Tax Simplicity and Universal Basic Income*

Cabral-Lassalle, Guillermo†

November 14, 2018

Latest version: https://github.com/cabralguiller/cabralguillermo/blob/master/JMP.pdf

Abstract

In this paper I analyze the convenience of a cash transfer program, or universal basic income (UBI), combined with a flat or linear marginal tax rate on income. Proponents of a policy of cash transfers argue that if combined with a simplification of welfare programs and the tax system, it could generate enough benefits through a reduction of administrative costs and a reduction of distortions, especially in the labor market.

This idea is not new in economics, and numerical results in models as in Mirrlees (1971) have lump-sum components that can be interpreted as UBI. Also, Mirrlees has noted that the optimal marginal non-linear tax rate is close to a linear tax system. I start by double-checking Mirrlees' calculations and determining that those numbers hold for a wide range of parameters. For a utilitarian planner, if a simplification of the tax system results in a value of 0.18% - 0.5% of GDP, a linear tax system is desirable.

Second, I analyse the effects of introducing uncertainty in the optimal non-linear tax system. I introduce uncertainty in the preference of the agent. I show that heterogeneity could be a factor in making the optimal non-linear tax system closer to a linear tax system. In the extreme case where heterogeneity increases to its maximum possible value, the optimal tax system tends toward a linear tax system.

*I am grateful to Larry Jnoes, V.V. Chari, and Christopher Phelan for their guidance and support throughout this project. I also thank the participants of the Growth and Development Workshop at the University of Minnesota for their helpful comments.

††University of Minnesota. 1925 Fourth Street South, Minneapolis, MN 55405 (email: cabr0025@umn.edu).

1 Introduction

Is giving cash directly to people a radical idea? Although this idea is not new[1], he growing interest in universal basic income recently[2] has created a debate that should be informed by economic theory. Among other arguments for UBI[3], I focus on the argument that UBI generates benefits by simplifying the tax system and welfare programs. This argument states that a lump-sum transfer combined with a constant marginal tax rate on income [4] is preferable to a tax system that includes several tax rate brackets and welfare programs for those at the bottom of the distribution. .

To answer the question above, I'll use the model developed by Mirrlees (1971). In this model the tax system is generally described as a lump-sum transfer to the individual (as a UBI), and a non-linear tax rate on income. In that paper, Mirrlees concludes that *"an approximately linear income-tax schedule, with all the administrative advantages it would bring, is desirable"*. Other papers like Farhi, Werning (2013) and Heathcote, Tsujiyama (2015) find that a linear tax is close the optimal non-linear one. The first thing I do in this paper is to try to replicate the findings of these studies for several different parameters, making sure that the results are not specific to those authors' unique calibrations. I find that the closeness is robust for several parameters for a utilitarian planner.

Next, I show that adding heterogeneity to the preferences can also generate a flatter non-linear marginal tax rate. I follow Lockwood, Weinzierl (2015), where they use a quasi-linear utility function, as in Diamond (1998), to simplify the Mirrlees model that allows for a simple characterization of the solution. With this model, I show that as the agents become more heterogeneous, the marginal tax rate becomes flatter.

Adding heterogeneity to the preferences make the problem harder to solve, since this adds more dimensions to the problem. As Judd et al. (2006) argue, the multidimensional case doesn't have a principle like single-crossing that allows us to use the first order approach. Also, numerical solutions become much harder to calculate since constraints are not concave. Therefore simplifications are necessary to characterize the solution.

Related literature. The literature related to the Mirrlees model is long and a good summary can be found in Mankiw, Weinzeirl, Yagan (2009). Among the papers that try to answer a similar question than this paper

[1] Friedman (1962) proposed something similar to UBI, a Negative Income Tax.

[2] Universal Basic Income can be defined as a program or policy that transfers income (normally in the form of cash) to all citizens or residents of a country or economy, regardless of their income or wealth.

[3] More notably in the last years: automation. But others include the possibility of reducing poverty and inequality, changing the welfare system that creates distortions in the labor market, simplifying the tax system, etc.

