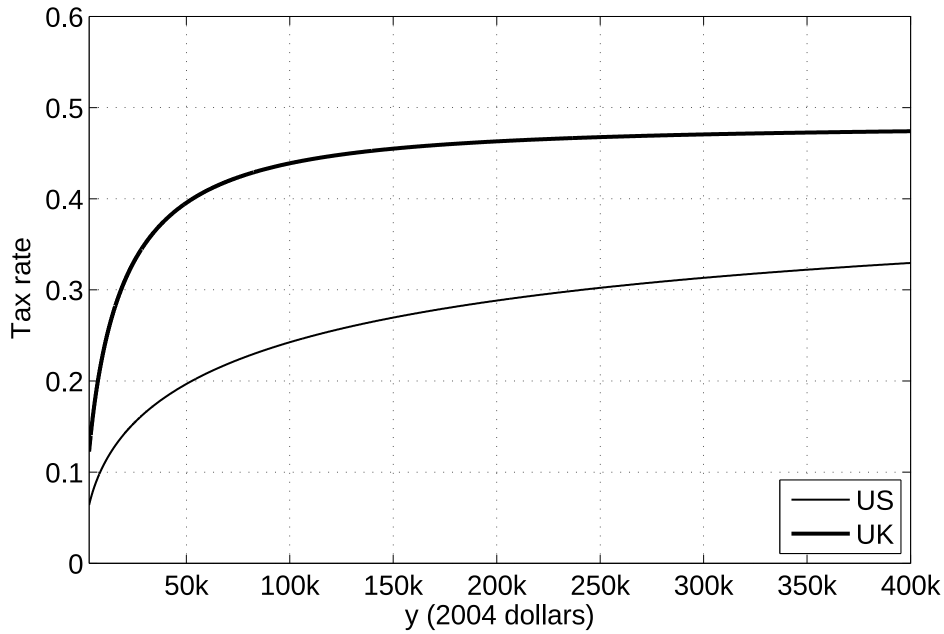

FIGURE A.3: Effective Marginal Income Tax in the U.S. and the U.K. 1995-2004

TABLE A.2: Estimation of Average Labor Income Tax Schedule

Coefficient	$\bar{T}(y)^{U.S.}$	$\bar{T}(y)^{U.K.}$
α_0	0.41865	0.48095
	(0.00085)	(0.00083)
α_1	0.50700	0.74726
	(0.00192)	(0.00158)
α_2	0.00099	0.00034
	(0.00001)	$(3.49e^{-06})$
n	545,454	1,154,609
R^2	0.8399	0.9301

$\bar{T}(y)^{U.S.}$: Average tax rate in the United States, $\bar{T}(y)^{U.K.}$: Average tax rate in the United Kingdom. Regression estimated by NLLS and robust standard errors. y is gross income measured in 2004 dollars and n is sample size.

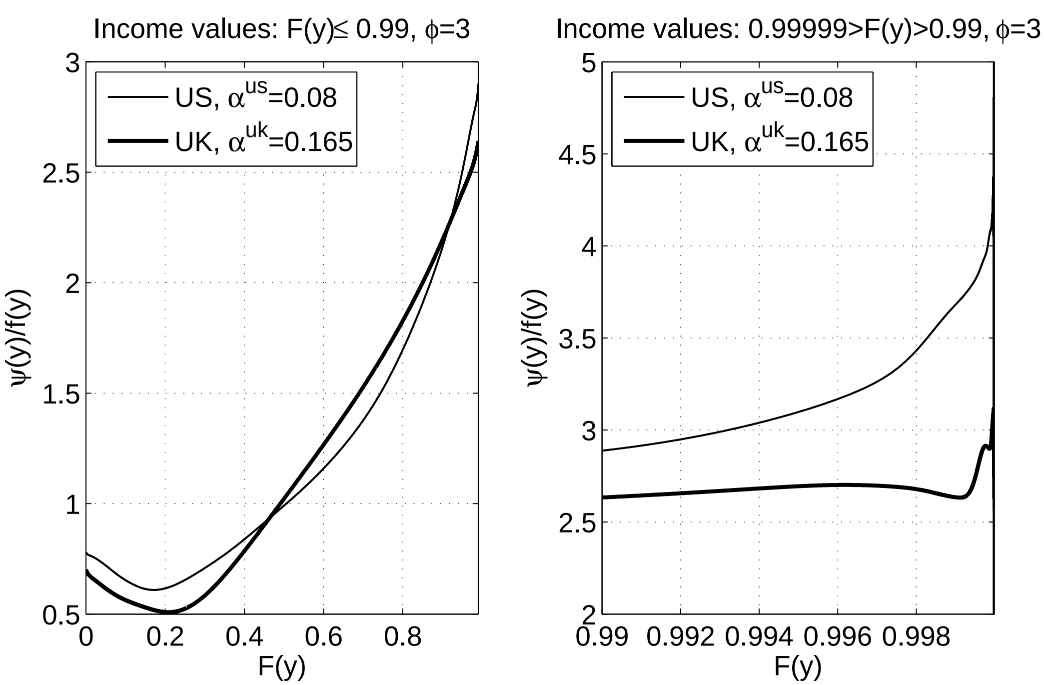

FIGURE A.4: Contribution to Consumption Externality in the U.S. and the U.K. 1995-2004

Appendix B

Appendix Chapter 3

B.1 Proofs

Proof of Proposition 6

Proof. Replicate proof of Proposition 7 setting $\eta(\theta) = 0 \quad \forall \theta \in \Theta$. □

Proof of Proposition 7

Proof. The first step is to transform the continuum of incentive compatibility constraints (3.3) into a first order condition. Let $\sigma(\theta) = \theta'$ and

$$W(\theta, \theta') \equiv u(c_n(\theta'), c_l(\theta')) - \alpha C - v\left(\frac{y(\theta')}{\theta}\right) \tag{B.1}$$

A necessary condition for truthful revelation of type is $\frac{\partial W(\theta,\theta')}{\partial \theta'}|_{\theta'=\theta} = 0$, therefore it follows that

$$u_{c_n}(c_n(\theta), c_l(\theta))\frac{\partial c_n(\theta)}{\partial \theta} + u_{c_l}(c_l(\theta), c_l(\theta))\frac{\partial c_l(\theta)}{\partial \theta} = v'\left(\frac{y(\theta)}{\theta}\right)\frac{y'(\theta)}{\theta} \quad \forall \theta \in \Theta \tag{B.2}$$

