Ensaio Econômico

EPGE

Escola

Brasileira de

Economia e

Finanças

Nº 819

ISSN 0104-8910

Progressive Consumption Taxes

Carlos E. da Costa, Marcelo R. Santos

Setembro de 2020

Progressive Consumption Taxes/
Carlos E. da Costa, Marcelo R. Santos - Rio de Janeiro:
FGV EPGE, 2020

44p. - (Ensaaios Econômicos; 819)

Inclui bibliografia.

CDD-330

Progressive Consumption Taxes

Carlos E. da Costa

FGV EPGE

carlos.eugenio@fgv.br

Marcelo R. Santos

INSPER

marcelors2@insper.edu.br

September 29, 2020

Abstract

In a static setting, whether consumption or labor income is progressively taxed is irrelevant for household choices and welfare. In a dynamic setting, however, these two forms of progressivity have markedly different implications for how earnings vary along the life-cycle: in a stylized life-cycle model, progressive income tax act reducing Frisch elasticities of labor supply whereas progressive consumption taxes act reducing the elasticity of intertemporal substitution. After showing that the latter leads to less inefficiencies in the stylized model than the former, we explore the consequences of replacing the current U.S. tax system by one in which labor income are linear and consumption taxes are progressive. We find welfare gains that exceed 10% in consumption equivalent variation terms for all possible specifications in steady state comparisons. Welfare gains are attained in all our specifications with either very small increases in the capital stock or even large declines. *Keywords:* Progressive consumption taxes; warm-glow motives. *J.E.L. codes:* E6; H3; J2.

FUNDAMENTAL tax reform is once again a hot topic in policy debate due to the pervasive inefficiencies and distortions that have been accumulating in the U.S. tax system. Among the most frequent advanced ideas is the move to a system more centered on consumption taxes. While we could refer to [Kaldor \(1956\)](#), as an early reference to the idea of moving to a system based on consumption taxes the appeal of such taxes dates as far back as Thomas Moore, which in his masterpiece, *Leviathan*, states, "It is fairer to tax people on what they extract from the economy, as roughly measured by their consumption, than to tax them on what they produce for the economy as roughly measured by their income."

Despite these views, the idea of moving to a system mostly based on consumption taxes is usually held back by its alleged regressivity. The goal of this paper is to study the potential

da Costa gratefully acknowledges financial support from CNPq Proc. 301140/2017-0. Santos gratefully acknowledges financial support from CNPq Proc. 311437/2014-1. We thank Pedro Cavalcanti, Felipe Iachan, Czar Santos and participants at the EPGE FGV Macro Group. All errors are our sole responsibility.

benefits from moving to a consumption based tax system that addresses its main shortcoming; we study a reform that replaces progressive income taxes by progressive consumption taxes.

In a static world, labor income and consumption taxes are equivalent. Simplicity, administrative ease, and many of the important issues related to real world tax systems discussed by [Slemrod and Gillitzer \(2014\)](#), should ultimately determine which system is best. Although we acknowledge that these considerations may justify or undermine a reform, we shall not have much to add to the subject. Instead, we focus on the different implications of these two types of progressivity in a dynamic setting. As we shall try to convey the potential benefits of progressive consumption taxes are so large as to justify a careful evaluation of its practical implementation.

Both [Kaldor \(1956\)](#) and [Hall and Rabushka \(2007\)](#) take income taxes to mean taxes on both labor and capital income, with a consequent focus on the reform's impact on savings. We, instead, distinguish labor and capital income taxes and assume that progressivity only applies to the former in the benchmark. That is not to say that the consequences of our reform are not important as we shall see. Another important aspect of our analysis is that we do not take into account the confiscatory aspect of a non-anticipated move to consumption taxes emphasized by [Correia \(2010\)](#), for example; our analysis is centered in steady state comparisons.

So, what are the forces that we explore here? First, note that the novelty of our work is to introduce progressivity in consumption taxes. Again, in a static world, whether we introduce progressivity in income taxes – concave earnings retention function – or in consumption – concave mapping from expenditures to consumption – is immaterial. The effects highlighted in this paper only arise when the intertemporal nature of household choices are taken into account. If labor income taxes are progressive, more taxes are paid when earnings are higher, whereas, when it is consumption taxes which are progressive, more taxes are paid when consumption is higher. This has important allocative consequences.

Indeed, consider the progressive tax schedules studied by [Heathcote et al. \(2017b\)](#) – henceforth HSV schedules, $T(z) = z - \xi z^{1-\tau}$, where z is labor income. In a very stylized life-cycle model – Section 1, we show that the main consequence of such taxes for household behavior is to reduce the Frisch elasticity of labor supply, making agents less responsive to the evolution of productivity that takes through one's life. The allocation of time or effort across ages is therefore distorted. When this system is replaced with one in which progressivity is on consumption taxes efficient intertemporal allocation of effort is restored. The downside is that we increase the intertemporal elasticity of substitution, thus increasing the desire for consumption smoothing.

The discussion above is based on a deterministic world with perfect capital markets.

The main goal of this paper is to provide a quantitative assessment of such reform, under more realistic conditions. We calibrate an overlapping generations economy where agents live a meaningful life-cycle facing shocks to their labor market productivity while self insuring through imperfect capital markets. While the progressivity of the current tax system provides insurance against earnings fluctuation, a progressive consumption schedule provides insurance against consumption fluctuation.

We find very large welfare gains.¹ Although large gains are due to the effects we have anticipated in our discussion, a very large impact of consumption tax progressivity on savings cannot be a priori discounted. To understand the logic, it is important to mention that agents in our model save not only for smoothing purposes but also because they derive direct utility from wealth accumulation through a warm glow motive. Such feature is introduced to account for the empirical wealth distribution. Indeed, non-homothetic utility for wealth has been incorporated in many macroeconomic models to account for the convexity of savings with respect to permanent income, an empirical regularity recently competently explored by [Straub \(2019\)](#), and its consequence for wealth accumulation.² As it turns, a progressive consumption tax leads the very rich to save even more. Because wealth is not 'eaten away' by the rich, they sacrifice more consumption which is not only turned into more capital accumulation but also left for the poor, thus strengthening the distributive (in terms of consumption Gini) impact of such reform. Yet, what we find in our baseline reform is that the capital stock increases by only 2.2% which can hardly explain the sizable gains we find.³

Still, since there is no consensus regarding the elasticity of bequest motives with respect to consumption taxes we run various robustness exercises.⁴ The exact value of welfare gains varies substantially across specifications but they remain very high in all exercises. In particular, when we assume that donors derive utility not by the bequest they leave but by the amount of consumptions that the donee can guarantee him or herself with the inherited bequest, welfare gains drop to 10.95% despite a 8.45% decline in the capital stock.

All variations considered welfare gains vary – from a minimum gain of 8.62% to a max-

¹In our baseline example, the unconstrained optimum generates a welfare gain of 13.34%. To put in perspective, [Conesa et al.'s \(2009\)](#) highly cited work finds a 1.33% gain. [Kapička's \(2020\)](#) introduction of history dependence on income taxes attains a maximum gain of 2.98%. [Conesa et al. \(2020a\)](#), 5.47% with differentiated commodity taxes, and [Conesa et al. \(2020a\)](#), 5.96% with differentiated taxes and a very generous Universal Basic Income.

²[Straub \(2019\)](#) shows how this convexity can partially account for the recent decline in real interest rates, the rising private wealth to GDP ratio and the fast increase in wealth inequality in the U.S.

³Again, it is useful to compare with [Conesa et al. \(2020a\)](#) and [Conesa et al. \(2020b\)](#) where large gains are also obtained. In [Conesa et al. \(2020b\)](#), the 5.96 increase in welfare is mostly explained by a 29.8% capital stock increase. These large increases in capital stock lead welfare gains to be more than dissipated in the transition.

⁴Unfortunately, most of the literature has emphasized the bequest responses to estate taxes, and we know of no study that directly explores responses to consumption taxes. As [Kopczuk \(2013\)](#) emphasizes, modeling details matter, and absent more evidence one has to focus on results which are robust across specifications.

imum gain of 14.49%. The lowest value arises when we restrict policies to those that do not increase the share of revenues raised by consumption taxes. This is a strong arbitrary restriction on tax instruments which induces an 18% decline in capital stock, which suggest large 'extra' welfare gains in a transition for the new steady state. If we allow the U.S. to become more like the rest of the OECD with regards to the relevance of consumption taxes, then steady state welfare gains return to two digits.

Following a brief literature review, in Section 1 we use a stylized model to provide the heuristics of our main results. Section 2 is where we write down and calibrate to the U.S. economy a model where agents live a meaningful life-cycle, are exposed to earnings risk, and face imperfect capital markets. In Section 3 we display our main results, which have their robustness assessed in Section 4. Section 5 compares our findings with those of [Conesa et al. \(2020a,b\)](#) where progressivity is attained by differentiated taxation of commodities. Section 6 is of a very speculative nature. There we digress about the difficulties associated with the practical implementation of progressive consumption taxes. Finally, Section 7 concludes.

Brief Literature Review

This paper is about fundamental tax reform. Famously [Vickrey's \(1947\)](#) "Agenda for Progressive Taxation" proposed replacing the yearly basis for income taxation by a tax system based on the notion of income averaging. More recently, the idea of transitioning to a consumption based tax system has found some support in a great deal due to [Hall and Rabushka's \(2007\)](#) flat tax and [Bradford's \(2005\)](#) X-tax proposals. We follow their lead in proposing a large overhaul in the U.S. tax system, but focus on the quantitative assessment of such reforms.

The focus on consumption taxes has its roots on actual proposals as the aforementioned flat tax and X-tax ideas. Of course, advocacy of consumption taxes dates back as far as Thomas Hobbes', while its potential regressivity was already recognized by John Stewart Mill ([cite](#)).⁵ We consider a model in which this regressivity is present. We allow agents to derive direct utility from accumulating wealth through a warm-glow bequest motive, first used by [Andreoni \(1989, 1990\)](#). By introducing non-homotheticity in these preferences we make consumption concave in permanent income, as shown by [Straub \(2019\)](#), and approximate the distribution of wealth found in the data, e.g., [De Nardi \(2004\)](#).

Non-homotheticity of warm-glow motives causes life-time-richer agents to consume a lower fraction of their life-time income than poorer agents, hence producing the type of regressivity that is referred to by the literature. What we do in this paper is to allow policy to impose progressivity directly on consumption taxes. This again has ancient roots. After

⁵Mill

pointing out the potential regressivity of consumption taxes, John Stewart Mill suggested the exemption of necessities as a way of introducing progressivity in consumption taxes. The idea of consumption tax progressivity was not abandoned, and many proposals have been advanced in the more than 150 years that has elapsed since Mill's work. In most cases, progressivity was attained in most proposals by a combination of income taxes (taxes on $z_t + r_{t-1}a_{t-1}$) with investment expenses (subsidies on $a_t - a_{t-1}$). Our work is agnostic about how progressivity is attained and we simply assume that it is equally feasible to introduce progressivity on income and on consumption taxes.

In a provocative post dating back a few years, John Cochrane proposes replacing the U.S. current income tax schedule by a progressive Value Added Tax. His main motivation is that by doing so many of the features that are added to the income tax system can be avoided and the tax code simplified. In his words, "An income tax made sense in 1914, when it was a small tax aimed only at high incomes, and when incomes were much easier to measure than consumption. But that is no longer the case." He continues, "We have a sales tax reporting mechanism, so adding or substituting VAT tax reporting is not that hard."⁶ Our work is silent about whether this view has merit. Indeed our approach does not capture the practical implementation issues that arise when one takes the tax system perspective, in the sense of [Slemrod and Gillitzer \(2014\)](#), and motivate Cochrane's post. For our purposes, any combination of taxes that lead to the budget constraints we evaluate here is equally valid.⁷

The main source of inefficiency that the schedules we study try to avoid are related to the allocation of effort across ages and states of nature. Along these lines, our work is reminiscent of [Vickrey's \(1947\)](#) income averaging approach. In a very stylized model we show that his proposal dominates ours since ours trades-off intertemporal distortions in effort and in consumption, whereas [Vickrey's](#) proposal is based on the lifetime value of both consumption and income, thus avoiding all intertemporal distortions. [Kapička \(2020\)](#), which is a modern take on these types of fundamental tax reforms, shows that, when compared to a history dependent tax, [Vickrey's](#) proposal remain optimal if productivity shocks are only transitory.