[4] The tax system is an affine system $T(y) = b + \tau y$, where b is the lump-sum component that can be positive (a lump-sum tax) or negative (a lump-sum transfer). For simplicity, I'll call this tax system a linear tax system.

there is Alari (2016), that compares universal transfers versus means tested in a partial equilibrium model, and finds that means tested are preferred. Saez (2002) shows that a negative income tax is optimal when behavioral responses are concentrated along the intensive margin. The paper by Damon, Marinescu (2017) analyzes the effects of the transfers of the Alaska Permanent Fund on the labor market. They do not find a negative effect on employment, and they do find an increase in part-time work. Marinescu (2018) analyzes the behavioral effects of cash transfer experiments realized in the past on labor, education, consumption, health and others. In general these experiments showed positive outcomes on the variables analyzed.

This paper is divided as follows. First, I solve numerically the standard Mirrlees model for different parameters and check if the linear system is "close" to the optimal non-linear marginal tax rate. Second, I describe the model with heterogeneity in preferences and show that heterogeneity can generate a flatter non-linear tax rate (closer to a constant marginal tax rate). Then I conclude.

2 Tax simplicity

In this section I'll confirm the conclusion found in Mirrlees (1971), that a linear tax system is close to the optimal non-linear tax system. The model is a standard Mirrlees model, which is described here briefly.

There is a continuum of individuals, each having the same preference: $u(c,l)=U(c)-V(l)$ defined over consumption x and labor l. Workers differ only in their productivity w, with density $f(w)$. A planner cannot observe productivity w, but can observe income $y=wl$. The agent maximizes its utility function $u(c,l)$ subject to the budget constraint:

$$c = y - T(y)$$

where $T(y)$ is the tax paid.

Government maximizes a social welfare function

$$W = \int_0^\infty G(u)f(w)dw$$

subject to the incentive compatibility constraints, and a resource constraint:

$$\int_0^\infty c(w)f(w)dw \leq \int_0^\infty y(w)f(w)dw - E$$

where E is the government expenditure (excluding transfers).

The parameters used are:

Table 1: Parameters of the model

Variable	Values	Source
Utility functions	Type 1: $u = log\left(c - \frac{l^{1+k}}{1+k}\right)$ Type 2: $log(c) - log\left(1 + \frac{l^{1+k}}{1+k}\right)$	Saez (2001)
Comp. elasticity$(1/k)$	$\{0.25, 0.5\}$	Saez (2001)
G (excluding transfers)	$\{0.1, 0.2, 0.3\}$	U.S. (2007) was 20.3%
People with ability $= 0$	5%	Social Security
Distribution $F(w)$	Log-normal up to 42$ per hour. Pareto (2) after that	CPS 2007, and Saez (2001)
Social welfare function	$G(u) = u,$	

To analyze the effects of moving from the full non-linear system to the linear system, I'll use two measures:

1. Welfare loss (%W), where $W(u(c(1-w\%), l); T, E) = W(u(\widetilde{c}, \widetilde{l}); \widetilde{T}, E)$. That is, moving to a linear system is equivalent to losing %W of consumption for every agent.

2. Saving needed (%S), where $W(u(c, l); T, E) = W(u(\widetilde{c}, \widetilde{l}); \widetilde{T}, E(1-\%S))$.

The first is the standard measure of welfare loss from not doing the optimal policy. The second is the one that is more suitable for the goal of this exercise. The second measure finds the savings in the economy (through simplification of the tax system, and welfare programs) that are necessary to make the planner indifferent between the non-linear and the linear tax systems.

Table 2: Numerical simulations of optimal taxes

Government spending (%GDP)	0.1		0.2		0.3	
	Compensated Elasticity		Compensated Elasticity		Compensated Elasticity	
	0.25	0.50	0.25	0.50	0.25	0.50
Type1 Utility						
Welfare loss	0.31%	0.37%	0.29%	0.43%	0.25%	0.42%
Saving needed (% GDP)	0.27%	0.33%	0.23%	0.34%	0.18%	0.34%
Type 2 Utility						
Welfare loss	0.50%	0.82%	0.50%	0.85%	0.51%	0.85%
Saving needed (% GDP)	0.38%	0.56%	0.49%	0.49%	0.52%	0.52%

Table 1 shows how close the two measures are for the linear tax compared to the non-linear one. In terms of welfare, using a linear tax system implies a reduction of equal to or less than 0.85% on average consumption. It should be noted that this comparison is with the optimal non-linear tax system, and not with the actual tax system. In Heathcote, Tsujiyama (2015)for example, they find that the linear tax system improves welfare compare to the actual system, and it achieves 71% of the welfare possible given by the non-linear system. So, these numbers should be interpreted as the distance in welfare between the linear system compared with the best the government can do, and not with the actual system of an economy.