Moreover, under truthful revelation $W(\theta) = u(c_n(\theta), c_l(\theta)) - \alpha C - v\left(\frac{y(\theta)}{\theta}\right)$ and thus, $W'(\theta) = u_{c_n}(c_n(\theta), c_l(\theta))\frac{\partial c_n(\theta)}{\partial \theta} + u_{c_l}(c_n(\theta), c_l(\theta))\frac{\partial c_l(\theta)}{\partial \theta} - v'\left(\frac{y(\theta)}{\theta}\right)\frac{y'(\theta)}{\theta} + v'\left(\frac{y(\theta)}{\theta}\right)\frac{y(\theta)}{\theta^2}$, which together with (B.2) becomes

$$W'(\theta) = v'\left(\frac{y(\theta)}{\theta}\right)\frac{y(\theta)}{\theta^2} \quad \forall \theta \in \Theta. \tag{B.3}$$

Define the expenditure function $e(W(\theta), c_l(\theta), y(\theta), C; \theta)$ to satisfy $W(\theta) = u(e, c_l) - \alpha C - v\left(\frac{y(\theta)}{\theta}\right)$. Thus, the planner problem can be restated as

$$\max_{W(\cdot), c_l(\theta), y(\cdot), C, \kappa} \int_\Theta W(\theta) g(\theta) d\theta \tag{B.4}$$

s.t

$$\int_\Theta c_l(\theta) f(\theta) d\theta + \int_\Theta e(W(\theta), c_l(\theta), y(\theta), C; \theta) f(\theta) d\theta = \int_\Theta y(\theta) f(\theta) d\theta \tag{B.5}$$

$$W'(\theta) = v'\left(\frac{y(\theta)}{\theta}\right)\frac{y(\theta)}{\theta^2} \quad \forall \theta \in \Theta \tag{B.6}$$

$$e_{c_l}(W(\theta), c_l(\theta), y(\theta), C; \theta) = \kappa \quad \forall \theta \in \Theta \tag{B.7}$$

$$C \equiv \int_{\Theta} \left[\omega e(W(\theta), c_l(\theta), y(\theta), C; \theta) + (1-\omega) c_l(\theta)\right] \psi(\theta) d\theta \tag{B.8}$$

The corresponding Lagrangian is

$$\mathfrak{L}(W(\theta), c_l(\theta), y(\theta), C, \kappa, \lambda, \mu(\theta), \gamma, \eta(\theta)) = \int_{\Theta} W(\theta) g(\theta) d\theta$$

$$-\lambda \int_{\Theta} \left[c_l(\theta) + e(W(\theta), c_l(\theta), y(\theta), C; \theta) - y(\theta)\right] f(\theta) d\theta$$

$$+ \int_{\Theta} \mu(\theta) \left[W'(\theta) - v'\left(\frac{y(\theta)}{\theta}\right) \frac{y(\theta)}{\theta^2}\right] d\theta$$

$$+\gamma \left[C - \int_{\Theta} [\omega e(W(\theta), c_l(\theta), y(\theta), C; \theta) + (1-\omega) c_l(\theta)] \psi(\theta) d\theta\right]$$

$$+ \int_{\Theta} \eta(\theta) \left[e_{c_l}(W(\theta), c_l(\theta), y(\theta), C; \theta) - \kappa\right] d\theta \tag{B.9}$$

Using integration by parts, it follows that

$$\int_{\Theta} \mu(\theta) W'(\theta) d\theta = \mu(\bar{\theta}) W(\bar{\theta}) - \mu(\underline{\theta}) W(\underline{\theta}) - \int_{\Theta} \mu'(\theta) W(\theta) d\theta \tag{B.10}$$

thus, we can reexpress the above Lagrangian as

$$\mathfrak{L}(W(\theta), c_l(\theta), y(\theta), C, \kappa, \lambda, \mu(\theta), \gamma, \eta(\theta)) = \int_{\Theta} W(\theta) g(\theta) d\theta$$

$$-\lambda \int_{\Theta} [c_l(\theta) + e(W(\theta), c_l(\theta), y(\theta), C; \theta) - y(\theta)] f(\theta) d\theta + \mu(\bar{\theta}) W(\bar{\theta}) - \mu(\underline{\theta}) W(\underline{\theta})$$

$$-\int_{\Theta}\mu'(\theta)W(\theta)d\theta-\int_{\Theta}\mu(\theta)v'\left(\frac{y(\theta)}{\theta}\right)\frac{y(\theta)}{\theta^2}d\theta$$

$$+\gamma\left[C-\int_{\Theta}[\omega e(W(\theta),c_l(\theta),y(\theta),C;\theta)+(1-\omega)c_l(\theta)]\psi(\theta)d\theta\right]$$

$$+\int_{\Theta}\eta(\theta)\left[e_{c_l}(W(\theta),c_l(\theta),y(\theta),C;\theta)-\kappa\right]d\theta \tag{B.11}$$

Assuming interior solution, it follows from first order conditions that

$W(\theta)$:

$$g(\theta)-\lambda f(\theta)e_W(W(\theta),c(\theta)_l,y(\theta),C;\theta)-\mu'(\theta)$$

$$-\gamma\omega\psi(\theta)e_W(W(\theta),c_l(\theta),y(\theta),C;\theta)=0 \tag{B.12}$$

$y(\theta)$:

$$-\lambda e_y(W(\theta),c_l(\theta),y(\theta),C;\theta)f(\theta)+\lambda f(\theta)-\frac{\mu(\theta)}{\theta^2}v'\left(\frac{y(\theta)}{\theta}\right)\left[1+\frac{1}{\epsilon(\theta)}\right]$$

$$-\gamma\omega e_y(W(\theta),c_l(\theta),y(\theta),C;\theta)\psi(\theta)=0 \tag{B.13}$$

$c_l(\theta)$:

$$-\lambda e_{c_l}(W(\theta),c_l(\theta),y(\theta),C;\theta)f(\theta)-\lambda f(\theta)-\gamma\omega e_{c_l}(W(\theta),c_l(\theta),y(\theta),C;\theta)\psi(\theta)$$

$$- \gamma(1-\omega)\psi(\theta) + \eta(\theta)e_{c_l c_l}(W(\theta), c_l(\theta), y(\theta), C; \theta) = 0 \tag{B.14}$$

C :

$$-\lambda \int_{\Theta} e_C(W(\theta), c_l(\theta), y(\theta), C; \theta) f(\theta) d\theta + \gamma - \gamma\omega \int_{\Theta} e_C(W(\theta), c_l(\theta), y(\theta), C; \theta)\psi(\theta) d\theta = 0 \tag{B.15}$$