Since we add to both [Vickrey's](#) and [Kapička's](#) life-cycle models a warm glow bequest motive progressive consumption taxes may have a very important role in incentivizing savings when such a motive for wealth accumulation is present. We use different elasticities of bequests with respect to consumption taxes and consider many restrictions on the allowable schedules to assess the robustness of our findings. For all optimal schedules and

⁶<https://johnhcochrane.blogspot.com/2017/04/a-progressive-vat.html>.

⁷Although the tax system which mimics consumption taxes by adding investment expenses to the income tax schedule is not a perfect substitute for the one analyzed here, one can think of adding features that make them so. E.g., by allowing income taxes to depend on investments and investment subsidies to depend on income.

all specification of preferences we find very large welfare gains.

The use of warm-glow preferences raises some important issues regarding the welfare criterion to be used — e.g., [Diamond \(2006\)](#); [Kopczuk \(2013\)](#). We take the conservative view of ignoring the flow utility generated by warm-glow.

By moving from income to consumption taxes a concern regarding policy evaluation is the role of wealth confiscation as emphasized by [Correia \(2010\)](#). We avoid this issue by focusing on steady-state comparisons.

Closest to our work is [Conesa et al. \(2020a,b\)](#). Progressivity in consumption taxes is attained in their model through the differentiated taxation of luxuries and necessities. The problem with such way of generating progressivity is that it is highly inefficient. Indeed, they consider preferences which are separable between consumption goods and leisure. A direct application of [Corlett and Hague's \(1953\)](#) classic result prescribe taxing more heavily the necessity. By subsidizing the good that should be taxed by purely efficiency reasons, one incurs in huge inefficiencies to attain redistribution. Again, our proposal avoids this drawback by directly imposing progressivity on all consumption.

Although we do not evaluate the transition to the new steady state, the welfare gains we find arise with either very little increase in the capital stock or even large declines. This suggests that the transition costs that undermine all reforms in [Conesa et al. \(2020a\)](#) and [Conesa et al. \(2020b\)](#) should not occur here.

1 Heuristics

The purpose of this section is to clarify the conceptual difference between consumption tax progressivity and labor income tax progressivity. We organize our discussion around constant progressivity schedules — $T(z) = z - \xi z^{1-\varrho}$, where z are earnings.⁸ Parameter ϱ controls progressivity and ξ is used to adjust the level of taxes. We shall refer to these schedules as HSV schedules in reference to [Heathcote et al. \(2017b\)](#), who has advocated its use as a good approximation of the U.S. system.

Let $\hat{c}$ denote the agent's expenditure and c , his consumption. In a static world, we have $c = \hat{c} \leq \xi z^{1-\varrho}$. Such budget constraint is, of course, equivalent to $c \leq \xi \hat{c}^{1-\varrho} = z$, in which labor income is no longer taxed while consumption expenditures are now progressively taxed. Since the budget sets are identical, the two systems are equivalent.

In a multi-period model, however, we can write

$$\sum_t \frac{\hat{c}_t - y_t}{(1+r)^t} \leq 0,$$

⁸[Musgrave \(1959\)](#); [Feldstein \(1969\)](#) which represent early uses of this parametrization for income tax schedules, emphasize their constant progressivity.

where $\hat{c}_t$ denotes period t expenditures and y_t , period t disposable income.

Typically, it is income which is progressively taxed. In which case, $\hat{c}_t = c_t$ and $y_t = \xi z_t^{1-\varrho}$. Note that we could equivalently have allowed for linear consumption taxes, by making $\hat{c}_t = c_t(1 + \tau^c)$ and defining $\hat{\xi} = \xi(1 + \tau^c)$. That is, whether the level parameter applies to consumption or income is irrelevant for the agent's program.⁹

But progressivity does matter. Indeed, consider the alternative tax system for which $c_t = \xi \hat{c}_t^{1-\varrho}$ and $y_t = z_t$. Now, more taxes are paid in periods in which the agent's consumption is higher, instead of periods in which earnings are higher. If we specialize our discussion to preferences of the form

$$u(c) = \frac{c^{1-\sigma}}{1-\sigma} \quad h(n) = \frac{n^{1+\gamma}}{1+\gamma},$$

then it is not hard to see that a *progressive income tax schedule reduces the Frisch elasticity of labor supply* from $1/\gamma$ to $(1 - \varrho)/(\gamma + \varrho)$, making labor supply more stable along the life-cycle. As for *progressive consumption taxes reduce the intertemporal elasticity of substitution* from $1/\sigma$ to $(\sigma(1 - \varrho) + 1)/(1 - \varrho)$, while leaving the Frisch elasticity unaltered.

The allocative consequences of the two types of progressivity are different and we shall now offer a very stylized framework to compare the two. In our quantitative assessments the forces isolated in this section are shown to be very important in a more realistic setting – see Figure 2.

We start by recalling that the dynamic problem has a static representation in which the agent derives utility from the present value of his or her consumption and, for each present value of income must endure some utility cost.

To simplify the notation, consider the no tax program under the assumption that $\beta(1 + r) = 1$,

$$\max \sum_{t=1}^T \beta^{t-1} [u(c_t) - v(n_t)] \quad \text{s.t.} \quad \sum_{t=1}^T \beta^{t-1} [c_t - n_t w_t] \leq 0.$$

It can equivalently be written as

$$V(C, Z, W) = A^{-1} \max_{C \leq Z} \left[\frac{C^{1-\sigma}}{1-\sigma} - \frac{A}{1+\gamma} \left(\frac{Z}{W} \right)^{1+\gamma} \right],$$

for

$$C = \sum_{t=1}^T \beta^t c_t, \quad Z = \sum_{t=1}^T \beta^t n_t w_t, \quad W = \left[\sum_{t=1}^T \beta^t w_t^{\frac{1+\gamma}{\gamma}} \right]^{\frac{\gamma}{1+\gamma}}, \quad A = \left[\frac{1 - \beta^T}{1 - \beta} \right]^{-\sigma}.$$

⁹In the calibrated model, this is no longer the case since we also assume a warm glow bequest motive.

The constant A accounts for the intertemporal aggregation: the 'technologies' for i) converting, for a given life-cycle productivity profile, $\{n_t\}_t$, a flow of effort costs $(1+\gamma)^{-1} \{n_t^{1+\gamma}\}$ into a present value of earnings, Z , and; ii) converting a present value of earnings into a flow of consumption utilities, $(1-\sigma)^{-1} \{c^{1-\sigma}\}_t$. This intertemporal aggregation in itself may lead to more or less effort by this 'static' representation of an agent making dynamic choices when compared to the same agent making true static choices, e.g., were he or she not able to access asset markets. When w_t is constant, it can be shown that whether more or less effort is generated only depends on whether $\sigma \lesseqgtr 1$. If $u(c) = \ln c$ intertemporal aggregation is irrelevant; the static representation of the dynamic problem leads to the same choices as the static problem. Note that there is nothing economically meaningful in this equivalence, it is just an adjustment which is necessary for a representation choice which is convenient for our discussion.

We can construct similar problems for the cases in which the budget constraints are

$$\sum_t \beta^t [c_t - \xi(n_t w_t)^{1-\varphi}] \leq 0, \text{ and } \sum_t \beta^t \left[\xi^{\frac{1}{\varphi-1}} c_t^{\frac{1}{1-\varphi}} - n_t w_t \right] \leq 0,$$

respectively.

A little algebra allows us to show that a progressive labor income tax is translated in the static model into a progressive static income tax $T(y) = y - \xi y^{1-\varphi}$ and a change in productivity. That is, under the progressive income the consumer solves

$$V(A, \tilde{W}) = A^{-1} \max_{C \leq \xi Z^{1-\varphi}} \left[\frac{C^{1-\sigma}}{1-\sigma} - \frac{A}{1+\gamma} \left(\frac{Z}{\tilde{W}} \right)^{1+\gamma} \right],$$

for

$$\tilde{W} = \left\{ \sum \beta^t w_t^{\frac{(1+\gamma)(1-\varphi)}{\gamma+\varphi}} \right\}^{\frac{\gamma+\varphi}{(1-\varphi)(1+\gamma)}}.$$

When progressive consumption is used instead, the consumer program becomes

$$V(\tilde{A}, W) = \tilde{A}^{-1} \max_{C \leq \xi Z^{1-\varphi}} \left[\frac{C^{1-\sigma}}{1-\sigma} - \frac{\tilde{A}}{1+\gamma} \left(\frac{Z}{W} \right)^{1+\gamma} \right],$$

for

$$\tilde{A} = \left[\frac{1 - \beta^T}{1 - \beta} \right]^{\frac{\sigma+\varphi(1-\sigma)}{1-\varphi}}.$$

Both could be compared to [Vickrey's](#) income averaging proposal, in which neither A nor W are affected by progressive taxation. Indeed, let us consider the income averaging

tax system that defines the intertemporal budget constraint

$$\sum_t \beta^t c_t \leq \hat{\xi} \left[\sum_t \beta^t (n_t w_t) \right]^{1-\xi},$$

where $\hat{\xi} = \xi [(1 - \delta^T)/(1 - \delta)]$, guarantees that all sequences of constant earnings and consumption that are feasible in one system are also in the other two.

It is then possible to prove the following proposition.

Proposition 1 *If, for all t , $(1 + r_t) = \beta^{-1}$ then [Vickrey's](#) income averaging and the progressive consumption systems lead to the same allocations. Moreover, both are strictly better than the one implemented by the progressive income tax schedule if w_t is not constant.*

When $(1 + r_t) \neq \beta^{-1}$, then [Vickrey's](#) system is strictly better than both. With $\beta < (1 + r)^{-1}$, agents choose an increasing consumption profile. Progressive consumption taxes act as an increasing consumption tax, which, as is well known is equivalent to capital taxation. This leads to an inefficient use of resources attained through labor. When $\beta > (1 + r)^{-1}$ then it acts as a capital subsidy, also leading to inefficiencies. If progressive income taxes are used instead, the intertemporal inefficiency in labor supply applies even when wages are constant.

2 Model

The discussion in Section 1 is based on an overly simplistic description of the economy made to isolate the key forces we want to study. We have assumed perfect capital markets, ruled out any type of uncertainty, and ignored any bequest motive capable of reproducing the pattern of inheritances found in the data. All these aspects will play a role as we assess the benefits of such a reform under a more realistic setting.

2.1 Demography

Each period, j , a new generation is born. Uncertainty regarding the time of death for each person is captured by the fact that each individual faces a probability ψ_{t+1} of surviving to the age $t + 1$ conditional on being alive at age t . Hence, an individual born in period j is alive in period $j + t$ with probability $\prod_{k=1}^t \psi_k$. We also assume that there is $T > 0$ such that $\psi_{T+1} = 0$.

We assume independence of individual death shocks and appeal to the law of large numbers to map the survival probability into the time invariant age profile of the population denoted $\{\mu_t\}_{t=1}^T$. Letting g_n denote the population growth rate, the fraction of agents t

years old in the population is found using the following law of motion

$$\mu_t = \frac{\psi_t}{1 + g_n} \mu_{t-1}$$

with $\mu_t \geq 0$, $\sum_{t=1}^T \mu_t = 1$.

Our focus is on working lives, hence an agents life starts at the age $t = 20$. We also consider $T = 90$. For most of our analysis we will focus on the steady-state allocations. Since it greatly simplifies the notation we drop all time indices, j , from aggregate variables and use t to represent age.

2.2 Technology

Technology is standard. The production side of the economy aggregates and the technology for producing the consumption good is summarized by a Cobb-Douglass production function with constant returns to scale. That is,

$$Y = BK^\alpha N^{1-\alpha}$$

where K is aggregate capital, N is aggregate efficient units of labor, and B is a scale parameter. Every period, the standing representative firm solves the static optimization problem

$$\max_{K,N} \{BK^\alpha N^{1-\alpha} - \delta K - wN - rK\}$$

where r is the rental rate of physical capital and w is the rental rate of efficiency units of labor, i.e. the wage rate. Note that we assume that the rental rate of capital is net of depreciation costs which are born directly by the firm.