The savings needed to make the planner indifferent is 0.52% of GDP or less, depending on the parameters. These numbers can be compared directly with administrative costs of welfare programs and tax systems. Also, there are other costs related to having a non-linear tax system.Many tax systems change their tax rates in a non-continuous fashion, and welfare programs depending on income of the participant can change drastically by an increase of a small amount of income. These facts create kinks and notches[5] in the choice set creating some behavioral responses by individuals. For example, East (2018) finds a decrease in labor supply by welfare recipients. Saez (2010) finds some accumulation of individuals (bunching) in kinks created by the Earned Income Tax Credit (EITC), but he argues that fiscal evasion can explain that bunching. More recently, with a richer database, Mortenson, Whitten (2016) find evidence of bunching in more kinks and this bunching is becoming more important over time. Also, Ruh, Staubli (2018) find strong evidence of bunching in Austria due to a notch created by the disability insurance program of that country.

In a recent document by the IDB[6] they find that 4.4% of GDP is lost because of inefficient public spending in Latin America, and 1.7% of that occurs in transfer programs. That 1.7% is greater than all the values estimated previously.

3 The role of Heterogeneous preferences on Optimal Income Taxation

In this section I'll show that heterogeneous preferences can be a sufficient condition to have a linear marginal tax rate on income. This result is in line with the results of Lockwood, Weinzierl (2015) and Judd et al. (2006), where they find that heterogeneous preferences reduce redistribution.

[5]A kink is a non-smooth change in the choice set, as we see in changes in the marginal tax rates. A notch is a discontinuous jump in the choice set, a feature that is normally generated by losses of benefits in welfare programs due to small increases in income. A review of kinks and notches can be found in Slemrod (2019)

[6]See Izquierdo, Pessino, Vuletin (2018)

Heterogeneous preferences present a challenge in solving the optimal problem, and even numerical solutions are hard to obtain, as explained in Judd et al. (2006). The special case considered in Lockwood, Weinzierl (2015), with simplifying assumptions, allows for a simple characterization of the solution. They consider the Mirrlees model with a setup that is similar to Diamond (1998). By using a specific form of heterogeneous preferences, they find a solution that is similar to that in Diamond (1998), but with a different form of welfare weights. I'll use this formulation and show that the non-linear optimal marginal tax rate on income becomes flatter as heterogeneity increases. The extreme case, where heterogeneity is increased to its maximum possible amount, shows a linear marginal tax rate.

3.1 A model with homogeneous preferences

In this subsection and the following one I'll follow closely Lockwood, Weinzierl (2015) to describe the model with homogeneous preferences and heterogeneous preferences, and show how they compare to each other.

Individuals have a utility that is linear in consumption c and non-linear in labor effort l given by $u(c,l) = c - l^{1+1/\gamma}$where γ is the constant elasticity of labor supply. They have an unobservable ability $n \geq 0$ so that gross income y is equal to nl. Thus we can rewrite the utility function in the following form:

$$U(c,y,n) = c - (y/n)^{1+1/\gamma} \tag{1}$$

The distribution of ability is given by the cumulative density function $F(n)$with density $f(n)$.

The planner selects the allocation $\{c(n), y(n)\}$to maximize the social welfare function W, solving:

$$W = \max_{\{c(n),y(n)\}} \int_0^\infty g(n)\, U(c(n), y(n), n)\, f(n) dn \tag{2}$$

where $g(n) \geq 0$ is the welfare weight for type n assigned by the planner. If the planner is utilitarian, $g(n) = 1$ for all n.

The maximization problem in 2 is subject to the resource constraint:

$$\int_0^\infty (y(n) - c(n))\, f(n)\, dn \geq E$$

where E is government spending,

The incentive compatibility (IC) constraints are:

$$U\left(c\left(n\right),y\left(n\right),n\right)\geq U\left(c\left(m\right),y\left(m\right),n\right),\ \forall\ ,m,n$$

As done in the literature, we can define the income tax function as $T\left(y\right)=y-c$.