κ :

$$\int_{\Theta} \eta(\theta) d\theta = 0 \tag{B.16}$$

together with the boundary conditions $\mu(\underline{\theta}) = \mu(\bar{\theta}) = 0$ and where $\epsilon(\theta) \equiv \frac{v'(\frac{y(\theta)}{\theta})}{v''(\frac{y(\theta)}{\theta})\frac{y(\theta)}{\theta}}$. Moreover, by implicit differentiation of $W(\theta)$ it follows that $e_W(W(\theta), c_l(\theta), y(\theta), C; \theta) = \frac{1}{u_{c_n}(c_n(\theta), c_l(\theta))}$, $e_y(W(\theta), c_l(\theta), y(\theta), C; \theta) = \frac{v'\left(\frac{y(\theta)}{\theta}\right)\frac{1}{\theta}}{u_{c_n}(c_n(\theta), c_l(\theta))}$, $e_{c_l}(W(\theta), c_l(\theta), y(\theta), C; \theta) = -\frac{u_{c_l}(c_n(\theta), c_l(\theta))}{u_{c_n}(c_n(\theta), c_l(\theta))}$ and $e_C(W(\theta), c_l(\theta), y(\theta), C; \theta) = \frac{\alpha}{u_{c_n}(c_n(\theta), c_l(\theta))}$. Moreover, observe that $e_{c_l c_l}(W(\theta), c_l(\theta), y(\theta), C; \theta) = \left[\frac{u_{c_l}(\theta)}{u_{c_n}(\theta)}\frac{u_{c_n c_l}(\theta)}{u_{c_n}(\theta)} - \frac{u_{c_l c_l}(\theta)}{u_{c_n}(\theta)}\right]$. The result follows after manipulating (B.12)-(B.16).

$\square$

B.1.1 Proof of Proposition 8

Proof. Taking first order conditions in agent's problem we have

$$\frac{T'(y(\theta))}{1 - T'(y(\theta))} = \frac{u_{c_n}(c_n^{eq}(\theta), c_l^{eq}(\theta))}{v'\left(\frac{y^{eq}(\theta)}{\theta}\right)\frac{1}{\theta}} - 1 \quad \forall \theta \in \Theta \tag{B.17}$$

and

$$\frac{u_{c_l}(c_n^{eq}(\theta), c_l^{eq}(\theta))}{u_{c_n}(c_n^{eq}(\theta), c_l^{eq}(\theta))} = 1 + \tau \quad \forall \theta \in \Theta \tag{B.18}$$

hence from (3.29)-(3.31) it follows that

$$\frac{u_{c_l}(c_n^{eq}(\theta), c_l^{eq}(\theta))}{u_{c_n}(c_n^{eq}(\theta), c_l^{eq}(\theta))} - 1 = \frac{\frac{\alpha(1-2\omega)}{\lambda}\int_\Theta \frac{\psi(\theta)}{B^*(\theta)}d\theta}{\int_\Theta \frac{f(\theta)}{B^*(\theta)}d\theta + \frac{\alpha\omega}{\lambda}\int_\Theta \frac{\psi(\theta)}{B^*(\theta)}d\theta} \quad \forall \theta \in \Theta \tag{B.19}$$

and

$$\frac{u_{c_n}(c_n^{eq}(\theta), c_l^{eq}(\theta))}{v'\left(\frac{y^{eq}(\theta)}{\theta}\right)\frac{1}{\theta}} - 1 = \frac{\alpha}{\lambda}\frac{\omega\psi(\theta)}{f(\theta)} + \frac{u_{c_n}(c_n^*(\theta), c_l^*(\theta))}{\theta f(\theta)}\left[1 + \frac{1}{\epsilon^*(\theta)}\right] I^*(\theta) \quad \forall \theta \in \Theta \tag{B.20}$$

Finally notice that the fact that the government balances its budget implies that

$$\int_\Theta c_n^{eq}(\theta) f(\theta) d\theta + \int_\Theta c_l^{eq}(\theta) f(\theta) d\theta = \int_\Theta y^{eq}(\theta) f(\theta) d\theta \tag{B.21}$$

Thus, from (B.19)-(B.21) we obtain $\{c_n^{eq}(\theta), c_l^{eq}(\theta), y^{eq}(\theta)\}_{\theta\in\Theta} = \{c_n^*(\theta), c_l^*(\theta), y^*(\theta)\}_{\theta\in\Theta}$.

$\square$

B.2 Solving the Model Numerically

I solve the model by casting it into a system of differential-algebraic equations. Let

$$r(\theta) \equiv \left(1 + \frac{1}{\epsilon(\theta)}\right)^{-1}\left[\frac{1}{v'(\frac{y(\theta)}{\theta})} - \frac{e_y(\theta)}{v'(\frac{y(\theta)}{\theta})}\left(1 + \frac{\gamma\omega}{\lambda}\frac{\psi(\theta)}{f(\theta)}\right)\right] \tag{B.22}$$

thus equation (B.13) becomes

$$r(\theta) = \frac{\mu(\theta)}{\lambda\theta^2 f(\theta)} \tag{B.23}$$

Differentiating (B.23) and using (B.22) we obtain

$$r'(\theta) = \frac{\mu'(\theta)}{\lambda f(\theta)\theta^2} - r(\theta)\left[\frac{2}{\theta} + \frac{f'(\theta)}{f(\theta)}\right] \tag{B.24}$$

Finally, using equation (B.12) we have

$$r'(\theta) = \frac{g(\theta)}{f(\theta)\lambda\theta^2} - \frac{e_W(\theta)}{\theta^2} - \frac{\gamma\omega}{\lambda}\frac{\psi(\theta)}{f(\theta)}\frac{e_W(\theta)}{\theta^2} - r(\theta)\left[\frac{2}{\theta} + \frac{f'(\theta)}{f(\theta)}\right] \tag{B.25}$$

From incentive compatibility we directly have

$$W'(\theta) = v'\left(\frac{y(\theta)}{\theta}\right)\frac{y(\theta)}{\theta^2}. \tag{B.26}$$

Thus, equations (B.25) and (B.26) are the differential equations of the system. The algebraic equations of the system are the following: from equation (B.14) we obtain

$$-e_{c_l}(\theta) - 1 - \frac{\gamma\omega\psi(\theta)e_{c_l}(\theta)}{\lambda f(\theta)} - \frac{\gamma(1-\omega)\psi(\theta)}{\lambda f(\theta)} + \eta(\theta)\frac{e_{c_l c_l}(\theta)}{\lambda f(\theta)} = 0 \tag{B.27}$$

from problem's restriction we have

$$e_{c_l}(\theta) - \kappa = 0. \tag{B.28}$$

Thus, equations (B.22), (B.27) and (B.28) are the algebraic equations of the system. The variables of the system are $[W(\theta), r(\theta), y(\theta), c_l(\theta), \eta(\theta)]'$. Once the value of these variables is known, it is straightforward to calculate the value of $c_n(\theta)$. The solution of the system has to satisfy the feasibility constraint and $\int_\Theta \eta(\theta)d\theta = 0$. This can be done by adjusting the value of κ and the vector of initial conditions.[1] Moreover, observe that an initial guess must be followed to calculate C, $\int_\Theta c_n(\theta)\psi(\theta)d\theta$ and $\int_\Theta c_n(\theta)f(\theta)d\theta$ which are required to obtain the solution of the system through the Lagrange multipliers of the system. We can iterate the previously guessed values until convergence.

