I Z A Institute
of Labor Economics
Initiated by Deutsche Post Foundation

DISCUSSION PAPER SERIES

IZA DP No. 10508

Time-Poor, Working, Super-Rich

Giacomo Corneo

JANUARY 2017

I Z A Institute
of Labor Economics

Initiated by Deutsche Post Foundation

DISCUSSION PAPER SERIES

IZA DP No. 10508

Time-Poor, Working, Super-Rich

Giacomo Corneo
Freie Universität Berlin and IZA

JANUARY 2017

IZA – Institute of Labor Economics

Schaumburg-Lippe-Straße 5–9
53113 Bonn, Germany

Phone: +49-228-3894-0
Email: publications@iza.org

www.iza.org

ABSTRACT

Time-Poor, Working, Super-Rich*

This paper revisits the standard model of labor supply under two additional assumptions: consumption requires time and some limited amount of work is enjoyable. Whereas introducing each assumption without the other one does not produce novel insights, combining them together does if the wage rate is sufficiently high. For top earners, work has a positive marginal utility at the optimum and above a critical wage level it converts into a pure consumption good. Their labor-supply curve is first backward bending and then vertical. This can justify an optimal marginal tax rate on top incomes equal to 100 percent. Top earners in the vertical half-line of the labor-supply curve optimally refrain from spending their entire income. At the macroeconomic level, this can generate a lack of effective demand. With some qualifications, these findings carry over to models that include savings and philanthropy.

JEL Classification: J22, H21, H24

Keywords: super-rich, labor supply, time allocation, effective demand, optimal taxation of top labor incomes

Corresponding author:
Giacomo Corneo
Department of Economics
Freie Universität Berlin
Boltzmannstr. 20
14195 Berlin
Germany
E-mail: giacomo.corneo@fu-berlin.de

* Helpful remarks by Felix Bierbrauer, Hans Brunner, Johannes König, Thomas Piketty, Wolfram Richter, John Roemer, and Harald Uhlig are gratefully acknowledged.

1 Introduction

In the U.S., the richest 0.01 percent households account for about five percent of total household income and eleven percent of total household wealth.[1] These super-rich with yearly incomes in excess of $10m represent a tiny fraction of the population but a significant share of the economy. This fact has spurred debates on the taxation of top incomes and the link between rising income concentration and the occurrence of macroeconomic crises. Despite the growing economic significance of the super-rich in large parts of the globe, not much has been done to develop economic models especially tailored to analyze their behavior. In particular, the standard neoclassical model of labor supply is still at the center of both public-finance analyses of optimal taxation and DSGE analyses of macroeconomic policies with heterogeneous agents. The aim of this paper is to propose two simple modifications of the standard labor supply model that make it better suited to portray the labor supply of the super-rich, to explore their main implications, and to point out the new insights offered by such a model.

Most economists would likely subscribe to the following two statements: (i) some work enhances well-being; (ii) consuming goods require time. Yet, both (i) and (ii) are neglected by the standard model of labor supply. Since the key properties of labor supply are deemed to be independent of (i) and (ii), by Ockham's razor they are dropped. I am going to show that for agents with very high wages - the working super-rich - this is unwarranted: combining assumptions (i) and (ii) generates insights that profoundly differ from those delivered by the standard model, with noticeable implications for tax policy and in the realm of macroeconomic management.

Assumption (i) - some work enhances well-being - has received an enormous empirical support from studies in psychology, sociology, economics and management science.[2] For one thing, work is a crucial source of identity and social relationships. Most people maintain that laziness is wicked, many believe that work is a duty towards society, and some think that hard working is virtuous. For another, individuals in control of their work secure some gratification from efficiently performing their work activities. Mastery of a working task is for those individuals a source of pleasure, pride, and personal fulfillment.

[1]See Saez (2015) and Saez and Zucman (2016). Atkinson and Piketty (2010) offer a historical perspective on the economic role of top incomes. In 2014, according to Forbes Magazine, the twenty-five highest-earning hedge funds managers and traders made on average $500m in personal income.