In this setup, as shown in Diamond (1998), the optimal tax function is characterized by the following first-order condition:

$$\frac{T'\left(y\left(n\right)\right)}{1-T'\left(y\left(n\right)\right)}=\frac{1+1/\gamma}{nf\left(n\right)}\left(G\left(n\right)-F\left(n\right)\right),\ \forall n \tag{3}$$

where $G\left(n\right)=\frac{\int_0^n g(m)f(m)dm}{\int_0^\infty g(m)f(m)dm}$, normalized so that $G\left(0\right)=0$ and $\lim_{n\to\infty}G\left(n\right)=1$.

3.2 A model with heterogeneous preferences

A simple model that can add heterogeneity and keep the simplicity of the solution is developed in Lockwood, Weinzierl (2015), which I describe shortly below.

The model is a modification of the previous subsection, where now an individual is defined by a two-dimensional type (w,θ), where $w\geq 0$ is now the unobservable ability so $y=wl$, and $\theta>0$ is an unobservable preference parameter. This parameter is assumed to have a population average equal to one, and can be thought of as a taste parameter scaling the disutility of labor relatively to consumption. The utility function of the agent is now $u\left(c,l\right)=c-\left(l/\theta\right)^{1+1/\gamma}$. As in the previous section, the utility can be rewritten as:

$$U\left(c,y,w,\theta\right)=c-\left(\frac{y}{w\theta}\right)^{1+1/\gamma} \tag{4}$$

Here it is important to note that in (4) the two unknown parameters to the planner, w and θ, are entered in the utility function specifically to help solve the problem. If two individuals (w',θ'),(w'',θ'') are such that $w'\theta'=w''\theta''$, they will behave as if they are the same individual.

Thus, the product $w\theta$ is a sufficient statistic for this problem and the planner can consider individuals as a function of the product $w\theta$. The planner then will choose allocations $\{c(w\theta), y(w\theta)\}$.

As in the previous section, it is assumed that the planner seeks to maximize the welfare function W using welfare weights $b(w,\theta)$, depending now on both parameters. Let $H(w,\theta)$ denote the joint probability distribution of ability and preferences, with density $h(w,\theta)$. The planner's problem is:

$$W = \max_{\{c(w\theta),y(w\theta)\}} \int_0^\infty \int_0^\infty b(w,\theta) U(c(w\theta), y(w\theta), w, \theta) h(w,\theta) dwd\theta \tag{5}$$

The resource constraint in this case is:

$$\int_0^\infty \int_0^\infty (y(w\theta) - c(w\theta)) h(w,\theta) dwd\theta \geq E$$

and the IC constraints.

I'll use the same assumption used in Lockwood, Weinzierl (2015).

Assumption: $b(w,\theta) = b(w,\theta')$for all θ and θ', or $b(w) \equiv b(w,\theta)$for all θ.

Under this assumption, (5)becomes:

$$W = \max_{\{c(w\theta),y(w\theta)\}} \int_0^\infty \int_0^\infty b(w) U(c(w\theta), y(w\theta), w, \theta) h(w,\theta) dwd\theta \tag{6}$$

Now, to transform this model to an equivalent version of the homogeneous preferences, it is useful to make a change in variables denoting the unified type by n, so that $n = w\theta$.

The change of variables implies some changes in notation. Let $\hat{H}(\theta,n)$ now be the joint distribution of preferences and unified type, with density $\hat{h}(\theta,n) = h(n/\theta,n)$. Also, let $f(n) = \int_0^\infty \hat{h}(\theta,n) d\theta$. Then substituting the variables, the resource constraint and the IC constraint in both models are the same.

Further, the planner problem can be written as:

$$W = \max_{\{c(n),y(n)\}} \int_0^\infty \int_0^\infty b(n/\theta) U(c(n), y(n), n) \hat{h}(\theta,n) d\theta dn =$$

$$\max_{\{c(n),y(n)\}} \int_0^\infty \left(\int_0^\infty \frac{b(n/\theta)\,\hat{h}(\theta,n)\,d\theta}{f(n)} \right) U(c(n), y(n), n)\, f(n)\, dn$$

or

$$\max_{\{c(n),y(n)\}} \int_0^\infty \hat{b}(n)\, U(c(n), y(n), n)\, f(n)\, dn \tag{7}$$

where $\hat{b}(n) = \left(\int_0^\infty \frac{b(n/\theta)\hat{h}(\theta,n)d\theta}{f(n)} \right)$.