B.3 A Case of Zero Labor Marginal Tax Schedule

Proposition 10. *Suppose $\omega = 0$, $g(\theta) = f(\theta)\forall\theta \in \Theta$ and*

i) $u(c_n, c_l) = \left[\eta c_n^{1-\frac{1}{\sigma}} + (1-\eta)c_l^{1-\frac{1}{\sigma}}\right]^{\frac{\sigma}{\sigma-1}}$,

ii) $\frac{(1-\eta)}{\eta}\left(\frac{c_l^*(\theta)}{c_n^*(\theta)}\right)^{-\frac{1}{\sigma}} = \frac{(1-\eta)}{\eta}\left(\frac{c_l^*(\theta')}{c_n^*(\theta')}\right)^{-\frac{1}{\sigma}} = q \quad \forall\theta, \theta' \in \Theta$

then $T'(y^*(\theta)) = 0\forall\theta \in \Theta$.

Proof. The first step is to show that in equilibrium, the marginal utility of c_n is equal across agents. Observe that

$$\frac{u_{c_n}(c_n(\theta), c_l(\theta))}{u_{c_n}(c_n(\theta'), c_l(\theta'))} = \frac{\left[\eta c_n(\theta)^{1-\frac{1}{\sigma}} + (1-\eta)c_l(\theta)^{1-\frac{1}{\sigma}}\right]^{\frac{1}{\sigma-1}} \eta c_n(\theta)^{\frac{1}{\sigma}}}{\left[\eta c_n(\theta')^{1-\frac{1}{\sigma}} + (1-\eta)c_l(\theta')^{1-\frac{1}{\sigma}}\right]^{\frac{1}{\sigma-1}} \eta c_n(\theta')^{\frac{1}{\sigma}}} \forall\theta, \theta' \in \Theta \quad \text{(B.29)}$$

[1]From the fact that $\mu(\bar{\theta})$ it must be that $r(\bar{\theta}) = 0$ whenever $\omega = 0$.

Rearranging we obtain

$$\frac{u_{c_n}(c_n(\theta), c_l(\theta))}{u_{c_n}(c_n(\theta'), c_l(\theta'))} = \frac{\left[\eta + (1-\eta)\left(\frac{c_l(\theta)}{c_n(\theta)}\right)^{1-\frac{1}{\sigma}}\right]^{\frac{1}{\sigma-1}}}{\left[\eta + (1-\eta)\left(\frac{c_l(\theta')}{c_n(\theta')}\right)^{1-\frac{1}{\sigma}}\right]^{\frac{1}{\sigma-1}}} \forall \theta, \theta' \in \Theta \tag{B.30}$$

Using assumption ii) we obtain that in equilibrium $\frac{u_{c_n}(c_n^*(\theta), c_l^*(\theta))}{u_{c_l}(c_n^*(\theta), c_l^*(\theta))} = 1$. Finally, imposing $\omega = 0$ and $g(\theta) = f(\theta)$ it follows from equations (3.25) and (3.27) that $I^*(\theta) = 0 \forall \theta \in \Theta$ which implies by equation (3.31) that $T'(y^*(\theta)) = 0 \forall \theta \in \Theta$. □

B.4 Additional Sensitivity Analysis

In this section I present a sensitivity analysis of some of the parameters not presented in Chapter 3. Figure B.1 shows comparative statics of changes in the jealousy parameter, α, when the skills distribution is Pareto. As expected, both taxes are increasing in this parameter. Also, observe that when α is higher, the luxury good consumption distribution experience sharper changes at the top upon the introduction of the positional flat tax. Figure B.2 shows changes in η, the parameter that measures the share of disposable income assigned to each good. Figure B.3 presents the endogenous distributions and taxes for different elasticity of substitution across goods, the parameter σ. Figure B.4 displays distributions and taxes for changes in the ratio $\psi(\theta)/f(\theta)$. Figure B.5 also displays changes in the ratio $\psi(\theta)/f(\theta)$. Notice that in this case, the previous ratio is decreasing in income! Figure B.6 shows the endogenous distributions and taxes for a very high elasticity of labor supply. Finally, Figure B.7 shows the endogenous distributions and taxes when the planner is not utilitarian and places more weight on low skilled agents.