The first order conditions for the firm's profit maximization problem are,

$$(1 - \alpha)BK^\alpha N^{-\alpha} = w, \tag{1}$$

and

$$\alpha BK^{\alpha-1} N^{-\alpha} - \delta = r. \tag{2}$$

2.3 Households

Preferences Individuals derive utility from consumption, c , and leisure, l . Preferences defined over random paths of (c_t, l_t) are represented by the time-separable von-Neumann

Morgenstern utility function,

$$\mathbb{E} \left[\sum_{t=1}^T \beta^{t-1} \left(\prod_{k=1}^t \psi_k \right) U_t(c_t, l_t) \right], \quad (3)$$

where β is the subjective discount factor, and $\mathbb{E}$ is the expectation operator conditional on information at birth.

We allow preferences over consumption-leisure bundles to vary with age by indexing the flow utility by t . Specifically, flow utility is of the form

$$U_t(c_t, l_t) = \frac{(c_t^{1-\rho_t} l_t^{\rho_t})^{1-\gamma} - 1}{1-\gamma}, \quad (4)$$

for $\rho_t \in (0, 1) \forall t, \gamma > 0, \gamma \neq 1$.

The fact that we allow the marginal rate of substitution between leisure and consumption to vary with age gives us more degrees of freedom to try to match the behavior of hours along the life-cycle. As ρ decreases, agents become more willing to forego leisure to obtain more consumption. Since lower ρ implies higher (in absolute terms) Frisch elasticity of leisure, the two effects compound to generate more variation in the Frisch elasticity of labor supply.

Another issue raised by our choice of time-varying preferences is that for a given n , the marginal utility of consumption varies with age. A perfectly smooth profile would no longer be optimal even if hours were constant. This choice of preferences some care in measuring and reporting welfare gains.

Note also that this specification for preferences implies a Frisch elasticity of labor supply which decreases with hours worked. Indeed, let ϵ^f denote the Frisch elasticity of labor supply. Then,

$$\epsilon_t^f = \frac{(1-\gamma)(1-\rho_t) - 1}{\gamma} \frac{1-n_t}{n_t}$$

Elasticities are, of course, crucial in the determination of optimal taxes. So, understanding how hours vary along the life-cycle will be important for welfare assessment.

Wealth in the Utility Function In order for the model to be able to reproduce the observed wealth distribution observed, we assume that individuals derive utility from leaving a bequest, a , to their children

$$v(a) = \eta_1 \left(1 + \frac{a}{\eta_2} \right)^{(1-\gamma_c)}. \quad (5)$$

The idea of including a utility for wealth as a motive for accumulation is not new, but

has been regaining traction in recent times. We motivate it through a utility from bequeathing which, first introduced by [Andreoni \(1989\)](#), has been called a 'warm glow' or 'joy-of-giving' motive. From the perspective of our work two things are of essence. First, it creates non-homotheticities that have been proven essential to generate a wealth distribution that is closer to what one observed in the data thanks to its functional form, described in (5), first proposed by [De Nardi \(2004\)](#). Second, a typical warm glow motivation implies that it is the amount received by the donee that matters for the donor not the amount that is donated. An alternative view, sometimes referred to as 'capitalistic spirit' in reference to [Weber's](#) *'The protestant ethic and the spirit of capitalism'* implies the opposite. One may think that the distinction is immaterial for our analysis since we do not consider estate taxation. Yet, when we move to a consumption based taxation the same amount of wealth bequeathed entails less consumption than in a world in which it is labor income that is taxed. Because neither of these reduced forms of capturing an independent desire for accumulation have a clear micro-founded basis, we do not have a clear guide for defining how bequests react to consumption taxes. So, while in our baseline case we assume preferences of the form (5) we also explore the alternative specification (15) in Section 4.

Labor Supply and Retirement Every period, individuals choose labor supply, consumption, and asset accumulation to maximize their objective, (3), subject to a budget constraint which we shall explain momentarily.

Each person has a unit time endowment which can be directly consumed in the form of leisure, l , or used in market related activities. An agent's period-by-period time constraint is $l_t + n_t = 1$. An individual of age t who works for n hours supplies to the market a total of $n_t s_t e^{(u+z_t)}$ efficiency units which are paid at a rental rate w . The variable $u \sim \mathcal{N}(0, \sigma_u^2)$ is a permanent component of an individual's skills. It is realized at birth and retained throughout one's life. On the other hand, z follows an AR(1) process, $z_t = \varphi_z z_{t-1} + \varepsilon_t$, with innovations $\varepsilon_t \sim \mathcal{N}(0, \sigma_\varepsilon^2)$.

Whereas u aims at capturing the heterogeneity at birth, the most relevant source of welfare variation, z is the main source of uncertainty affecting long term choices. The parameter φ_z accommodates the empirically observed persistence of productivity shocks. Finally, s_t is what we call the age-efficiency profile.

Labor productivity shocks are independent across agents. As a consequence, there is no uncertainty regarding the aggregate labor endowment even though there is uncertainty at the individual level. Retirement is mandatory at the age of 65, or $t = 46$.

Asset Accumulation Besides choosing how much leisure to consume, individuals trade a risk free asset which holdings we denoted by a_t . Asset holdings are subject to an exogenous lower bound. More precisely, for our main exercise, we follow [Conesa et al. \(2009\)](#)

in assuming that agents are not allowed to contract debt at any age, so that the amount of assets carried over from age t to $t + 1$ is such that $a_{t+1} \geq 0$. Because no agent can hold a negative position in assets at any time, we assume without loss that asset takes the form of capital, $a_t = k_t$, as in [Aiyagari \(1994\)](#).

As we shall make clear, there is exogenous (as well as endogenous) variation in productivity along the life-cycle. Consumption smoothing thus provides a reason for one to accumulate assets. Another aspect of choices is that individuals may resort to self-insurance to protect themselves against the uncertainty on labor income. Savings will be, to some extent, motivated by precautionary reasons. By precluding borrowing we potentially affect the valuation of future social security benefits as compared to what we would observe in a world with perfect capital markets.

In addition, households are born with initial wealth endowment a_1 drawn from an endogenous distribution that integrates up to the overall amount of wealth bequeathed in the economy by the deceased households. In order to account for the intergenerational persistence in wealth inequality, we assume that the draw is correlated with the permanent component of the elderly's skills. In particular, let $\hat{a}_u$ and $\sigma_{\hat{a}_u}^2$ denote the mean and the variance of wealth bequeathed by the deceased households with permanent skill u , respectively. Thus, an agent at the first age with permanent skill u draws her initial wealth from a log-normal distribution with mean $\hat{a}_u$ and variance $\sigma_{\hat{a}_u}^2$.¹⁰

Budget Constraints To write each agent's flow budget constraint we need to specify the fiscal policy that is being used by the Government. In our case, it is important to distinguish between the current fiscal policy, needed to calibrate the model, and the ones we evaluate. The current tax system will be the benchmark for our studies.

2.4 The Government

Tax revenues are raised to finance an exogenous flow of expenditures, G , and possibly to cover the social security deficit. The government levies taxes on capital income, consumption, and labor income. We assume that capital income is taxed at a flat rate τ_k . Labor income taxes are given by $T(y) = \max\{y - \xi y^{1-\varepsilon}, 0\}$. This specification was first proposed by [Musgrave \(1959\)](#) and since adopted by [Feldstein \(1969\)](#); [Bénabou \(2000, 2002\)](#); [Heath-](#)

¹⁰The reason why we model the intergenerational links in the form of both accidental and voluntary bequests this way is to avoid the inclusion of an additional state variable regards to the parents ability as is done in [De Nardi \(2004\)](#). We assess the importance of bequest for our results in section 4.3.

cote et al. (2017a), to name a few. Similarly, consumption taxes are specified as

$$T(\hat{c}) = \max \left\{ \hat{c} - \left(\frac{\hat{c}}{1 + \tau_c} \right)^{\frac{1}{1+\varphi_c}}, 0 \right\}, \quad (6)$$

The government also runs a social security system with contributions and benefits that are equal in equilibrium. Retirement benefits are given by

$$b(x) = \begin{cases} \theta_1 x, & \text{if } x \leq y_1 \\ \theta_1 y_1 + \theta_2 (x - y_1), & \text{if } y_1 \leq x \leq y_2 \\ \theta_1 y_1 + \theta_2 (y_2 - y_1) + \theta_3 (x - y_2), & \text{if } y_2 \leq x \leq y_{\max} \\ \theta_1 y_1 + \theta_2 (y_2 - y_1) + \theta_3 (y_{\max} - y_2), & \text{if } x \geq y_{\max} \end{cases} \quad (7)$$

where x is the individual's average life-cycle earnings and $\theta_1 > \theta_2 > \theta_3$.

The law of motion for x is given by

$$x_{t+1} = \frac{[t-1]x_t + \tau_{ss} \min\{y_t; y_{\max}\}}{t}, \quad (8)$$

and $x_T = x$.

In addition, government provides individuals a minimum consumption, $\underline{c}$. We assume that transfers, tra , are conditional on individuals available resources. In particular, following Hubbard et al. (1995), we specify:

$$tra = \max\{\underline{c} - [1 + (1 - \tau_k)r]a - \tilde{y} + T(\tilde{y}), 0\} \quad (9)$$

This equation implies that government transfers fill the gap between illiquid resources – which may include not only wealth and labor income, but also other government transfers such as social security benefits – and the consumption floor. Thus, individuals can always consume at least $\underline{c}$. Equation (9) is intended to be a model counterpart for means-tested programs such as Food Stamp, AFDC, Section 8 housing assistance, Medicaid and SSI.

Recursive Formulation of Households' Problem

A crucial definition in all that follows is that of consumption expenditure, $\hat{c}$, which is related to consumption through $c = \hat{c} - T^c(\hat{c})$. Using the definition of $\hat{c}$ above, we define the

agents' budget constraint at individual state $\omega_t := (a_t, u, z_t, x_t) \in \Omega$ as

$$\mathcal{B}(\omega) := \{(\hat{c}, a', l) \in \mathbb{R}_+^2 \times [0, 1] \mid \hat{c} + a' \leq [1 + (1 - \tau_k)r] a + \tilde{y} - T(\tilde{y}),$$

$$\tilde{y} = y - \tau_{ss} \min\{y, y_{\max}\}, y = [1 - l]se^{(u+z)}\}.$$

Let $V_t(\omega_t)$ denote the value function of an individual aged $t < T + 1$, where $\omega_t = (a_t, u, z_t, x_t) \in \Omega$ is the individual state. If the agent has not yet attained the retirement age, his or her optimization problem at age t can be recursively represented as follows. Let $\omega' = (a', u, z', x')$, then,

$$V_t(\omega) = \max_{(\hat{c}, a', l) \in \mathcal{B}(\omega)} U_t(c, l) + \beta \mathbb{E}_{z'} [V_{t+1}(\omega')], . \quad (10)$$

Although we have only used individual state variables in ω , prices are, of course, also relevant state variables. To solve the model we will need to find the equilibrium prices by explicitly taking into account how they enter the policy functions associated with (10). Still for economy in notation we drop them from ω .

For retired agents, no labor supply decision is made. Moreover, we assume that, after retirement agents start deriving utility from bequests as well. A retired agents' value function is, therefore,

$$V_t(\omega) = \max_{(\hat{c}, a', l) \in \mathcal{B}(\omega)} U_t(c, 1) + \beta [\psi_{t+1} \mathbb{E}_{z'} V_{t+1}(\omega') + (1 - \psi_{t+1}) \nu(a')]. \quad (11)$$

Finally, considering that agents die for sure at age T , we have $V_{T+1}(\omega_{T+1}) = \nu(a)$.

2.5 Recursive competitive equilibrium

Recall that we are describing our economy in the steady state. This has allowed us to drop time subscripts and use t to denote age, instead.

At each point in time, agents differ from one another with respect to age their age, t , and their individual state, $\omega = (a, u, z, x)$. Agents of age t identified by ω , are distributed according to a probability measure λ_t defined on Ω , as follows. Let $(\Omega, \mathcal{F}(\Omega), \lambda_t)$ be a probability space, where $\mathcal{F}(\Omega)$ is the Borel σ -algebra on Ω : for each $\eta \subset \mathcal{F}(\Omega)$, $\lambda_t(\eta)$ denotes the fraction of agents aged t that are in η .

Given the age distribution, λ_t , $Q_t(\omega, \eta)$ induces the age $t + 1$ distribution λ_{t+1} as follows. $Q_t(\omega, \eta)$ determines the probability of an agent at age t and state ω to transit to the set η at age $t + 1$. $Q_t(\omega, \eta)$, in turn, depends on the policy functions in (10), and on the exogenous stochastic process for z .