Note that (7)is equivalent to (2), with $\hat{b}(n)$ instead of $g(n)$.

Now, the optimal income tax will be given by equation (3), with

$$G(n) = \frac{\int_0^n \hat{b}(m)\, f(m)\, dm}{\int_0^\infty \hat{b}(m)\, f(m)\, dm} \tag{8}$$

So, the introduction of heterogeneity on preferences in this model creates another model that is equivalent to the original one, but with a modified welfare weight that will depend on the distribution of the preference parameter.

3.3 The role of heterogeneity on marginal income tax rates

I now proceed to show an analytical and a numerical result, showing the role of heterogeneity on the marginal income tax rate. The result applies for Pareto distributions, an important distribution for this literature since it is use to estimate the tail of the ability distribution as in Saez (2001).

I'll consider the simplified version of the model where the ability distribution is independent of the preference parameter. Assume the ability is distributed Pareto with parameter α, that is, w has support $[w_0, +\infty]$ and the pdf is $\frac{\alpha w_0^\alpha}{w^{\alpha+1}}$. The preference parameter$\theta$is distributed uniformly between $[\theta_L, \theta_H]$. Specifically, I'll consider the case where $\theta_L = 1 - \epsilon$ and $\theta_H = 1 + \epsilon$ for $\epsilon < 1$.

Let:

$$b(w) = \kappa w^{-\eta} \tag{9}$$

be the welfare weight assigned by the planner to the individual with ability w, for positive values of $\kappa$and η.

In the following proposition I state the main result, that the marginal tax rate goes to a constant for all n above some level of wages w.

Proposition: *Consider the income tax function (3) that solves the planner problem in (7), where $G(n)$ is given by (8). Further, let the welfare weight be given by (9) and the unified type distribution be given by the joint distribution of w and θ,where w is distributed Pareto and $\theta \sim U[1-\epsilon, 1+\epsilon]$ is uniformly distributed.*

Then, as $\epsilon \to 1$, $T' \to C$ for all $n \geq w_0\theta_H$, where C is a constant.

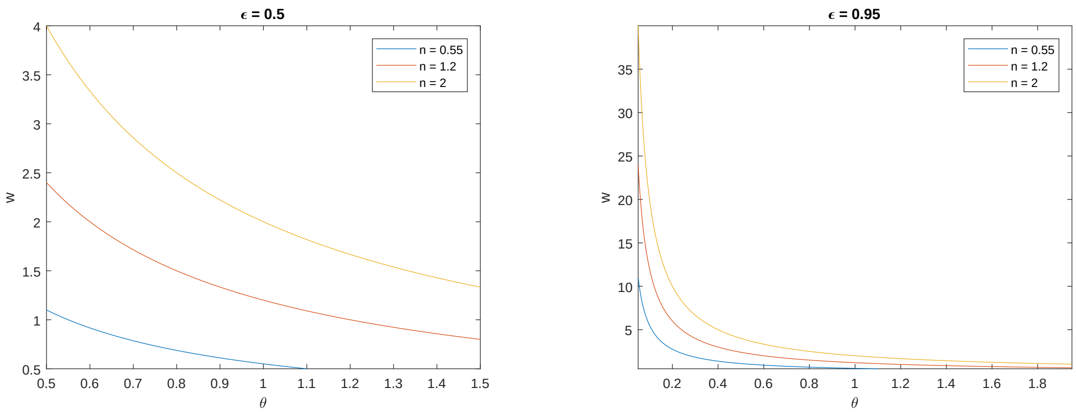
Proof: see appendix.

To understand this result we must notice that, as ϵ increases, every unified type n includes more and more w types (in the sense that for every n type, several w types are part of the unified n type). If $\epsilon = 0$, then $n = w$. But, if $\epsilon > 0$, for every n all the $w \; \epsilon \; [n/(1+\epsilon), n/(1-\epsilon)]$. Then, the welfare weight $G(n)$ includes several wages w and their corresponding values $b(w) = \kappa w^{-\eta}$. In fact, as $\epsilon \to 1$, for every $n \geq w_0\theta_H$ the unified type n includes all the w types from $n/\theta_H$to ∞.