[2]In psychology, see e.g. Deci and Ryan (1980) and Csikszentmihalyi and LeFevre (1989). An extensive discussion of the sociological literature is offered by Baron (1988). Kreps (1997) gives an insightful account from an economist's perspective.

A taste for efficacy is likely to have been selected by nature because of its survival value. The presumption that some work enhances well-being suits especially well the working super-rich because they do not have to work, enjoy a great latitude in choosing the type of working activity they perform, and are compelled to show that they deserve their riches. Therefore, I will make the assumption that labor effort, up to some level, increases utility.[3] It is well known that by itself this is an innocuous modification of the standard neoclassical model of labor supply. Its key properties depend on the marginal rate of substitution between consumption and leisure and this will be positive at the individual optimum even if inframarginal units of labor add to the agent's utility up to some point. However, this may cease to hold if the time-consuming nature of consumption is simultaneously taken into account.

Assumption (ii) - consumption requires time - made its inroad in economics mainly through Becker (1965) who put forward the time opportunity costs of consumption activities on top of their market prices. By way of an example, the total economic cost of enjoying a movie at the cinema may amount to the money disbursed for the ticket plus the value of the moviegoer's time. As shown by Kleven (2004), taking into account that different consumption goods carry different time coefficients has implications for optimal commodity taxation: commodities that require more consumption time should be taxed more heavily. The literature on optimal income taxation often proceeds with the tacit assumption that the agents' time endowments are infinite; otherwise, assumptions on preferences are made that guarantee an interior solution. In reality, every individual has a finite total time available and, despite the rise in life expectancy, its long-run rate of growth is significantly lower than the long-run rate of growth of personal incomes. Hence, the labor supply model I propose will capture the notions that the time required for consumption has an opportunity cost and that total consumption time cannot exceed the time endowment of the individual. Arguably, this restriction is only relevant for the super-rich, whose time endowments are not a large multiple of the time endowments of the bulk of the population, but whose incomes are.

Incorporating assumptions (i) and (ii) in an otherwise standard model of labor supply yields the following new insights. First, agents with very high wages optimally stop working at a point where their marginal utility of work is still positive. Those agents forgo enjoyable work in order to have more time for their consumption activities. Second, the

[3] This assumption plays a key role in very different contexts studied by Corneo and Rob (2003) and Funk (2015). In Corneo and Rob (2003) it is used in order to explain why public firms offer wage contracts with weaker incentives than their private counterparts. In Funk (2015) it is used in relation to human capital accumulation in order to show the possibility of a persistent division of society into two groups, the educated rich and the uneducated poor.

labor supply of such top earners is backward bending even in the absence of income effects. For them, the time constraint binds and the increased consumption which is made possible by a wage increase reduces the time available for work. Third, the optimal top marginal tax rate on labor income is 100 percent if the cutoff level of income for the top tax bracket is high enough. As the labor-supply curve of the super-rich is backward bending, a higher tax rate increases their labor supply, and thus the tax revenue that can be redistributed to the poor. Fourth, the super-rich may optimally leave some income unspent because they lack the time to consume it. They earn such an excessive income because they enjoy their working activities. At the macroeconomic level, this joy of working implies that aggregate demand falls short of aggregate supply. Since the insufficiency of effective demand increases with the income share of the super-rich, a rising income concentration can increase the risk of a macroeconomic crisis.

The quest for parsimony that invites theorists to neglect assumptions (i) and (ii) should therefore not be embraced too soon if one wants to analyze the labor supply of top earners and obtain policy recommendations. In the case of the super-rich there are no systematic and reliable empirical data on their work and consumption behavior. This makes a careful evaluation of the mechanisms put forward by the theoretical literature all the more relevant. Models that feature (i) and (ii) should thus complement models that highlight other aspects of reality in order to arrive at robust policy conclusions.[4]

The remainder of the paper is organized as follows. Section 2 incorporates the assumptions (i) and (ii) discussed above in the standard labor supply model and derives novel insights concerning agents with very high wages. Sections 3 and 4 extend the model of section 2 in order to assess the robustness of its insights to the inclusion of savings and private transfers, i.e. ways to allocate income that do not hinge on time availability to the same extent that consumption activities do. Section 3 studies a model of non-overlapping generations of super-rich linked by altruistic transfers; section 4 examines charitable giving in a warm-glow setting.