The following equilibrium definition refers to the benchmark. Analogous definitions

apply to the economy under the policy experiments we run.

Definition Given the policy parameters, a **recursive competitive equilibrium** is a collection of value functions $\{V_t(\omega)\}$, policy functions for individual asset holdings $d_{a,t}(\omega)$, for consumption $d_{c,t}(\omega)$, for labor supply $d_{n_w,t}(\omega)$, prices $\{w, r\}$, age dependent but time-invariant measures of agents $\lambda_t(\omega)$, transfers ϵ such that:

- i) $\{d_{a,t}(\omega), d_{n_w,t}(\omega), d_{c,t}(\omega)\}$ solve the dynamic problems in (10);
- ii) individual and aggregate behaviors are consistent, that is:

$$\begin{aligned} K &= \sum_{t=1}^T \mu_t \int_{\Omega} d_{a,t}(\omega) d\lambda_t \\ N &= \sum_{t=1}^T \mu_t \int_{\Omega} d_{n_w,t}(\omega) s_t(\omega) \exp(u + z_t) d\lambda_t \\ C &= \sum_{t=1}^T \mu_t \int_{\Omega} \{d_{c,t}(\omega)\} d\lambda_t; \end{aligned}$$

- iii) $\{w, r\}$ are such that they satisfy the optimum conditions (2) and (1);
- iv) The final good market clears:

$$C + G + \delta K = K^\alpha N^{1-\alpha};$$

- v) the probability measure $\lambda_t(\omega)$ follows the law of motion:

$$\lambda_{t+1}(\eta) = \int_{\Omega} Q_t(\omega, \eta) d\lambda_t \quad \forall \eta \in \mathcal{F}(\Omega);$$

- vi) the overall amount of wealth bequeathed by the deceased households correspond to the aggregate initial wealth endowment:

$$\mu_1 \int_{\Omega} a_1(\omega) d\lambda_1 = \sum_{t=T_r}^T \mu_t \int_{\Omega} (1 - \psi_{t+1}) d_{a,t}(\omega) d\lambda_t$$

where T_r denote the retirement age.

vii) taxes are such that the government's budget constraint,

$$\begin{aligned} \sum_{t=1}^T \mu_t \int_{\Omega} T(d_{n_w,t}(\omega) s_t(\omega) \exp(u + z_t) - \\ T_{ss}(d_{n_w,t}(\omega) s_t(\omega) \exp(u + z_t))) d\lambda_t + \tau_c C + \tau_k rK + \\ \sum_{t=1}^T \mu_t \int_{\Omega} T_{ss}(d_{n_w,t}(\omega) s_t(\omega) \exp(u + z_t)) d\lambda_t = G + B \quad (12) \end{aligned}$$

are satisfied every period, where B are the total benefits that are due, given the social security system that we consider.

Item (vii) is redundant if conditions (i)–(vi) hold.

2.6 Calibration

To carry out our quantitative analysis, we need first to find values for all the parameters of the model. We accomplish this by calibrating the model to the U.S. economy.

The population age profile $\{\mu_t\}_{t=1}^T$ depends on the population growth rate, g_n , the survival probabilities, ψ_t , and the maximum age, T , that an agent can live. Agents enter the economy at age 20 and live for 71 years, $T = 71$, so that the real maximum age is 90 years.

Data on survival probability by age were extracted from [Bell and Miller \(2005\)](#). Given the survival probabilities, the population growth rate is chosen so that the age distribution in the model replicates the dependency ratio observed in the data. By setting $g_n = 0.009$, the model generates a dependency ratio of 23.27%, which is close to the dependency ratio observed in the data for the year 2017.

To calibrate the preference parameters we proceed as follows. First, we choose the discount factor β in such a way that the equilibrium of our benchmark economy implies a capital-output ratio around of 3.0, which is the value observed in the data. Then we fix the parameter γ to 4.0, from micro evidence, and choose the share of leisure in the utility function, ρ_t , to match average hours for different age groups. In particular, we assume that $\rho_t = \rho_0 + \rho_1 t$. To calibrate ρ_0 , we use the average working hours for ages 19 – 40 and for ρ_1 the average between 41 – 60. The first group works on average 37.86 while the second 40.37 of their time endowment. For the last 5 years we specify a new profile $\rho_t = \rho_{60} + \rho_2 t$. We calibrate ρ_2 to match the average hours during those last five years equal to 35.16.¹¹

The parameter ϕ_1 represents the weight on the utility from bequeathing. Since it measures the strength of bequest motives, we calibrate ϕ_1 so that the ratio of the median wealth held by the individuals aged 75 and above to that of all individuals is 1.8, as reported in the

¹¹The data for hours worked are from IPUMS (U.S. Department of Commerce, Bureau of the Census 2015).

Statistical Abstract of the United States (US Census Bureau 2009). The term ϕ_2 is the shifter of bequests in utility function. It reflects the extent to which bequests are luxury goods, affecting the bequest distribution, especially the high end of it. Thus, we follow [De Nardi and Yang \(2014\)](#) and set ϕ_2 to match the 90th percentile of bequest distribution normalized by income as reported by ?.

The parameters that characterized the stochastic component of individuals productivity are $(\sigma_u^2, \varphi_z, \sigma_\epsilon^2)$. Several authors have estimated similar stochastic process for labor productivity. Controlling for the presence of measurement errors and/or effects of some observable characteristics such as education and age, the literature provides a range of $[0.88, 0.96]$ for φ_z and of $[0.10, 0.25]$ for σ_ϵ^2 . In this article, we rely on the estimates of [Kaplan \(2012\)](#), setting $\varphi_z = 0.94$ and $\sigma_\epsilon^2 = 0.016$. Then, σ_u^2 was chosen for the Gini index for labor income in the model to match its counterpart in the data, which is nearly 0.43. The value obtained for σ_u^2 is in line with the estimates in [Kaplan \(2012\)](#) who provides a point estimate of 0.056 for this parameter. We discretize the two shocks in order to solve the model, using seven states to represent the permanent shock and eleven states for the persistent shock. For expositional convenience, we refer to the two extremes of the grid for the permanent shock as low and high ability type. We assume that individuals start the lifecycle with the mean level of the idiosyncratic shock z .

The values of the technological parameters (α, δ) are also in Table 1. We chose a value for α based on the U.S. time series data from the National Income and Product Accounts (NIPA). The depreciation rate, in turn, is obtained by $\delta = \frac{I/Y}{K/Y} - g$. We set the investment-product ratio I/Y equal to 0.25 and the capital-product ratio K/Y equal to 3.0. The economic growth rate, g , is constant and consistent with the average growth rate of GDP over the second half of the last century. Based on data from Penn-World Table, we set g equal to 2.3%, which yields a depreciation rate of 6.0%.

The age-efficiency profile for the exogenous model is set to be consistent with the values estimated in [Kaplan \(2012\)](#), which are based on the average hourly earnings by age in the PSID. We use a second order polynomial to smooth this profile and extend it to cover ages from 16 to 65.

Finally, we specify the others parameters related to government activity. First, we set government consumption, G , to 18% of the output of the economy under the baseline calibration. Following the literature, we assume a consumption tax of 6% and a capital income tax rate of 30%.¹² The parameter φ , which governs the progressivity of the labor income taxes, is calibrated to match the actual average tax rates. Marginal tax rates are chosen to raise enough revenue to finance government consumption. The value we find for ζ is 0.83.¹³

¹²See, for example, [Fuster et al. \(2007\)](#)

¹³Recall that $\zeta(1 - \varphi)y^{-\varphi}$ is the marginal retention rate, $1 - T'(y)$.

As for Social Security, we set the normal retirement age to $T_r = 46$, which corresponds to the real age of 65. Contributions are of the form $T_{ss}(y) = \tau_{ss} \min\{y, y_{\max}\}$, where $y_{\max}$ is the contribution ceiling. We set the payroll tax rate, τ_{ss} , to 6.2%. The marginal replacement rates in the progressive Social Security payment schedule $(\theta_1, \theta_2, \theta_3)$ are set at their actual respective values of 0.9, 0.32 and 0.15. The bend points where the marginal replacement rates change (y_1, y_2) and the maximum earnings $y_{\max}$ are set equal to the actual multiples of mean earnings used in the U.S. Social Security system so that y_1 , y_2 and $y_{\max}$ occur at 0.21, 1.29 and 2.42 times average earnings in the economy.

3 Main Findings

Our main findings are displayed in Table 2. Before we discuss them it is necessary to define some key statistics.

Welfare The welfare variation (CEV) is calculated as follows: Let $V_1(\omega)$ denote the expected utility of an agent who starts life at state ω under the policy we aim at evaluating. Then, define

$$V_1^0(\omega) = \mathbb{E} \left[\sum_{t=1}^T \prod_{s=1}^t \psi_s (1 + \Delta)^{(1-\rho_t)(1-\gamma)} U_{t,0}(c_t, 1 - n_t) \right]$$

where $U_{1,0}(c_t, 1 - n_t)$ is the flow utility attained by the agent under the benchmark at age t . Our relevant measure of welfare variation is

$$CEV = \min_{\Delta} [\mathbb{E}_{\omega} V_1^0(\omega) - \mathbb{E}_{\omega} V_1^1(\omega)]. \quad (13)$$

Note how we take full advantage of our iso-elastic specification for preferences.

CEV Decomposition Aiming at understanding the source of welfare gains we decompose the CEV in variations that are due to improved insurance and those that are due to a more efficient use of aggregate resources. Let $C_{0,t}$ and $L_{0,t}$ denote average consumption and average hours worked by t years old agents at the benchmark, respectively. Let also $C_{1,t}$ and $L_{1,t}$ denote the same statistics at the alternative tax system. We may, in this case, implicitly define Δ_{lev} through

$$\sum_t \beta^{t-1} \prod_{j=1}^t \psi_t \left(((1 + \Delta_{lev}) C_{0,t})^{1-\rho_t} (1 - L_{0,t})^{\rho_t} \right)^{1-\gamma} = \sum_t \beta^{t-1} \prod_{j=1}^t \psi_t \left(C_{1,t}^{1-\rho_t} (1 - L_{1,t})^{\rho_t} \right)^{1-\gamma}.$$

For $(c_{t,0}, l_{t,0})_t$, the benchmark equilibrium allocations, and $(c_{t,1}, l_{t,1})_t$, the equilibrium allocations under the alternative policy, implicitly define p_0 and p_1 , through

$$\sum_t \beta^{t-1} \prod_{j=1}^t \psi_t \left(((1-p_0) C_{0,t})^{1-\rho_t} (1-L_{0,t})^{\rho_t} \right)^{1-\gamma} = \mathbb{E} \left[\sum_t \beta^{t-1} \prod_{j=1}^t \psi_t \left(c_{t,0}^{1-\rho_t} (1-l_{t,0})^{\rho_t} \right)^{1-\gamma} \right],$$

and

$$\sum_t \beta^{t-1} \prod_{j=1}^t \psi_t \left(((1-p_1) C_{1,t})^{1-\rho_t} (1-L_{1,t})^{\rho_t} \right)^{1-\gamma} = \mathbb{E} \left[\sum_t \beta^{t-1} \prod_{j=1}^t \psi_t \left(c_{t,1}^{1-\rho_t} (1-l_{t,1})^{\rho_t} \right)^{1-\gamma} \right],$$

In both expressions $\mathbb{E}$ denotes the unconditional expectation operator.¹⁴ Then,

$$\Delta_{unc} \equiv \frac{1-p_1}{1-p_0} - 1.$$

It is important to note that Δ_{unc} only takes into account smoothing across agents of the same age. The fact that preferences are age-dependent implies that complete smoothing across ages is not optimal.

Warm Glow and Welfare The use of a warm glow bequest motive some subtle issues for normative analysis. In particular, should the utility gain from bequeathing be included in welfare evaluation?

In contrast with other form of externalities, if it is the act of bequeathing that is valued by agents, then the final allocation of resources need no longer suffice to define welfare. As [Diamond \(2006\)](#) explains. "The mathematical formulation of warm glow preferences would also apply to a particular structure of ordinary externalities. In that case, we would count the utility of own consumption, w . What distinguishes a warm glow, and so the question of design of the social welfare function, is the distinction that actual resource use is the same with public and private public good supply, but individual experience is different in these two cases."