The economic intuition of this result rests on the fact that the parameter θgenerates noise to the planner, where he is not able to distinguish what kind of $w-$type the agent is. The more variation (the bigger ϵ), the more noise there is. The only exception to the constant marginal tax rate is for those types $n \; \epsilon \; [w_0\theta_L, w_0\theta_H]$.Those types are different because there is only a subset of θvalues that can be combined with w to get n. For example, the unified type $n = w_0\theta_L$has a unique value for θ that generates that value of n (which is θ_L). It can be seen in the figure below that for $n = 0.55$ there is no w^* that makes $n = w * \theta_H$.

The figures below show the w-types that every n-type implies for different values of θ. If ϵ is small, fewer w-types are included in each n-type. If ϵ is close to 1, more w-types are included making the modified welfare weight $\hat{b}(n)$ more alike for different n.

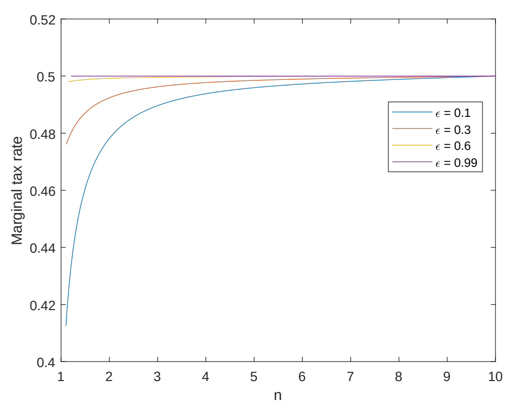
Figure 1: w types included in each n type for different values of e



The exercise is done for $\theta \ \epsilon \ [1-\epsilon, 1+\epsilon]$and $w \ \epsilon \ [0.5, \infty]$

To see how the marginal tax rate changes with ϵ. The next graph shows the evolution of the marginal tax rate for different values of ϵ.

Figure 2: Marginal tax rate for different values of ϵ



Values used in this graph: $w0 = 1$, $\alpha = 2$, $\kappa = 1.5$, $\gamma = 1$.

4 Conclusion

In this paper I have calculated the closeness of a linear tax system to the optimal non-linear tax system. I find that a linear tax system is close to the optimal non-linear tax system for a wide range of parameters. As others have noted, the linear system is better, in terms of welfare for the planner, than the actual tax system in the US. So, this closeness seems to hold not only for some specifications of the parameters model, but for a wide range of them. An obvious conclusion is that countries where administrative costs are high, and distortions created by welfare programs are important, have more to gain with a linear tax system.

Further, the characterization of the optimal non-linear tax system requires a lot of information that might not be available to the planner. In the second part of the paper I introduce heterogeneity into the preferences of the individuals, and I find that heterogeneity makes the non-linear tax system flatter. A key result of this paper is that heterogeneity could make the tax system completely linear.

Given these results, the answer to the question posed at the beginning is that it is not radical to give cash directly to people. A lump-sum transfer, that can be interpreted as UBI, is present on the Mirrlees model. Further, the non-linear tax system is already close to the linear tax system. If we consider that individuals have heterogeneous preferences, the non-linear tax system gets even closer to the linear tax system.

This paper should not be interpreted as a proof of the convenience of a policy like UBI; however, the paper does indicate that UBI should be considered seriously.

Further research should include dynamic aspects, as well as other uncertainties that the planner mayface, such as not knowing exactly the distribution of abilities.

Appendix.

Proof of proposition.