2 Labor supply with time scarcity and joy of work

2.1 Laissez faire

Define the super-rich as agents with a sufficiently high wage rate w, to be specified later; the analysis of the model for any $w > 0$ is relegated to Appendix A.1. Agents' utility

[4] For instance, Piketty et al. (2014) highlight responses to top marginal tax rates that occur through tax avoidance and compensation bargaining. Other aspects studied by the literature and neglected in the current paper include the role of occupational choice, innovation, winner-takes-all compensation, status seeking, and migration.

depends on consumption and labor, $U(c,l)$, where $c \geq 0$ is consumption and $l \geq 0$ is labor effort. Labor supply is assumed to reduce utility if and only if labor exceeds some strictly positive level $\widetilde{l} \equiv \arg\max U(c,l)$. Following findings in empirical psychology, one might interpret $\widetilde{l}$ as the level of labor effort such that the individual experiences boredom if $l < \widetilde{l}$ and anxiety if $l > \widetilde{l}$ (Csikszentmihalyi and LeFevre, 1989). In order to exhibit the implications of this assumption in a crystal-clear fashion, I assume away income effects and posit

$$U = c + \alpha l - \frac{\beta}{2} l^2, \tag{1}$$

so that $\widetilde{l} = \alpha/\beta$. As it will become clear, introducing income effects would just reinforce the results to be shown. The quadratic disutility from labor is only for the sake of simplicity and could be replaced with any concave function reaching its maximum at a strictly positive level of l.

Normalizing to one the price of consumption, the budget constraint of the individual reads:

$$c \leq wl. \tag{2}$$

Similarly to Becker (1965), consumption takes time according to a time coefficient θ that captures the required input of time per unit of consumption. Denoting by T the time endowment, the time constraint of the individual is:

$$\theta c + l \leq T. \tag{3}$$

Throughout the paper the following two restrictions on parameter values are posited: $T > \widetilde{l}$ and $\alpha > 1/\theta$. The first one is necessary in order for the marginal utility of labor to become negative; the second one is necessary in order for underconsumption to be a rational choice.

The linear way in which consumption enters (3) expedites the analysis by ensuring the existence of an interior solution, and could be relaxed. It can be seen as arising from a model with a large number of consumption activities performed at varying quality levels entailing different time requirements. Specifically, let $J \in \mathbb{N}^+$ denote the number of consumption activities and denote by c_j the quality of consumption activity j and by $T_j(c_j)$ the time spent on that activity, $j = 1, ...J$. The individual chooses which consumption activities to perform and at which quality level. As soon as an activity j is performed, a fixed amount of time is required. This fixed amount includes not only the time that is necessary to learn and perform that activity (e.g. the time to make a big catch in an off-shore fishing), but also the time to search for the goods necessary for the consumption activity, select them, and complete the corresponding transactions with suppliers. At any point in time, markets exist that supply standardized goods of varying quality that allow the individual to increase the quality of her consumption activity j up to a level $\bar{c}_j$. Increasing the quality of consumption beyond that level is possible but requires goods that are not supplied in ordinary markets and / or have to be assembled in an innovative way especially for the buyer. The design of such top-quality consumption goods demands an imagination effort on the side of the consumer in order to figure out what she wants and communicate and discuss her desires with specialized providers. Examples include super-yachts, mega-mansions, and art collections.[5]

For given unit price of qualities and assuming normality, increasing total consumption expenditure brings about an increasing quality in the consumption activities performed by the individual. As long as only markets for standardized goods are used, an increased consumption expenditure need not increase the time devoted to consumption because the individual can substitute time with better commodities. If total expenditure is large enough, all performed consumption activities require customized items in order to raise their quality beyond the level that is attainable in ordinary markets. The relationship between total consumption expenditure and consumption time may then be as the one depicted in Figure 1.