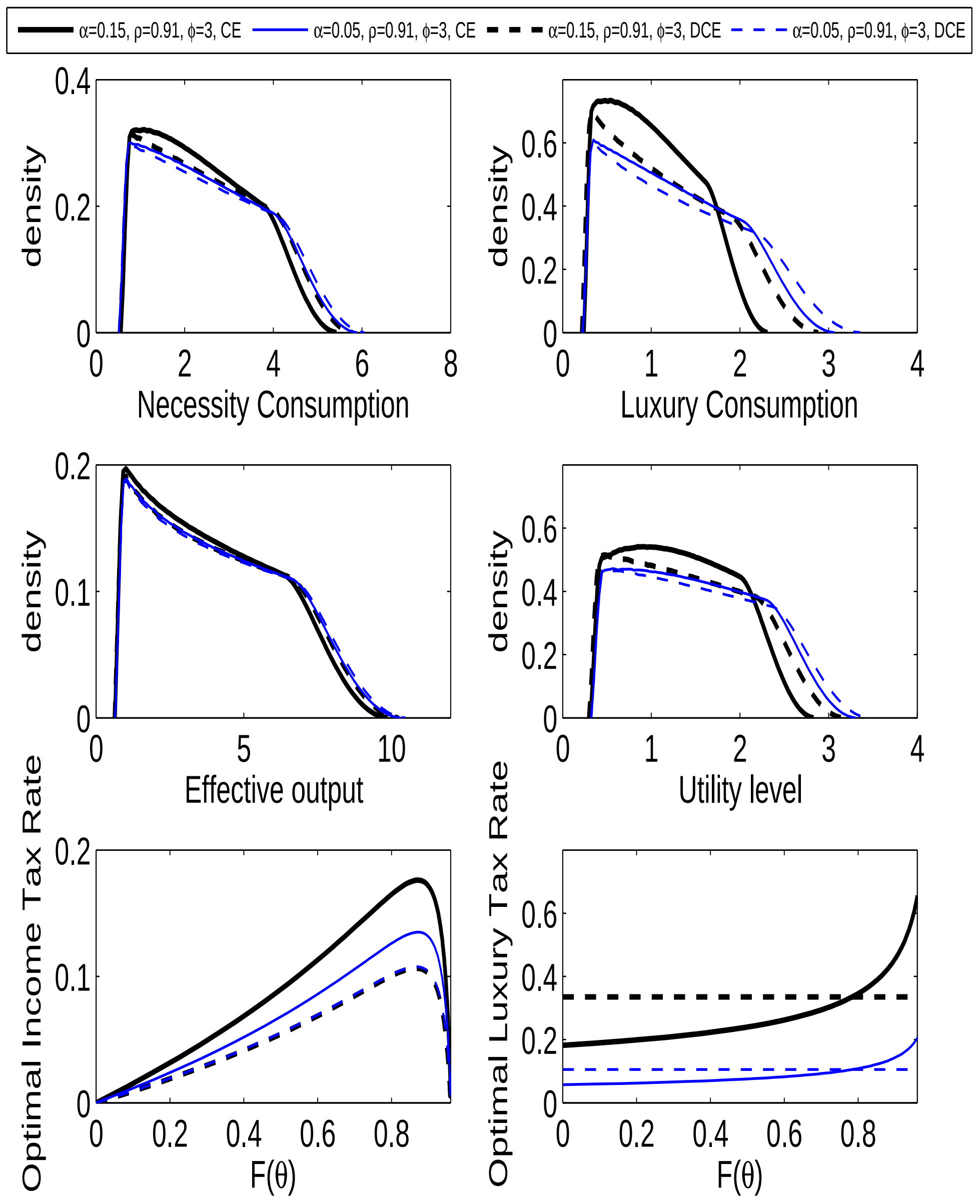

FIGURE B.1: Endogenous distributions and optimal taxes when distribution of skills is Pareto and preferences are non-homothetic. Changes in α.

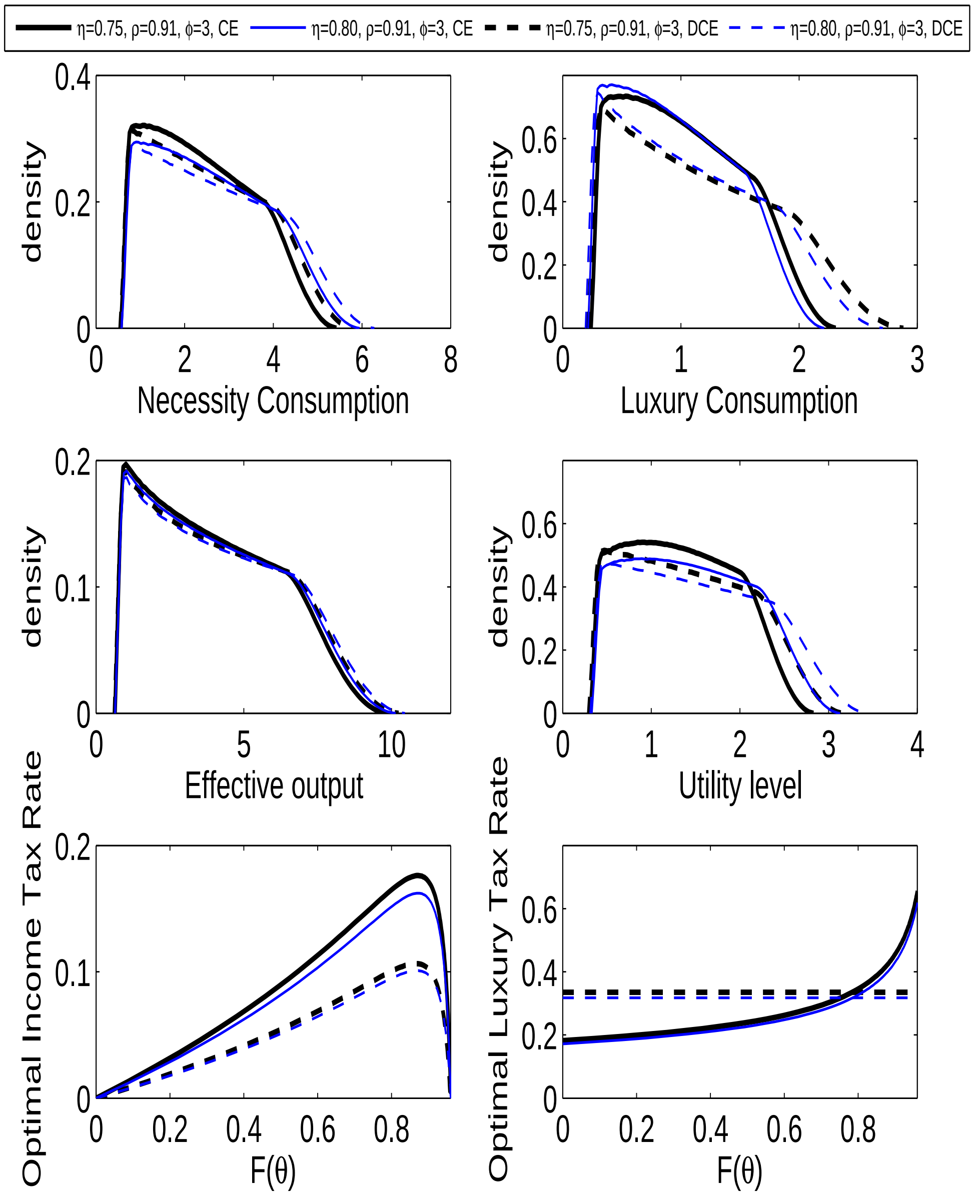

FIGURE B.2: Endogenous distributions and optimal taxes when distribution of skills is Pareto and preferences are non-homothetic. Changes in η. When $\eta = 0.80$ we have $\lambda_{\Theta} = 0.15\%$, $\lambda_L = 1.21\%$ and $\lambda_H = -5\%$