At the end of the day, the normative criterion choice matters for welfare evaluation. As is apparent from (13) the welfare assessment we use ignores the intrinsic (warm glow)

¹⁴That is, we take expectations with respect to all possible permanent types and histories.

utility derived from the act of bequeathing. If we consider aggregate welfare gains, our measure is probably on the conservative side for the type of reform we study. It may, however, bias how we measure the distribution of gains against the richest agents.

Insurance To assess the amount of consumption insurance for each possible policy we define

$$\phi = 1 - \frac{\text{Cov}(d\tilde{c}_{i,t}, dz_{i,t})}{\text{Var}(dz_{i,t})} \quad (14)$$

where $\tilde{c} = \log(c)$. The statistic ϕ is 0 whenever every change in z is translated directly into a variation in consumption. At the other extreme, when there is full insurance, idiosyncratic exogenous changes in productivity do not translate into changes in consumption and ϕ is equal to 1.

Benchmark with Consumption Taxes The findings for our baseline reform are displayed in Table 2. The first four lines in the table display the values for the policy parameters at each alternative tax system. In the next six lines the findings for aggregate variables are reported. Next, the revenue distribution block, explains the relative importance of each tax base. The consequence of each system for consumption smoothing is accounted for by the next two blocks. First, the Gini block displays this representative cross-sectional inequality statistic for income before and after taxes, assets and consumption. In the following block the insurance coefficient measures how consumption responds to variations in each agent's productivity changes. It requires longitudinal information.

The first column displays all the relevant statistics for the benchmark allocation. In the second we consider use the same progressivity parameter from the benchmark system but let the progressivity apply to consumption taxes, instead. The level parameter is adjusted to keep tax revenues at the benchmark level.

Hours are slightly increased in benchmark with consumption taxes relative to benchmark, but there is also a 5% wage increase which indicates that simply moving to progressive consumption taxes has a small negative impact on hours.

The large change in the capital stock that arises by simply moving to consumption taxes raises the question of what features matter for our results. We expect to see an important impact on savings; for the preference specification of Section 1 the coefficient of relative prudence would be increased by 50% with the optimal progressive consumption taxes. Still, the move to a consumption tax cannot be just assumed to be inconsequential in the presence of a warm glow motive for asset accumulation. The findings in a revenue preserving reform displayed in Table 5 provides further hints to the importance of this channel.

3.1 Optimal Tax Systems

If both distribution and insurance matter, then any tax system trades-off improvements in these two dimensions with potential inefficiencies due to the distortions generated by taxation. The first important thing to notice about the current (benchmark) system is that, from a Utilitarian perspective it errs in the direction of sacrificing too little efficiency and promoting too little redistribution/insurance. This is made clear by our findings regarding the optimal labor income tax schedule. Indeed, if we compare the allocation resulting from optimal labor income taxes with the benchmark allocation we find the efficiency CEV to be 4.02% lower which is more than compensated by a 6.49% increase in the gains from better smoothing as captured by the CEV uncertainty statistics.¹⁵

Such a perspective is an important one to bear in mind if we are to understand the results that follow. In our heuristic account of the forces at play, we have emphasized the efficiency gains from moving to a progressive consumption taxes based system. Yet, as one can see, the bulk of the welfare gains, 11.10%, for the optimal consumption tax system are found in the CEV uncertainty statistics. The efficiency gains are explored by the optimal system to provide better insurance and redistribution. Indeed, reinforcing the evaluation above is the labor/hour ratio that compares productivity adjusted hours with total hours: for all but the unrestricted consumption taxes case, the ratio is reduced when compared to the benchmark: the current system provides too little insurance/redistribution for the amount of inefficiencies it creates. To understand how these changes are induced by policy one has to consider the top four rows in Table 2. Progressivity increases from 0.090 in the benchmark to 0.220 when we move to optimal labor income taxes (0.130, if taxes are restricted to be non-negative). These numbers are even larger for the case of consumption taxes. Figure 3 displays the static budget set generated by each tax schedule. It is interesting to note that the benchmark schedule is very close to the optimum if progressivity applies to labor income and taxes are restricted to be non-negative.

Figure 3 is very useful if our goal is to compare between systems that share the same tax base. Not so much to compare progressivity across systems that differ with regards to the base since not all earnings are consumed along the life-cycle. This is why we have displayed the budget sets associated with the different tax bases in different panels. To compare across bases we shall now examine progressivity by means of the fraction of revenues raised from each earnings quintile. This approach addresses the aforementioned concerns at the cost of focusing on endogenous objects.

In our main experiment we completely eliminate labor income taxes which leads to

¹⁵Another sign that it is not though efficiency gains that the allocation is improved is the slight reduction in the labor/hours statistics. The tax system is discouraging effort from those with higher productivity more at the optimum than at the benchmark.

a large increase in revenues raised through consumption taxes.¹⁶ Progressivity in consumption taxes is apparent in Table 4, where we measure taxes paid against labor earnings. Interestingly, the most progressive system under this metric is exactly the one with unrestricted progressive consumption tax schedule. This is all the more significant once one notes that the table displays revenues against income which is not too highly correlated with consumption under this schedule – see Figure 3. When unrestricted income taxes are used that we see a highly progressive income tax schedule – in particular note that the four lowest deciles in the income distribution pay negative taxes on average – which might lead us to expect overall progressivity to be at its maximum. However, because we keep capital income tax rates fixed and wealth becomes more concentrated, the unrestricted consumption tax system ends up being even more progressive despite the fact that capital income taxes now play a lesser role due to a reduction in the return to capital.

The large welfare gains associated with the reform are, therefore, in a great part due to better consumption smoothing, both in the form of better distribution (of consumption) and more insurance. The improvement in insurance is displayed in the last three rows of Table 2. Insurance is higher in all reforms, but the highest improvement occurs when we use the unrestricted progressive consumption schedule. In this case ϕ increases from 0.349 in the benchmark to 0.474 at the optimum. For this system, insurance is best for the low type agents. The combined effect of better insurance and distribution induced by progressive consumption schedules is captured by a substantially lower consumption Gini, despite the fact that the same system induces the largest after tax income Gini and largest wealth Gini of all systems. These effects are, perhaps, better illustrated by regressing consumption against income. While the coefficient hovers around .6 whenever income taxes are used, and goes down to around 0.3 when progressive consumption taxes are used, instead.

Table 3 provides further details regarding wealth distribution. The most egalitarian wealth distribution is attained with the unrestricted progressive income tax schedule. In contrast, progressive consumption taxes worsens the wealth distribution even when compared to the benchmark. It is in fact possible to check that the normalized wealth distribution associated with the unrestricted optimal consumption tax is dominated in the second order stochastic dominance sense by the benchmark distribution which is, in turn, dominated by the distribution induced by the unrestricted labor income tax. Hence, if the likelihood of fundamental tax reform depends on these currently hotly debated statistics we have only reasons to be pessimistic. Ironically, if a specific group were to oppose the reform, it should be the high productivity group, which would have a substantial welfare loss – $CEV = -12.70\%$ –, which under a utilitarian metric is more than compensated by

¹⁶See however Section 4.1, where we allow revenues to be raised by linear labor income taxes.

the extraordinary gains – $CEV = 40.34\%$ – obtained by the lowest productivity agents in society.

Beyond the impact on redistribution/insurance this fundamental tax reform leads to important changes in aggregate variables. GDP is almost 10% lower than the benchmark in the allocation induced by the optimal unrestricted progressive consumption taxes despite an increase in the capital income ratio increases by more than 13%. Of course once one accounts for the drop in GDP, then one can see that the capital stock has increased but only by 4%, and that the biggest chunk in the GDP change is explained by a reduction in hours of more than 18%. The large increase in progressivity highlighted in Figure 3 which still matters from a life-cycle perspective – recall the presentation in Section 1 – can partly explain this finding. Further evidence on this regard is provided by the fact that an even larger drop in GDP occurs when unrestricted labor income taxes are used. Interestingly, the decline in hours in this case is slightly larger than the one for optimal positive progressive consumption taxes which has similar progressivity – see Figure 3.

Most important for our analysis, however, is how hours adjust along the life-cycle, our original motivation for studying this reform. In this case, Figure 2 makes it plain that while the increase in progressivity that characterize optimal labor income taxes makes hours less responsive to changes in productivity, the move to optimal taxes make them substantially more responsive.

4 Robustness

The move to a system based on progressive consumption taxes has the potential to generate very large welfare gains. Our baseline model is in most aspects canonical. Yet, one feature which may raise some concern is the direct impact of wealth on utility as captured by the warm glow bequest motive.

We have used a specialized version of warm glow motives in which it is only the amount bequeathed that matters for the agent. Consumption taxes change the relative price of consumption with respect to wealth, thus increasing the incentives for wealth accumulation. Beyond this, the presence of a warm glow motive raises some important issues for normative analysis. As explained by [Kopczuk \(2013\)](#), under this assumption about bequest motives welfare improvements are possible by simply taxing bequest and returning to the agents. We eliminate this latter source of spurious gains by not including the utility associated with the warm glow term in the welfare evaluation, but we cannot ignore the fact that our specification has important consequences for choices that are not as well settled as other aspects of our model. At issue is the fact that most studies focus on the bequest or wealth taxation. We are not aware of any study directly aimed at assessing the

consequences of consumption taxes on bequests.¹⁷ Hence, to assess the robustness of our findings we consider different variations of our experiments that focus on the impact of consumption taxes on welfare accumulation.

First, we restrict the amount of revenues that can be raised by consumption taxes. Second, we modify preferences to allow the bequest motive to depend on the present value of consumption that wealth can purchase instead of wealth itself. Third, we eliminate the impact of tax changes on savings that arise from the warm glow motive by holding fixed the benchmark policy functions. Finally, we shut down all general equilibrium effects on prices.

4.1 Restricting the role of Consumption Taxes

For the first set of robustness checks we restrict the amount of taxes raised through consumption taxes. Our findings are summarized in Table 5.

We start by repeating the exercise from Section 3 in which we use the benchmark labor income tax schedule and progressivity, but now we require the total amount of revenues to be raised via consumption taxes to be held constant at 16.60% of total revenues. CEV drops from 5.27% to 'only' 2.15%.¹⁸ Despite the fact that revenues raised from consumption taxes are held fixed the capital/income ratio still increases by almost 3%. Wealth is disproportionately accumulated by those with higher income, hence progressivity and not only the level of consumption taxes matter for capital accumulation. All consumption smoothing statistics – Consumption Gini and Insurance Coefficient – decline with this reform. In a sense it is unfair to say that progressivity is preserved with this reform. Since not all earnings are consumed, progressivity applies to a smaller share of agents' lifetime income. Moreover, because the rich save more, progressivity is significantly reduced. So these results really apply to a policy that is very far from the optimum.

If we maintain the restriction that revenues raised by consumption taxes cannot exceed 16.60%, but allow progressivity to be optimally chosen, then welfare gains are again very high, $CEV = 8.62\%$. These welfare gains arise *despite a 9% reduction in the capital stock!* The emphasis is important. If we restrict consumption tax revenues, we attain lower, albeit still very large, welfare gains in the steady state comparison. Note that this lower gain is accompanied by a large decline in capital stock, whereas in our baseline exercise the capital

¹⁷Diamond (2006); Kopczuk (2013) both comment on the normative shortcomings of relying on a reduced form motive. Because we do not know the underlying reasons why people hold on to wealth in ways that cannot be accounted for by pure life-cycle or pure altruistic motives, we have no *a priori* views on how different tax schemes should affect choices. Most studies have focused on how bequests respond to estate taxes. Yet, without a micro-foundation for the reduced form it is not clear what these findings have to say about behavioral responses to consumption taxes.

¹⁸To put in perspective, recent work by Kapička (2020) finds a maximum of 2.98% gain from a fundamental tax reform which introduces history dependence on labor income taxes.

stock slightly increases. The large drop in the capital stock suggests very large additional welfare gains if the transition is taken into account.

Next, we fix the level term, ζ , in the labor income schedule at the value that preserves average income taxes at the benchmark. We then follow with the same procedure. We find that welfare gains remain very high, $CEV = 11.48\%$, which are however 15% lower than the welfare gains obtained in our main exercise, $CEV = 13.34\%$. The lower welfare gains occur with a stock of capital that is 4% lower than the one in our baseline reform.