In that figure, $\bar{c}$ stands for the level of expenditure above which all consumption activities require customized goods that have to be designed for the consumer. At low levels of the consumption expenditure c the curve is rather flat: increasing expenditure has a negligible impact on the required amount of time because the individual mainly replaces goods of lower quality with ones of higher quality. Beyond some level of expenditure, additional activities may be performed, each one requiring an additional fixed amount of time to be learnt (e.g. playing golf, enjoying operas, hunting the fox). This first part of the curve in Fig. 1 - for expenditure levels well below $\bar{c}$ - may capture the time-consumption pattern that is typical for the overwhelming majority of individuals. At a substantially higher level of expenditure, increasing total expenditure comes along with new consumption styles that have to be invented because quality can only be raised by means of custom-made goods. It is at this point that "time-to-design" enters the picture. Raising total expenditure beyond $\bar{c}$ implies that the entire consumption bundle

[5] The documentary film "The Queen of Versailles" (http://www.bloomberg.com/news/articles/2012-03-15/versailles-the-would-be-biggest-house-in-america) gives some insights about the amount of time invested by a super-rich in order to specify and choose the distinctive features of his new residence, one of the largest single-family houses in the United States.

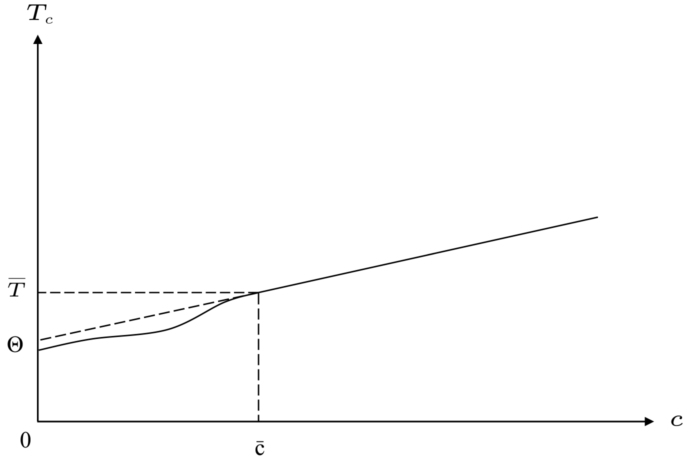

Figure 1: Total consumption time as a function of expenditure.

of the individual is tailor-made. Improving the quality of any consumption activity is only possible by adding new special features and gadgets that increase the individual's utility. Any improvement requires an additional imaginative effort and additional time to communicate the buyer's desires to the providers - or to the intermediaries in charge, including personal secretaries and household staff. In a first approximation, the increase in consumption time required by a marginal increase of consumption expenditure can thus be thought of as constant. Thus, for a super-rich with a consumption expenditure $c > \overline{c}$, total consumption time can be written as

$$T_c = \overline{T} + \theta(c - \overline{c}) = \Theta + \theta c,$$

where $\Theta \equiv \overline{T} - \theta\overline{c}$ is a constant. In this interpretation, θ is the asymptotic time coefficient of consumption once all existing possibilities of substitution of time has been exhausted. Redefining T in (3) as the total time available to the individual minus Θ yields the retained linear specification of the time constraint.

The problem of a super-rich is to maximize (1) subject to (2) and (3).[6] Let (l^*, c^*) denote the solution to that maximization problem. The following result describes the optimal labor supply of the super-rich as depending on their wage rate.

Proposition 1. *(i) There exists a wage rate* w_+ *such that* $\frac{\partial U(c^*,l^*)}{\partial l} > 0$ *for any* $w > w_+$*. (ii) There exists a wage rate* $\widehat{w} > w_+$ *such that* $c^* < wl^*$ *for any* $w > \widehat{w}$*.*

[6] In this problem, any utility from pure free time - leisure without consumption goods - is neglected. As shown in Appendix A.2, this is without any significant loss of generality.