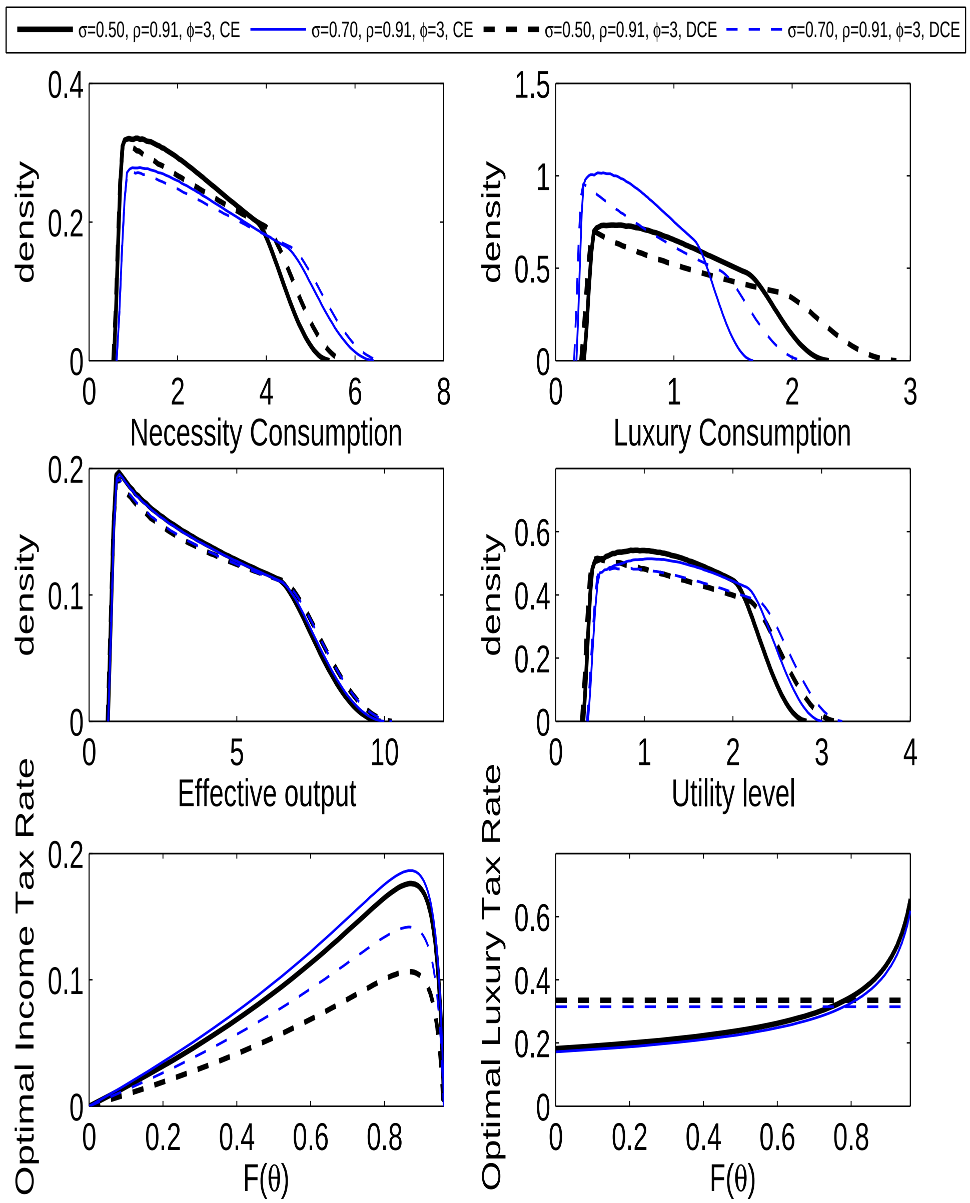

FIGURE B.3: Endogenous distributions and optimal taxes when distribution of skills is Pareto and preferences are non-homothetic. Changes in σ. When $\sigma = 0.70$ we have $\lambda_\Theta = 0.10\%$, $\lambda_L = 0.87\%$ and $\lambda_H = -3.93\%$