A more relevant drop, more than 50%, would occur if consumption taxes were constrained to be non-negative. It is not hard to understand why. When taxes cannot be negative, the scope for progressivity for a given revenue requirement is reduced. Income taxes, by further reducing the revenue requirements lead to lower progressivity in consumption taxes. With unrestricted consumption taxes this effect is avoided and progressivity is only slightly lower than in our main policy reform.

With regards to the impact on aggregate variables, the increase in the capital/income ratio is only 2/3 of the one which occurs in our main policy reform. This leads to a slightly larger drop in GDP – -11% against -10% – despite a slightly smaller reduction in hours -17% against -18% . Small differences in insurance and consumption Gini helps explain why welfare gains are smaller in this case. It is interesting to note how the use of income taxes, by reducing the scope for progressivity of restricted consumption taxes, worsens their consumption smoothing role. In our main policy, reform the largest reduction in the consumption Gini occurs when taxes are restricted to be non-negative. When labor income taxes are required to be used the Gini coefficient is raised from 0.225 to 0.302, which explains the bulk of welfare losses.

All in all, requiring some of the revenue to be raised by income taxes lowers welfare gains both through the reduced progressivity that characterizes the optimal reform in this case and through lower savings that are partially explained by these changes in progressivity, but possibly also by the differences in impact through the warm glow motive.

4.2 The bequest motive

[Kopczuk \(2013\)](#), emphasizes the importance of how one models the warm glow motive for the behavioral consequences of estate taxation. In particular, is it wealth bequeathed or is it wealth inherited that is valued by individuals? Although we do not consider estate taxation, an analogous issue arises regarding the amount of consumption that wealth affords.

Hence, to further assess the importance of our chosen warm glow specification, we assume that agents are interested in how much their inheritance is translated into con-

sumption for their descendants. Formally, we modify (15) to account for taxes using

$$v(a) = \eta_1 \left(1 + \frac{a - (1+r)T(ra)/r}{\eta_2} \right)^{(1-\gamma_c)}, \quad (15)$$

where $T(\cdot)$ is given by equation 6. Our findings are summarized in Table 6.

If agents are concerned with the consumption equivalent of their wealth instead of wealth per se, this reduces the reform's impact on capital accumulation with a drop of four percentage points in the capital/income ratio, which is very similar to what we found in the alternative policy reform. Welfare gains are lower, but still substantial – $CEV = 10.95\%$.

While difference in preferences directly explain some of the changes, a possible indirect effect through a different choice of progressivity cannot, in principle, be ignored. Yet, what we see is a very mild decrease in progressivity. This is also reflected in similar values for the consumption Gini and the insurance coefficients. A larger drop in GDP, therefore, seems to account for most of the difference in welfare gains.

4.3 Constraining Agents' Choices

For the next series of robustness we consider some restricted choices that agents can make. Table 7 displays our main findings.

First, we borrow the policy function from the benchmark, $d_{a,t}(\cdot)$ and generate an asset accumulation path from $a' = d_{a,t}(\omega)$, but allow the agents to re-optimize by choosing an adjustment, $\hat{a}'$, which, however, does not take into account the warm glow motive for asset accumulation. In the first column of Table 7 we assume that only upward adjustments, $\hat{a}'$ are possible. Second, we allow arbitrary adjustments that allow $\hat{a}' < 0$ but respect $\hat{a}' + a' \geq 0$. The most striking aspect about this latter case is how similar the findings are when compared to our baseline findings for unrestricted consumption taxes in Table 2. Welfare gains are slightly *larger* than in our baseline reform. The capital stock is about the same highlighting the fact that the most relevant impacts of our reform cannot be accounted for by the presence of a warm glow motive. Third, we consider the case in which no capital accumulation adjustments are allowed. This is slightly different from allowing agents to solve in every period and in every state of nature a static labor supply problem where the budget constraint is $\hat{c} \leq nw + B$ for $B = [1+r]a - a'$, the benchmark asset process. If this was the case agents which were not constrained at the benchmark would respond less to the reform since it would not be the Frisch elasticity but the Marshallian elasticity that would determine the labor supply responses. Here, however, it is the policy function that is borrowed from the baseline exercise; it does not respond to the contemporaneous choice of effort.

The capital/output ratio displays a large increase in the last exercise, which highlights

a fact that we have been emphasizing all along. Since wealth accumulation is unchanged the only reason why capital/output ratio is because hours decline responding to a large increase in progressivity. When compared to an equally progressive labor income tax schedule, marginal tax rates are smaller in states of nature in which agents were saving (high productivity states) and larger in states in which they were decumulating. This emulates some of the efficiency gains in the distribution of effort across time and states of nature. Welfare gains remain very large.

The final exercise, displayed in the last column of Table 2 measures the 'mechanical' impact of our reform. Because agents do not respond to the reform, the planner has a lot of latitude to increase welfare. Reliance on consumption taxes is lower in this case, which again suggests that it is not through changes in bequest that welfare gains are attained.

4.4 Partial equilibrium

When prices are held fixed, the dampening effects on prices are muted and all allocative consequences become stronger – Table 6. Capital income ratio now increases more than 30%, with an increase in capital stock of about 25%. Average hours drop slightly more than when prices change and welfare gains reach 14.45%. The real variable that best explains the additional gains in the case of consumption taxes is GDP, which now drops 'only' four and a half percentage points.

Interestingly, relevant gains relative to the case in which general equilibrium effects are taken into account only occur in the case of unrestricted progressive consumption taxes. For both the restricted and unrestricted labor income taxes, welfare gains are slightly larger in partial equilibrium. However, if we require consumption taxes to be non-negative the absence of price changes dampens welfare gains from 9.04 to 7.48%, which is still a very large number. But why is it only the unrestricted progressive consumption tax reform that benefits from holding prices fixed? It seems that the benefit from higher interest rates outweighs the losses from lower wages only when savings are sufficiently higher, which appears to occur only in this case.

We can draw some lessons from these variations of our main exercise. First, a better understanding of how consumption taxes affect warm-glow motives is important for the assessment of welfare gains. Yet, even in the worst case scenario explored here, welfare gains are very large. Second, and also relevant, if consumption taxes are restricted to be non-negative then labor income taxes lead to substantially lower welfare gains. This effect is mostly mechanic, lower revenue requirements restrict the scope for progressivity when consumption taxes cannot be negative. Finally, the fact that the *level* of consumption taxes affect savings indicate that the warm-glow motive plays a role in our result.

5 Progressivity Through Non-uniform Commodity Taxation

Our findings suggest very significant welfare gains from using progressive consumption taxes. Of course, for consumption taxes to be progressive in the sense that we are implying here one must be able to verify who consumed. While we postpone until the next section a more detailed discussion about implementation, it is important not to confuse our proposal with [Mill's \(1844\)](#) idea of exempting (or subsidizing) necessities to create progressivity in consumption taxes. Indeed, recent works by [Conesa et al. \(2020a,b\)](#) introduce progressivity in consumption taxation by differentiated commodity taxes.¹⁹ Most of their findings suggest welfare losses from replacing the current system with one in which progressivity is attained by differentiated taxation. As we shall recall, in theory, this is to be expected. Differentiated taxation is a very inefficient way of creating progressivity.

The idea of using differentiated commodity taxation for distributive purposes is formalized in [Diamond and Mirrlees' \(1971\)](#) seminal work. Essentially, they derive an expression for optimal taxation that generalizes [Ramsey's](#) formulae to account for agent heterogeneity. While pure efficiency concerns prescribe that the consumption of all goods should be discouraged proportionately by the tax system, in the presence of heterogeneity, discouragement relative to the efficient benchmark should be smaller for commodities which consumption is positively correlated with the marginal social value of income.

It is important, however, to note that if efficiency and equity considerations move in different directions, then the commodity taxes become a poor instrument for mimicking progressivity. This is exactly the case in [Conesa et al. \(2020a,b\)](#). Indeed, [Corlett and Hague \(1953\)](#) show that, in a three commodities world – two consumption goods and leisure –, the [Ramsey](#) proportional discouragement principle is equivalent to the prescription of taxing more heavily the good which is more complementary with the untaxed good, leisure. Under separability, assumed by [Conesa et al. \(2020a,b\)](#), the cross compensated effect between goods of different groups is proportional to income effects.²⁰ This leads to the prescription that necessities should be more heavily taxed, which is exactly the opposite of what equity would prescribe. Efficiency would, therefore, be sacrificed even if taxation was uniform, let alone with subsidization of necessities.²¹

$$U_t(c_t, l_t) = \frac{(c_t^{1-\rho_t} l_t^{\rho_t})^{1-\gamma} - 1}{1-\gamma}, \quad (16)$$

¹⁹[Gordon \(1989\)](#) discusses the use of commodity taxes to mimic poll taxation. This is *not* the rationale in [Conesa et al. \(2020b\)](#) where a universal basic income is available. [Conesa et al. \(2020b\)](#) show that a UBI that is 10 times as large as the current level of targeted transfers can generate a 5.96% welfare gain, if financed with differentiated consumption taxes.

²⁰See [Deaton and Muellbauer \(1980\)](#).

²¹Even when preferences are not separable, [Sah \(1983\)](#) has provided limits to how much redistribution is possible through commodity taxes. In simple parametrizations little redistribution is possible.

where $c_t = (c_{1,t} - \underline{c})^\xi c_{2,t}^{1-\xi}$.

Welfare gains are not negligible as one can see in Table 9 – 2.52% when labor income taxes are eliminated and 2.66% when they are kept in place. For the case in which labor income taxes are still used, welfare gains are attained despite a small decline, almost 2%, in the capital stock, which suggest that these gains may survive the transition between steady states. As for the case without labor income taxes, the welfare gains occur with an increase of more 20% in the capital stock.

We can draw two lessons from these exercises. First, that differentiated taxes are very useful to complement labor income taxes but are a very poor substitute for it. Second that implementing progressivity in consumption taxes with differentiated taxation leads to much lower gains than it is possible with an explicit progressive tax. In the next section we digress about the possible difficulties as well as omitted potential benefits from moving to progressive consumption taxes.

6 Digression: Implementation Concerns

Our quantitative assessments show substantial welfare gains from a fundamental tax reform that moves progressivity to consumption taxes. The goal of this section is to raise some of the practical difficulties that must be addressed for actual implementation.

Interpersonal tax arbitrage A concern regarding progressive consumption taxes is that agents may engage in tax arbitrage by transferring expenditures from agents with high expenditures to agents with low expenditures.

To simplify our discussion let us focus on a system where all expenditures up to a certain value, say US\$12,000 a year, is exempt and all expenditures above this value are taxed at a 20% rate. Then, an agent who is planning on spending more than US\$12,000, can offer an agent who is planning on spending say US\$11,000 to buy US\$1,000 in her name and pocket US\$100. In fact, the agent might be induced to reduce his consumption to benefit from the arbitrage. Indeed, assume that the agent earns US\$12,000 for two years and would optimally save zero. Let us say that another person who is spending well above US\$12,000 proposes this agent to reduce his consumption by US\$ 2,000 in year one in exchange for a return R on this value to be paid in year two. Note that the agent who was not paying taxes will now pay $.2 \times 2,000 \times R$ in taxes. Hence, just to keep the net present value constant an excess return of $R = .25$ must be paid.

These transactions are, of course possible, but they entail potentially large costs, particularly since these agreements will not be enforced by law.

International Arbitrage Another possible concern is that an agent who did not pay income taxes would leave the country and spend his income abroad to avoid paying taxes. As we have mentioned, linear income taxes could still be used and heavily relied on. What is relevant is that progressivity applies to consumption taxes not income taxes.

Still, average consumption taxes should not be lower than current ones.²² In this case, two alternative ways for reducing the risk of international arbitrage would be either to tax all legal international money transfers could be taxed at the baseline consumption tax rate, or taxing all withdrawals, or transfers to other persons' accounts, from the international accounts. In either case, information on these transactions should be used to determine the exemption to which the agent would be entitled.

Period Length As we have seen, the length of the period is crucial. At the limit, if we allow the period length be the whole life, we are in a [Vickrey \(1947\)](#) world in which whether it is progressive income or progressive consumption taxes is irrelevant. We assume that progressivity applies to yearly consumption to keep it comparable to the current rules that apply to income taxes.

Durables Related to the length of a period is the treatment of durable goods. A special provision for goods whose half life, according to the depreciation allowance rules, exceed a year could be easily incorporated to the system. Consumption of this good for taxable purposes would be spread through time.