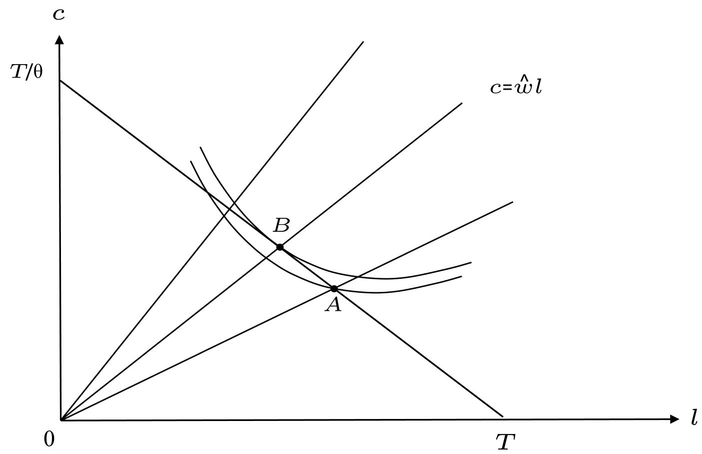

Figure 2: Optimum at three wage rates.

Proof.

Denote by l' the amount of labor such that constraints (2) and (3) are simultaneously binding, i.e.

$$l' = \frac{T}{1 + w\theta}. \tag{4}$$

In the (l, c)-space, the contour of the opportunity set is strictly increasing in l if $l < l'$ and strictly decreasing in l if $l > l'$ (see Fig. 2).

In order to prove (i), notice that $\frac{\partial U(c^*,l^*)}{\partial l} \leq 0$ requires $T \geq \widetilde{l}$ and $l' \geq l^* \geq \widetilde{l}$. Eq. (4) defines l' as a strictly decreasing function of w that goes to zero as w goes to infinity. Therefore, it exists a critical wage w_+ such that $l' \geq \widetilde{l}$ is violated for all $w > w_+$. For those wage levels, $\frac{\partial U(c^*,l^*)}{\partial l} > 0$ must hold.

In order to prove (ii), assume $w > w_+$ and consider Figure 2. Point A represents the optimum for some wage; at that point both constraints are binding and $l^* = l'$. Increasing the wage rate makes the budget line rotate anti-clockwise and allows the individual to reach higher indifference curves. Since $\alpha > 1/\theta$, there exists a wage rate $\widehat{w}$ such that point B is reached at which both constraints are binding and the time constraint is tangential to the highest indifference curve that can be reached by the individual. Further wage increases beyond $\widehat{w}$ leave the optimum unchanged at B. Therefore, at such wage levels $wl^* > c^*$. QED

As the wage rate grows, work undergoes a metamorphosis in this model. At ordinary wage rates, only the budget contraint (2) is binding at the optimum. At top wage rates, only the time contraint (3) is binding at the optimum. In between, both constraints bind.[7] Whereas at ordinary wages work is a means to earn one's livelihood, at top wages it is an end in itself, competing with consumption activities as an alternative use of time. At wages in the intermediate range, it shares both natures of means and end.

Within this intermediate range, work gradually converts into a consumption activity: individuals receiving more than w_+ derive utility from their last hour of work. They optimally refrain from expanding pleasant work because doing so would reduce the time available for their consumption activities.

As income effects are assumed away, the labor-supply curve of ordinary earners is increasing with the wage. This does not hold true for earners in the intermediate range. If both constraints are binding, their labor supply equals l' which, as shown by (4), is decreasing with the wage. The backward bending of the labor-supply curve is not due to income effects - which have been assumed away - but to the fact that the time constraint becomes binding if the wage and hence the individual's expenditure for consumption activities are large enough. Then, a higher wage leads to more consumption and, mechanically, to less time devoted to work.

For wages larger than $\widehat{w}$, the budget constrain is slack and the optimal labor supply strikes a balance between the marginal utility gain from personal fulfillment on the job and the marginal utility loss from less time for consumption activities. Thus, l^* is independent of the wage rate and equals

$$\frac{\alpha - \theta^{-1}}{\beta} \equiv \widehat{l}. \tag{5}$$

The time coefficient of consumption, θ, is now equal to the opportunity cost of consumption. The larger θ is, the smaller the amount of consumption that has to be given up for an additional hour of enjoyable work, and the larger the optimal labor supply.[8]

Super-rich with $w > \widehat{w}$ do not consume the entire amount of numeraire good they earn. This is not because they are satiated - their marginal utility from consumption is strictly positive - but because they have not enough time to spend their earnings. They optimally leave some earnings unspent because they are not willing to forsake time of personally fulfilling work in order to consume more.