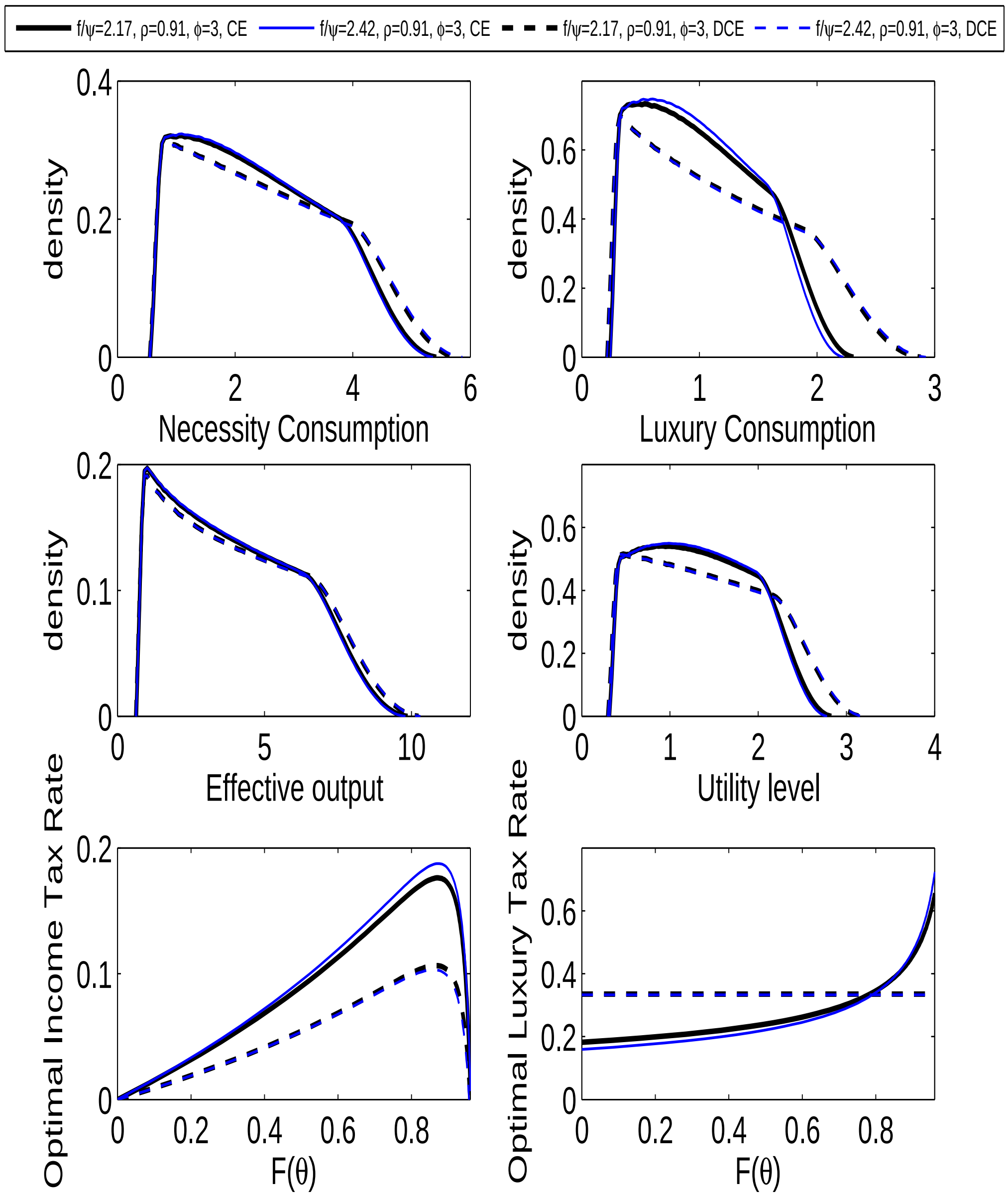

FIGURE B.4: Endogenous distributions and optimal taxes when distribution of skills is Pareto and preferences are non-homothetic. Changes in k_ψ. When at the top, $\psi/f = 2.42$ we have $\lambda_\Theta = 0.21\%$, $\lambda_L = 1.62\%$ and $\lambda_H = -6.77\%$

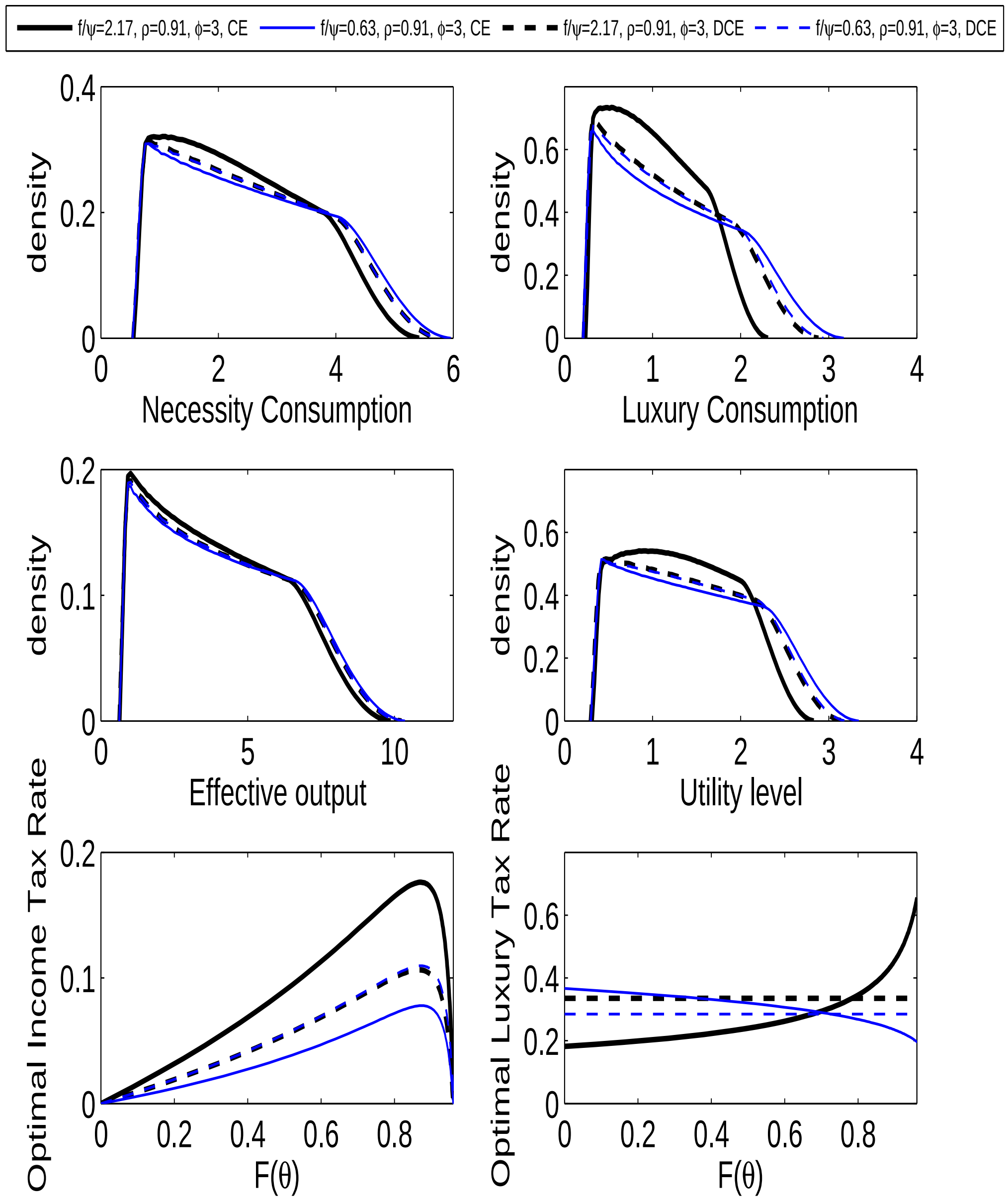

FIGURE B.5: Endogenous distributions and optimal taxes when distribution of skills is Pareto and preferences are non-homothetic. Changes in k_ψ. When at the top, $\psi/f = 0.63$ and at the bottom $\psi/f = 1.20$ we have $\lambda_\Theta = 0.03\%$, $\lambda_L = -0.47\%$ and $\lambda_H = 2.51\%$