Privacy Concerns Finally, for an agent to be entitled to the rebates that generate progressivity they would have to inform the tax authorities about their consumption choices. This raises concerns about privacy. While the tax authority might be required to destroy this information as soon as the agent gets his or her refund, there is still a chance that this information may be misused. Agents whose expenditures exceed the threshold for rebate may choose which ones are reported, and may also forgo some money to retain their privacy. Still this is not a complete solution and this raises many of the concerns that have already been voiced regarding private companies.

Informality Although informality is usually not a top concern in developed economies, it cannot be ignored. More so if one thinks about implementation in less developed countries.

²²Lowering consumption taxes would lead to a distribution from the poor to the wealthy. Although we do not discuss transition here, such restriction on consumption taxes should be imposed in any serious discussion of the topic.

Informality is mostly concentrated in low quality goods consumed by the poorest agents.²³ A progressive consumption tax reduces the gains from consuming goods produced in the informal sector exactly by the agents who currently do so. This suggests a force in the opposing direction from the ones above; taking informality into account *increases* the appeal of progressive consumption taxes.

7 Conclusion

Most tax systems generate progressivity through non-linearities in labor earnings taxes. This type of progressivity discourages effort at moments of high productivity on a persons' life-cycle creating a severe misallocation of effort. If progressivity is, instead, imposed on consumption taxes this inefficiency is eliminated. This alternative system creates, in contrast, more consumption smoothing than it would be optimal without these taxes. In a very stylized model with no uncertainty and perfect capital markets we show that progressive consumption taxes are strictly better than progressive income taxes when productivity varies along the life-cycle.

When we depart from the stylized economy and consider an economy calibrated to match the main statistics for the U.S. economy we find that the theoretical underpinnings are translated into huge welfare gains. Our baseline exercise produce welfare gains is steady-state comparisons that are two to three times larger than what [Costinot and Rodríguez-Clare \(2018\)](#) find for all gains from international trade, for example. These gains are attained with either very small increases or large declines in the capital stock, depending on the specification for preferences.

Progressivity in consumption taxes are not a feature of tax systems around the world. Practical concerns that are abstracted in optimal tax theory, in general – see [Slemrod and Gillitzer \(2014\)](#) –, and in our model, in particular, may undermine these taxes. The sheer size of welfare gains found in our study suggest that we should spend more resources on trying to carefully address these concerns. All in all, we believe that fundamental tax reforms that entail consumption progressivity are among the most promising ones.

References

- S Rao Aiyagari. Uninsured idiosyncratic risk and aggregate saving. *Quarterly Journal of Economics*, 109(3):659–684, 1994. [2.3](#)
- James Andreoni. Giving with impure altruism: Applications to charity and ricardian equivalence,. *Journal of Political Economy*, 97(6):1447–1458, 1989. [\(document\)](#), [2.3](#)

²³See [Bachas et al. \(2020\)](#).

- James Andreoni. Impure altruism and donations to public goods: A theory of warm-glow giving. *Economic Journal*, 100:464–477, 1990. ([document](#))
- Pierre Bachas, Lucie Gadenne, and Anders Jensen. Informality, consumption taxes and redistribution. NBER Working Paper 27429, National Bureau of Economic Research, June 2020. [23](#)
- Felicite Bell and Michael Miller. Life tables for the united states social security area 1900–2100. Actuarial Actuarial Study 120, Social Security Administration, 2005. [2.6](#)
- Roland Bénabou. Unequal societies: Income distribution and the social contract. *American Economic Review*, 90(1):96–129, 2000. [2.4](#)
- Roland Bénabou. Tax and education policy in a heterogeneous-agent economy: What levels of redistribution maximize growth and efficiency? *Econometrica*, 70(2):481 – 517, 2002. [2.4](#)
- David F. Bradford. A tax system for the twenty-first century. In Alan J. Auerbach and Kevin A. Hassett, editors, *Toward Fundamental Tax Reform*. American Enterprise Institute, 2005. ([document](#))
- Juan Carlos Conesa, Sigiri Kitao, and Dirk Krueger. Taxing capital? not a bad idea after all. *American Economic Review*, 99(1):25–48, 2009. [1](#), [2.3](#)
- Juan Carlos Conesa, Bo Li, and Quian Li. Welfare implications of switching to consumption taxation. *Journal of Economic Dynamics and Control*, (forthcoming), 2020a. [1](#), [3](#), ([document](#)), [5](#)
- Juan Carlos Conesa, Bo Li, and Quian Li. Universal basic income and progressive consumption taxes. Mimeo. Stony Brook, January 2020b. [3](#), ([document](#)), [5](#), [19](#)
- W.J. Corlett and D. C. Hague. Complementarity and the excess burden of taxation. *Review of Economic Studies*, 21:21–30, 1953. ([document](#)), [5](#)
- Isabel Correia. Consumption taxes and redistribution. *American Economic Review*, 100(4): 1673–1694, 2010. ([document](#))
- Arnaud Costinot and Andrés Rodríguez-Clare. The us gains from trade: Valuation using the demand for foreign factor services. *Journal of Economic Perspectives*, 32(2), 2018. [7](#)
- Mariacristina De Nardi. Wealth inequality and intergenerational links. *Review of Economic Studies*, 71(3):743 – 768, 2004. ([document](#)), [2.3](#), [10](#)
- Mariacristina De Nardi and Fang Yang. Bequests and heterogeneity in retirement wealth. *European Economic Review*, 72:182 – 196, 2014. [2.6](#)
- Angus S. Deaton and John Muellbauer. *Economics and Consumer Behaviour*. Cambridge University Press, 1980. [20](#)
- Peter A. Diamond. Optimal tax treatment of private contributions for public goods with and without warm glow preferences. *Journal of Public Economics*, 90:897–919, 2006. ([document](#)), [3](#), [17](#)

- Peter A. Diamond and James A. Mirrlees. Optimal taxation and public production ii: Tax rules. *American Economic Review*, 61(3):261–278, 1971. [5](#)
- Martin Feldstein. The effects of taxation on risk taking. *Journal of Political Economy*, 77(5):755 – 764, 1969. [8](#), [2.4](#)
- L. Fuster, A. Imrohoroglu, and S. Imrohoroglu. Elimination of social security in a dynastic framework. *Review of Economic Studies*, 74:134–45, 2007. [12](#), [1](#)
- James P.F. Gordon. A many good, many person, corlett-hague tax rule. *Economics Letters*, 29:157 – 161, 1989. [19](#)
- Robert E. Hall and Alvin Rabushka. *The Flat Tax*. Hoover Institution Press, 2007. ([document](#))
- Jonathan Heathcote, Kjetil Storesletten, and Ginaluca Violante. Optimal tax progressivity: An analytical framework. *Quarterly Journal of Economics*, 132(4):1693 – 1754, 2017a. [mimeo](#). [2.4](#)
- Jonathan Heathcote, Kjetil Storesletten, and Giovanni Violante. Optimal progressivity with age-dependent taxation. [mimeo](#), July 2017b. ([document](#)), [1](#)
- Nicholas Kaldor. *The Expenditure Tax*. George Allen and Unwin, 1956. ([document](#))
- Marek Kapička. Quantifying the welfare gains from history dependent income taxation. [mimeo](#). CERGE-EI, Prague, January 2020. [1](#), ([document](#)), [18](#)
- Greg Kaplan. Inequality and the lifecycle. *Quantitative Economics*, 3(11-014):471–525, 2012. [2.6](#), [1](#)
- Wojciech Kopczuk. Taxation of intergenerational transfers and wealth,. In Alan J. Auerbach, Raj Chetty, Martin S. Feldstein, and Emmanuel Saez, editors, *Handbook of Public Economics*, volume 5, pages 329 – 390. Elsevier Science Publishers B.V. (North-Holland), 2013. [4](#), ([document](#)), [4](#), [17](#), [4.2](#)
- John Stewart Mill. *Essays on Some Unsettled Questions of Political Economy*. London: Longmans, Green, Reader, and Dyer., 1844. [5](#)
- Richard A Musgrave. *The Theory of Public Finance*. New York, 1959. [8](#), [2.4](#)
- Frank Ramsey. A contribution to the theory of taxation. *Economic Journal*, 37:47–61, 1927. [5](#)
- Raaj K. Sah. How much redistribution is possible through commodity taxes? is possible taxes? *Journal of Public Economics*, 20:89 – 101, 1983. [21](#)
- Joel Slemrod and Christian Gillitzer. Insights from a tax-systems perspective. *CESifo Economic Studies*, 60(1):1 – 30, 2014. ([document](#)), [7](#)
- Ludwig Straub. Consumption, savings, and the distribution of permanent income. [Mimeo](#) Harvard, June 2019. ([document](#)), [2](#)
- William S. Vickrey. *Agenda for Progressive Taxation*. The Ronald Press Company, 1947. ([document](#)), [1](#), [1](#), [6](#)

Max Weber. *The protestant ethic and the spirit of capitalism*. Scribner, New York, 1958. [2.3](#)

A Tables and Figures

Table 1: Parameter Values - Baseline Calibration

Parameter	Value	Source/Target
β	0.99	$K/Y = 3$
γ	4.00	Micro evidence
ρ_t	-	life cycle profile of mean hours
σ_u^2	0.06	Gini index of 0.43
φ_z	0.94	Kaplan (2012)
σ_ϵ^2	0.016	Kaplan (2012)
δ	0.06	see text
α	0.36	NIPA
B	0.90	$w = 1$
τ_c	0.05	Fuster et al. (2007)
τ_k	0.30	Fuster et al. (2007)
τ_{ss}	0.062	Social security program
ξ	0.83	Balanced budget
ϱ	0.09	Average tax rates
η_1	-24.5	Wealth distribution
η_2	50.6	Wealth distribution

Results

Variable	Benchmark	Benchmark w/ cons. taxes	Consumption		Labor	
			Positive	Negative	Positive	Negative
ζ	0.831	1.000	1.000	1.000	0.806	0.728
ϱ	0.090	0.000	0.000	0.000	0.130	0.220
ϱ_c	0.000	0.099	0.350	0.500	0.000	0.000
τ_c	0.060	0.294	0.690	1.128	0.060	0.060
Y	0.661	0.707	0.647	0.596	0.642	0.587
K/Y	3.028	3.333	3.405	3.429	2.998	2.927
Average Hours	0.394	0.395	0.367	0.322	0.389	0.360
labor/hours	1.250	1.266	1.231	1.286	1.237	1.238
w	1.012	1.068	1.081	1.086	1.006	0.993
r	5.88%	4.80%	4.57%	4.59%	6.00%	6.20%
Revenue distribution						
Labor income	39.72%	0%	0%	0%%	40.76%	44.29%
Consumption	16.60%	56.80%	60.36%	63.28%	16.24%	15.03%
Capital	26.33%	24.90%	22.65%	20.90%	26.00%	24.79%
Gini						
Income Before Taxes	0.555	0.568	0.561	0.579	0.427	0.449
Income After Taxes	0.418	0.454	0.441	0.459	0.386	0.351
Consumption	0.324	0.342	0.279	0.261	0.333	0.301
Assets	0.684	0.718	0.704	0.715	0.664	0.612
Insurance coefficient						
Low	0.376	0.356	0.358	0.512	0.376	0.465
all	0.349	0.311	0.445	0.474	0.390	0.458
High	0.360	0.325	0.439	0.488	0.400	0.480
Welfare						
CEV	—	5.27%	9.04%	13.34%	0.31%	2.51%
CEV low type	—	7.72%	8.56%	40.34%	-0.53%	5.26%
CEV high type	—	3.81%	-5.19%	-12.70%	-2.30%	-9.65%
CEV uncertainty	—	-0.69%	6.28%	11.10%	1.70%	6.49%
CEV efficiency	—	5.97%	2.74%	2.23%	-1.41%	-4.02%

Table 2: **Current and Optimal Tax System:** The first column describes the benchmark economy for the statistics explained in the text. The second, relates to when the labor income progressivity in the benchmark is instead applied to consumption taxes. The other columns describe the optimum for consumption and labor income taxes. Positive denotes taxes that are restricted to be positive and negative, unrestricted taxes.

Wealth distribution					
	Bottom 20%	20-40%	40-60%	60-80%	Top 20%
Benchmark	0.43	3.09	7.92	17.47	71.07
Opt. cons. positive	0.23	2.25	7.21	17.00	73.10
Opt. cons. negative	0.14	2.02	6.66	17.13	74.02
Opt. labor positive	0.58	3.57	8.68	18.26	68.91
Opt. labor negative	1.06	4.75	10.54	20.15	63.47

Table 3: **Wealth distribution** The table describes the cross-sectional wealth distributions for the benchmark and each optimized tax system.