Given the distribution for w and θ the distribution of the unified-type n is given by:

$$f(n) = \begin{cases} \frac{\alpha w_0^\alpha}{n^{\alpha+1}} \frac{1}{\theta_H - \theta_L} \left[\frac{(n/w_0)^{\alpha+2} - \theta_L^{\alpha+2}}{\alpha+2} \right] \text{ if } n/w_0 < \theta_H \\ \frac{\alpha w_0^\alpha}{n^{\alpha+1}} \frac{1}{\theta_H - \theta_L} \left[\frac{\theta_H^{\alpha+2} - \theta_L^{\alpha+2}}{\alpha+2} \right] \text{ if } n/w_0 \geq \theta_H \end{cases}$$

The modified welfare weight $\hat{b}(n) = \left(\int_0^\infty \frac{b(n/\theta)\hat{h}(\theta,n)d\theta}{f(n)} \right)$ is then:

$$\hat{b}(n) = \begin{cases} \frac{\kappa(\alpha+2)}{n^\eta(\eta+1)} \left[\frac{\theta_H^{\eta+1} - \theta_L^{\eta+1}}{(n/w_0)^{\alpha+2} - \theta_L^{\alpha+2}} \right] \text{ if } n/w_0 < \theta_H \\ \frac{\kappa(\alpha+2)}{n^\eta(\eta+1)} \left[\frac{\theta_H^{\eta+1} - \theta_L^{\eta+1}}{\theta_H^{\alpha+2} - \theta_L^{\alpha+2}} \right] \text{ if } n/w_0 \geq \theta_H \end{cases}$$

Then $\hat{b}(n) f(n) = \frac{\kappa \alpha w_0^\alpha}{n^{\alpha+\eta+1}(\eta+1)} \left[\theta_H^{\eta+1} - \theta_L^{\eta+1} \right]$.

From (8)we have:

$$G(n) = \frac{\int_0^n \hat{b}(m) f(m)\, dm}{\int_0^\infty \hat{b}(m) f(m)\, dm} = \frac{\int_0^n \frac{\kappa \alpha w_0^\alpha}{m^{\alpha+\eta+1}(\eta+1)} \left[\theta_H^{\eta+1} - \theta_L^{\eta+1} \right] dm}{\int_0^\infty \frac{\kappa \alpha w_0^\alpha}{m^{\alpha+\eta+1}(\eta+1)} \left[\theta_H^{\eta+1} - \theta_L^{\eta+1} \right] dm}$$

$$= \frac{\frac{\kappa \alpha w_0^\alpha}{(\eta+1)} \left[\theta_H^{\eta+1} - \theta_L^{\eta+1} \right] \int_{w_0\theta_L}^n \frac{1}{m^{\alpha+\eta+1}} dm}{\frac{\kappa \alpha w_0^\alpha}{(\eta+1)} \left[\theta_H^{\eta+1} - \theta_L^{\eta+1} \right] \int_{w_0\theta_L}^{+\infty} \frac{1}{m^{\alpha+\eta+1}} dm} = \frac{\int_{w_0\theta_L}^n \frac{1}{m^{\alpha+\eta+1}} dm}{\int_{w_0\theta_L}^{+\infty} \frac{1}{m^{\alpha+\eta+1}} dm}$$

$$= \frac{\frac{1}{w_0\theta_L} - \frac{1}{n^{\alpha+\eta}}}{\frac{1}{w_0\theta_L}} = \frac{n^{\alpha+\eta} - w_0\theta_L}{n^{\alpha+\eta}}$$

From (3) we have that:

$$\frac{T'(y(n))}{1 - T'(y(n))} = \frac{1 + 1/\gamma}{n f(n)} (G(n) - F(n))$$

$$= \frac{1+1/\gamma}{n\frac{\alpha w_0^\alpha}{n^{\alpha+1}}\frac{1}{\theta_H-\theta_L}\left[\frac{\theta_H^{\alpha+2}-\theta_L^{\alpha+2}}{\alpha+2}\right]}\left(\frac{n^{\alpha+\eta}-w_0\theta_L}{n^{\alpha+\eta}}-\left(1-\frac{\theta_H^{\alpha+2}-\theta_L^{\alpha+2}}{(\alpha+2)(\theta_H-\theta_L)}\left(\frac{w_o}{n}\right)^\alpha\right)\right)$$

$$= \frac{(1+1/\gamma)(\theta_H-\theta_L)(\alpha+2)}{\alpha\left[\theta_H^{\alpha+2}-\theta_L^{\alpha+2}\right]}\left(\frac{n}{w_0}\right)^\alpha\left(\frac{n^{\alpha+\eta}-w_0\theta_L}{n^{\alpha+\eta}}-1+\frac{\theta_H^{\alpha+2}-\theta_L^{\alpha+2}}{(\alpha+2)(\theta_H-\theta_L)}\left(\frac{w_o}{n}\right)^\alpha\right)$$

$$= \frac{(1+1/\gamma)(\theta_H-\theta_L)(\alpha+2)n^\alpha}{\alpha w_0^\alpha\left[\theta_H^{\alpha+2}-\theta_L^{\alpha+2}\right]}\left(\frac{n^{\alpha+\eta}-w_0\theta_L}{n^{\alpha+\eta}}-1\right)+\frac{(1+1/\gamma)}{\alpha}$$