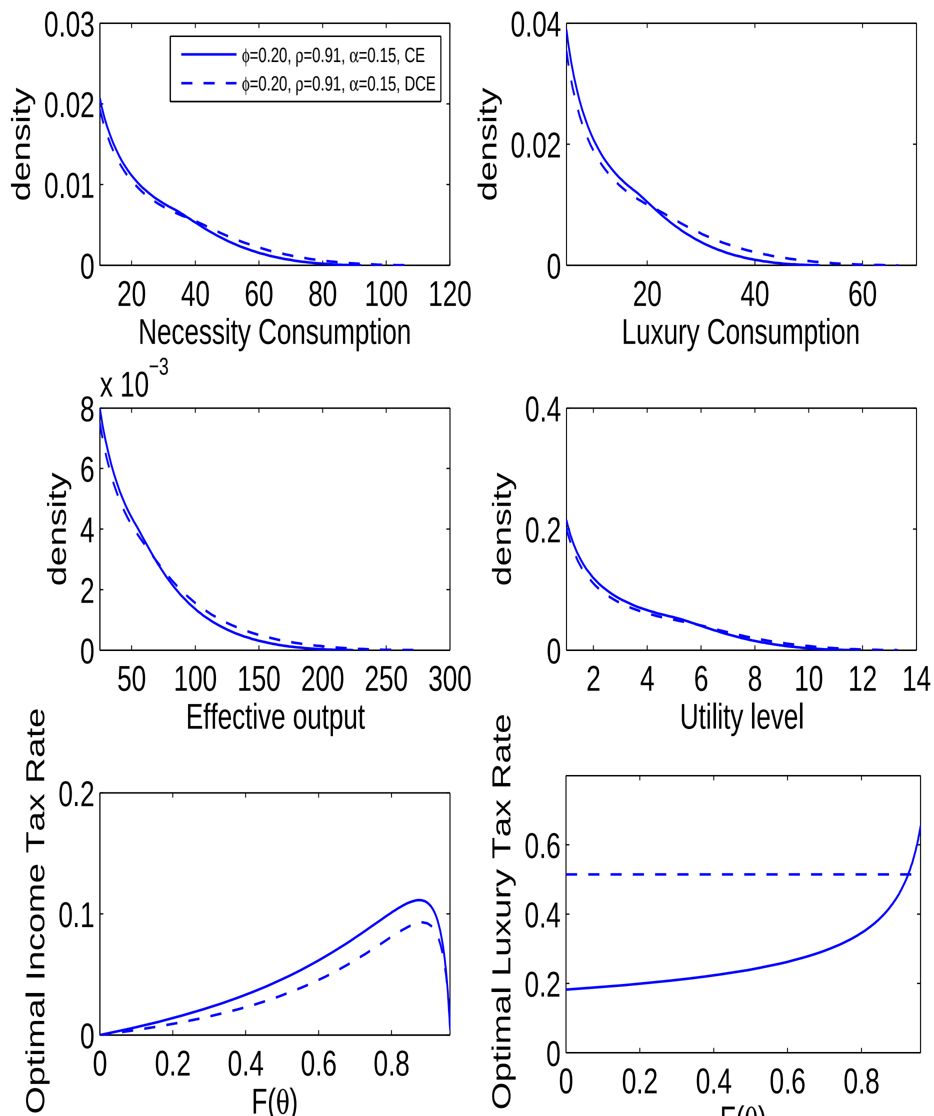

FIGURE B.6: Upper tail of endogenous distributions and optimal taxes when distribution of skills is Pareto and preferences are non-homothetic. A very elastic labor supply ($\phi = 0.20$). We obtained $\lambda_{\Theta} = 0.32\%$, $\lambda_L = 1.85\%$ and $\lambda_H = -0.93\%$

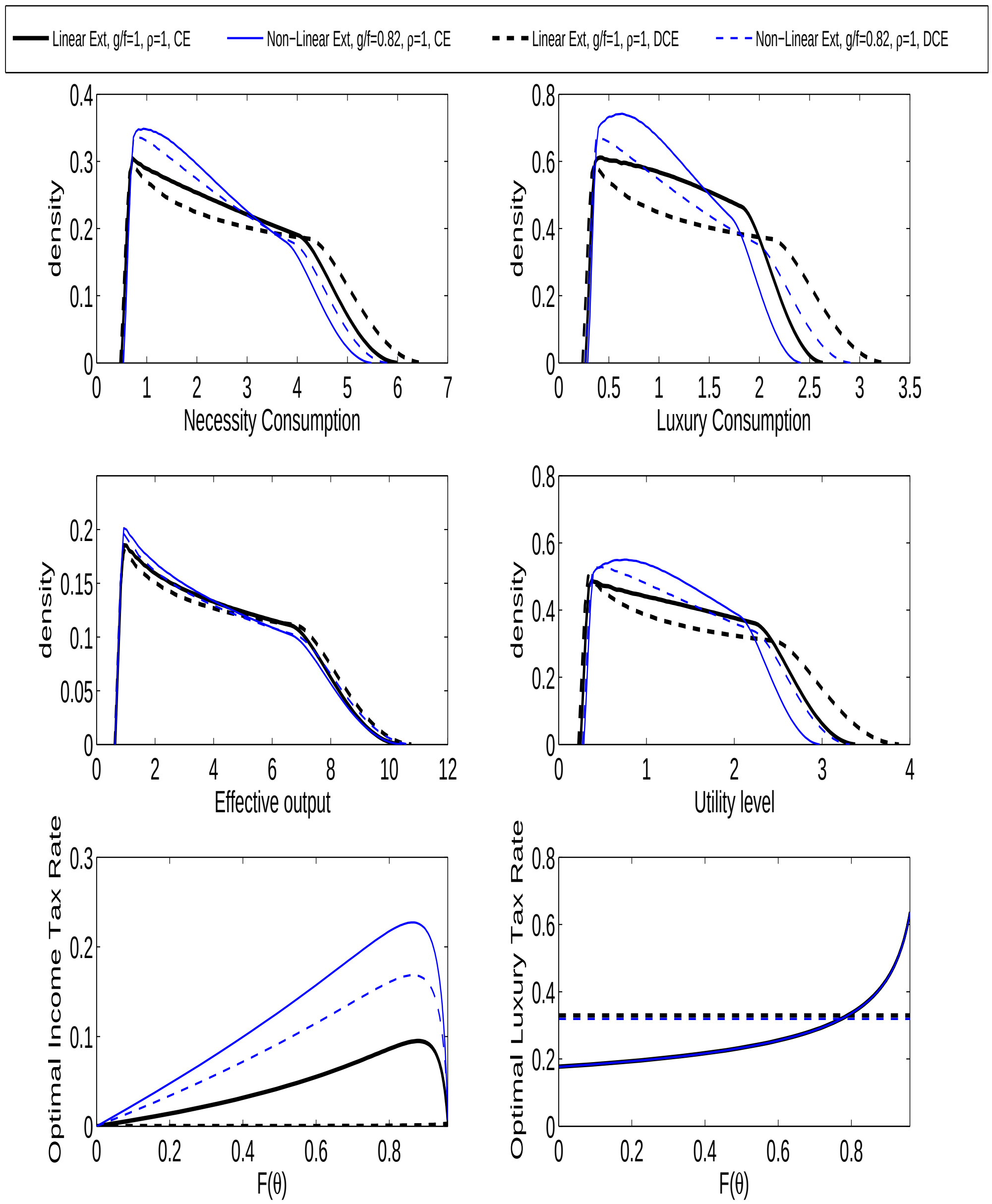

FIGURE B.7: Endogenous distributions and optimal taxes when distribution of skills is Pareto and preferences are homothetic. Changes in k_g from $g/f = 1$ to $g/f = 0.82$ at the top.