Benchmark					
	Bottom 20%	20-40%	40-60%	60-80%	Top 20%
Consumption taxes	10.82	13.51	15.03	21.27	39.36
Labor taxes	0.54	3.74	9.11	19.30	67.29
Capital taxes	2.80	8.74	10.94	20.17	57.34
Total	3.32	7.30	10.88	19.97	58.52
Optimum consumption negative					
	Bottom 20%	20-40%	40-60%	60-80%	Top 20%
Consumption taxes	-1.59	6.30	7.42	19.37	68.51
Labor taxes	0.00	0.00	0.00	0.00	0.00
Capital taxes	1.79	9.07	10.94	20.45	57.73
Total	-0.56	7.14	8.49	19.70	65.22
Optimum consumption positive					
	Bottom 20%	20-40%	40-60%	60-80%	Top 20%
Consumption taxes	2.34	8.77	6.82	18.18	63.88
Labor taxes	0.00	0.00	0.00	0.00	0.00
Capital taxes	2.64	9.78	10.04	20.27	57.25
Total	2.19	9.05	7.70	18.72	62.34
Optimum labor income negative					
	Bottom 20%	20-40%	40-60%	60-80%	Top 20%
Consumption taxes	12.16	14.87	16.00	21.19	35.75
Labor taxes	-4.79	-1.17	5.37	18.52	82.07
Capital taxes	4.08	11.12	13.19	21.35	50.25
Total	1.31	5.82	9.90	19.93	63.05
Optimum labor income positive					
	Bottom 20%	20-40%	40-60%	60-80%	Top 20%
Consumption taxes	11.21	14.09	15.21	21.28	38.21
Labor taxes	0	1.62	7.24	18.39	72.75
Capital taxes	3.16	9.48	11.41	20.60	55.33
Total	2.93	6.16	9.89	19.56	61.45

Table 4: **Distribution of taxes paid** The table displays the fraction of taxes paid by each quantile of labor income distribution for the benchmark and the optimized tax systems. The statistics are provided for each tax base as well as total taxes.

Results						
Variable	Benchmark			Optimal		
	Income	Positive	Negative	Positive	Negative	Negative*
ζ	0.831	0.869	0.867	0.919	0.919	0.831
ϱ	0.090	0.000	0.000	0.000	0.000	0.000
ϱ_c	0.000	0.099	0.099	0.250	0.450	0.400
τ_c	0.060	0.135	0.145	0.377	0.852	0.599
Y	0.661	0.679	0.671	0.659	0.589	0.579
K/Y	3.028	3.119	3.115	3.241	3.293	3.138
Average Hours	0.394	0.399	0.388	0.385	0.325	0.330
labor/hours	1.250	1.249	1.268	1.228	1.287	1.283
w	1.012	1.029	1.028	1.052	1.061	1.033
r	5.88%	5.54%	5.56%	5.11%	4.93%	5.47%
Revenue distribution						
Labor income	39.72%	39.66%	40.13%	23.95%	11.88%	45.21%
Consumption	16.60%	16.60%	16.60%	34.33%	40.72%	16.60%
Capital	26.33%	26.06%	25.85%	24.39%	21.84%	22.79%
Gini						
Income Before Taxes	0.555	0.562	0.569	0.556	0.578	0.576
Income After Taxes	0.418	0.442	0.450	0.435	0.455	0.450
Consumption	0.324	0.342	0.344	0.302	0.270	0.279
Assets	0.684	0.715	0.719	0.707	0.713	0.712
Insurance coefficient						
Low	0.376	0.332	0.396	0.328	0.514	0.528
all	0.349	0.320	0.322	0.389	0.463	0.454
High	0.360	0.336	0.335	0.407	0.477	0.466
Welfare						
CEV	—	-0.01%	2.15%	4.10%	11.48%	8.62%
CEV low type	—	4.69%	12.45%	5.19%	37.34%	36.31%
CEV high type	—	1.71%	0.68%	-1.92%	-12.27%	-12.34%
CEV uncertainty	—	-1.21%	-0.14%	2.92%	10.09%	9.12%
CEV efficiency	—	1.20%	2.30%	1.47%	1.37%	-0.51%

Table 5: **Current and Optimal Tax System: restricted consumption tax revenues.** The first three columns display the consequences of applying the progressivity parameter from labor income taxes to consumption taxes while retaining revenues from consumption taxes at the benchmark level. Columns 4 to 6 describe optimal progressive consumption taxes for three different types of restrictions. In columns 4 e 5 average income tax rates are required to remain at the same level as the benchmark. In column 6, revenues from consumption taxes are required to remain at the benchmark level.

Results			
Variable	Benchmark	Positive	Negative
ζ	0.831	1.000	1.000
φ	0.090	0.000	0.000
ϱ_c	0.000	0.300	0.450
τ_c	0.060	0.600	0.994
Y	0.661	0.646	0.581
K/Y	3.028	3.333	3.295
Average Hours	0.394	0.371	0.329
labor/hours	1.250	1.233	1.257
w	1.012	1.069	1.062
r	5.88%	4.80%	4.92%
Revenue distribution			
Labor income	39.72%	0%	0% ⁰
Consumption	16.60%	59.72%	60.04%
Capital	26.33%	23.25%	21.57%
Gini			
Income Before Taxes	0.555	0.561	0.573
Income After Taxes	0.418	0.440	0.451
Consumption	0.324	0.292	0.265
Assets	0.684	0.690	0.687
Insurance coefficient			
Low	0.376	0.358	0.505
all	0.349	0.424	0.474
High	0.360	0.418	0.517
Welfare			
CEV	—	8.13%	12.12%
CEV low	—	7.67%	36.18%
CEV high	—	-3.54%	-10.88%
CEV uncertainty	—	5.17%	9.99%
CEV efficiency	—	2.93%	2.11%

Table 6: **Current and Optimal Tax System: Adjusted Preferences** The benchmark economy is calibrated to the preferences described in (15) and the findings are reported in the first column. The other two columns refer to the optimal progressive consumption taxes.

Results

Variable	Constrained	beq. (pos.)	beq.	beq. + assets	beq. + assets +labor sup
ζ		1.000	1.000	1.000	1.000
ϑ		0.000	0.000	0.000	0.000
ϑ_c		0.500	0.500	0.500	0.500
τ_c		1.131	1.109	1.150	0.820
Y		0.614	0.598	0.591	0.661
K/Y		3.544	3.412	3.390	3.029
Average Hours		0.329	0.326	0.327	0.394
labor/hours		1.274	1.279	1.265	1.250
w		1.106	1.082	1.078	1.012
r		4.16%	4.55%	4.62%	5.89%
Revenue distribution					
Labor income		0%	0%	0%	0%
Consumption		63.21%	62.90%	62.89%	56.37%
Capital		20.56%	21.20%	21.13%	26.32%
Gini					
Income Before Taxes		0.556	0.577	0.561	0.573
Income After Taxes		0.451	0.457	0.447	0.450
Consumption		0.247	0.262	0.249	0.297
Assets		0.696	0.712	0.683	0.714
Insurance coefficient					
Low		0.504	0.510	0.519	0.529
all		0.456	0.477	0.453	0.457
High		0.458	0.485	0.434	0.417
Welfare					
CEV		13.62%	13.75%	11.69%	18.80%
CEV low		40.09%	39.50%	35.89%	44.33%
CEV high		-13.24%	-12.21%	-14.05%	-7.66%
CEV uncertainty		9.12%	11.34%	11.67%	12.10%
CEV efficiency		1.48%	2.38%	0.00%	6.68%

Table 7: **Optimal Tax System: Decomposition** This table describes the optimal policies when agents choices are restricted. In the first two columns agents policy function for asset accumulation are borrowed from the benchmark instead of calculated within the system. Adjustments in asset accumulation are allowed (in the first column only positive adjustments) but the warm glow motive is eliminated from preferences to determine the adjustment. In column 3, nos asset adjustments are allowed. Finally, in column 4, neither labor supply adjustments are allowed.

Results

Variable	Benchmark	Consumption		Labor	
		Positive	Negative	Positive	Negative
ζ	0.831	1.000	1.000	0.805	0.725
ϱ	0.090	0.000	0.000	0.130	0.220
ϱ_c	0.000	0.350	0.500	0.000	0.000
τ_c	0.060	0.590	1.008	0.060	0.060
Y	0.661	0.686	0.631	0.638	0.573
K/Y	3.028	3.826	3.983	2.954	2.771
Average Hours	0.394	0.369	0.315	0.389	0.363
labor/hours	1.250	1.215	1.282	1.238	1.239
avg productivity	1.000	1.000	1.000	1.000	1.000
w	1.012	1.012	1.012	1.012	1.012
r	5.88%	5.88%	5.88%	5.88%	5.88%
Revenue distribution					
Labor income	39.72%	0%	0%	41.78%	47.48%
Consumption	16.60%	48.63%	51.02%	16.21%	14.96%
Capital	26.33%	35.25%	34.43%	24.91%	21.33%
Gini					
Income Before Taxes	0.555	0.534	0.553	0.550	0.549
Income After Taxes	0.418	0.440	0.463	0.401	0.366
Consumption	0.324	0.283	0.266	0.333	0.232
Assets	0.684	0.673	0.674	0.667	0.625
Insurance coefficient					
Low	0.376	0.348	0.525	0.378	0.458
all	0.349	0.449	0.493	0.388	0.453
High	0.360	0.453	0.500	0.398	0.475
Welfare					
CEV	—	7.43%	14.49%	0.43%	2.86%
CEV low type	—	3.36%	39.71%	-0.64%	6.02%
CEV high type	—	-3.21%	-11.26%	-2.34%	-9.53%
CEV uncertainty	—	5.48%	19.24%	1.68%	6.31%
CEV efficiency	—	1.94%	2.76%	-1.25%	-3.43%

Table 8: **Current and Optimal Tax System: Partial Equilibrium** In all the exercises prices (r and w) are held fixed.

Results

Variable	Benchmark	with labor tax	without labor tax
ζ	0.834	0.834	1.000
φ	0.090	0.090	0.000
τ_{c_1}	0.060	-0.189	-0.122
τ_{c_2}	0.060	0.22	0.380
Y	0.712	0.699	0.785
K/Y	3.040	3.033	3.346
Average Hours	0.442	0.431	0.453
labor/hours	1.198	1.208	1.221
avg productivity	1.000	1.000	1.000
w	1.014	1.013	1.071
r	5.84%	5.87%	4.76%
Revenue distribution			
Labor income	39.54%	38.76%	0%
Consumption	16.67%	18.05%	55.80%
Capital	26.20%	25.25%	25.37%
Gini			
Income Before Taxes	0.547	0.550	0.563
Income After Taxes	0.400	0.403	0.440
Consumption	0.325	0.320	0.344
Assets	0.668	0.669	0.708
Insurance coefficient			
Low	0.467	0.486	0.418
all	0.416	0.428	0.344
High	0.408	0.408	0.317
Welfare			
CEV	—	2.66%	2.52%
CEV low	—	16.39%	12.83%
CEV high	—	-3.11%	4.88%
CEV uncertainty	—	5.69%	-3.16%
CEV efficiency	—	-3.05%	5.68%

Table 9: **Model with two goods: Current and Optimal Tax System** The first column displays the relevant statistics for the two good economy calibrated to the U.S. current (benchmark) tax system. The last two columns display optimal policies. For the first, labor taxes are kept at the current level. For the second, labor income taxes are eliminated.

Figure 1: Tax schedules

Figure 2: Tax schedules

Figure 3: Consumption function.

Os artigos publicados são de inteira responsabilidade de seus autores. As opiniões neles emitidas não exprimem, necessariamente, o ponto de vista da Fundação Getulio Vargas.

EPGE Escola Brasileira de Economia e Finanças

Diretor Geral: Rubens Penha Cysne

Vice-Diretor: Aloisio Araujo

Coordenador de Graduação: André Arruda Villela

Coordenadores de Pós-graduação Acadêmica: Humberto Moreira & Lucas Jóver Maestri

Coordenadores do Mestrado Profissional em Economia e Finanças: Ricardo de Oliveira Cavalcanti & Joísa Campanher Dutra