The first term of the equation goes to zero as $\epsilon \to 1$, since the parentheses term goes to zero (while the term not in parenthesis goes to a constant).

So:

$$\frac{T'(y(n))}{1-T'(y(n))} \to \frac{(1+1/\gamma)}{\alpha},\ \forall n \geq w_0\theta_H$$

References

Better spending for better lives: how Latin America and the Caribbean can do more with less / edited by Alejandro Izquierdo, Carola Pessino and Guillermo Vuletin.

Diamond, Peter A. "Optimal Income Taxation: An Example with a U-Shaped Pattern of Optimal Marginal Tax Rates." The American Economic Review 88, no. 1 (1998): 83-95.

East, Chloe N., 2018. "Immigrants' labor supply response to Food Stamp access," Labour Economics, Elsevier, vol. 51(C), pages 202-226.

Farhi, Emmanuel, and Ivan Werning. "Insurance and Taxation over the Life Cycle." The Review of Economic Studies 80, no. 2 (283) (2013): 596-635.

Friedman, Milton, and Rose D. Friedman. 1962. Capitalism and freedom.

Heathcote, Jonathan and Tsujiyama, Hitoshi, Optimal Income Taxation: Mirrlees Meets Ramsey (February 2015). CEPR Discussion Paper No. DP10380.

Jones, Damon and Ioana Marinescu. 2017. "The Labor Market Impacts of Universal and Permanent Cash Transfers: Evidence from the Alaska Permanent Fund." Working paper.

Judd, Kenneth L., Ding Ma, Michael A. Saunders and Che-Lin Su. "Optimal Income Taxation with Multidimensional Taxpayer Types." (2006).

Lockwood, Benjamin B. & Weinzierl, Matthew, 2015. "De Gustibus non est Taxandum: Heterogeneity in preferences and optimal redistribution," Journal of Public Economics, Elsevier, vol. 124(C), pages 74-80.

Mankiw, N. Gregory, Matthew C. Weinzierl, and Danny Yagan. "Optimal Taxation in Theory and Practice." Journal of Economic Perspectives 23, no. 4 (Fall 2009): 147–174.

Marinescu, Ioana. (2018). "No Strings Attached: The Behavioral Effects of U.S. Unconditional Cash Transfer Programs," NBER Working Papers 24337, National Bureau of Economic Research, Inc.

Mirrlees, J. A. "An Exploration in the Theory of Optimum Income Taxation." The Review of Economic Studies 38, no. 2 (1971): 175-208.

Mortenson, Jacob A. and Andrew E. Whitten. "How Sensitive Are Taxpayers to Marginal Tax Rates ? Evidence from Income Bunching in the United States." (2016).

Paulus, Alari, (2016), The antipoverty performance of universal and means-tested benefits with costly take-up, No 16/12, ImPRovE Working Papers, Herman Deleeck Centre for Social Policy, University of Antwerp

Ruh, Philippe & Staubli, Stefan, 2018. "Financial Incentives and Earnings of Disability Insurance Recipients: Evidence from a Notch Design," CEPR Discussion Papers 12979, C.E.P.R. Discussion Papers.

Saez, Emmanuel. "Do Taxpayers Bunch at Kink Points?" American Economic Journal: Economic Policy 2, no. 3 (2010): 180-212.

Saez, Emmanuel. (2001). Using Elasticities to Derive Optimal Income Tax Rates. Review of Economic Studies. 68. 205-29.

Saez, Emmanuel. (2002). “Optimal Income Transfer Programs: Intensive Versus Extensive Labor Supply Responses”. The Quarterly Journal of Economics. 117. 1039-1073.

Slemrod, Joel. (2013). Buenas notches: Lines and notches in tax system design. eJournal of Tax Research. 11. 259-283.