Seen it through the lens of general equilibrium theory, the introduction of a time

[7] This is formally proven in Appendix A.1, which also provides the explicit closed-form solutions for optimal consumption and labor, as well as the critical wages, $w_+ = (\beta T - \alpha)/\alpha\theta$ and $\widehat{w} = (\beta T - \alpha + \theta^{-1})/(\alpha\theta - 1)$.

[8] If, contrary to my assumption, $\alpha < 1/\theta$, the budget constraint always binds at the individual optimum. Then, the labor supply curve has no vertical half-line but an asymptote: as w goes to ∞, l^* goes to zero. In that case, the consumption level asymptotically converges to T/θ.

constraint implies that the consumption set of agents is bounded. In this case, preferences cannot be locally nonsatiated, and the first fundamental theorem of welfare economics fails.[9] By way of an example, consider an economy with firms and two groups of agents, one with productivity strictly lower than $\widehat{w}$ and one with productivity strictly higher than $\widehat{w}$. An allocation supported by a relative wage equal to relative productivity is a competitive equilibrium. But that equilibrium is not a Pareto optimum because one could transfer some numerarire good from the second group to the first one without decreasing the utility of the latter and making the former strictly better off.

In a monetary economy where money is used as a medium of exchange, it is natural to interpret situations where the budget constraints of some agents do not bind as demand-constrained allocations rather than competitive equilibria. If money is the institutionally necessary counterpart of any transaction, top earners with a wage in excess of $\widehat{w}$ will optimally refrain from spending their entire money income.[10] Casting result (ii) of Prop. 1 in such a monetary framework has a remarkable macroeconomic implication. Let N denote the size of the workforce and $f(w)$ the density of the skill distribution. Aggregating the budget constraints across all workers implies that aggregate demand falls short of aggregate supply by

$$\Delta = N \int_{\widehat{w}}^{\infty} \left[\frac{\widehat{l}(1 + w\theta) - T}{\theta} \right] f(w) dw.$$

The larger the wage share of those earning more than $\widehat{w}$, the larger is Δ, the lack of aggregate demand. This is consistent with the relationship between high inequality and the occurrence of macroeconomic crises that is sometimes put forward in policy debates.[11] Potentially, the erosion of effective demand highlighted by this model may be quantitatively significant. If the share of unspent income of the super-rich equals, say, 20 % and their income represents five percent of national income, then this would generate a wedge between aggregate supply and aggregate demand equal to one percentage point of national income.

[9] See e.g. Mas-Colell et al. (1995, ch. 16).

[10] This function of money could be modeled using a liquidity constraint along the lines of Grandmont and Younes (1972).

[11] See Kumhof et al. (2015) for a discussion of the stylized facts. In their model, the channel linking the income distribution to crises is the debt leverage at the bottom of the distribution. I examine the role of savings in section 3.

2.2 Optimal top marginal tax rate

The top marginal tax rate is the one at which incomes above a certain threshold are taxed.[12] Its optimal level is usually assessed applying the theory of optimal taxation, which builds on the standard labor supply model.[13] According to this theory, the marginal tax rate on the highest income should be zero if the government knows that income level ex ante. A positive optimal top marginal tax rate obtains if the maximum income subject to taxation is ex ante unknown to the government and uncertainty about the top of the income distribution is captured by an unbounded distribution of skills.[14] By contrast, even a confiscatory top marginal tax rate can be optimal if one incorporates the assumptions that some labor is enjoyable and consumption requires time. It suffices that the threshold income be such that the labor-supply curve of those subject to the top marginal tax rate is backward bending or vertical.

To see it formally, assume that the income tax schedule has a top tax bracket that starts at a cutoff level $\overline{y}$. The tax liability at that income level equals $\bar{t}$. Denoting by $\tau \in [0, 1]$ the top marginal tax rate, the budget constraint for individuals in the top tax bracket reads

$$c \leq wl - \bar{t} - \tau(wl - \overline{y}) = w(1 - \tau)l - \bar{t} + \tau\overline{y}. \tag{6}$$

This inequality now replaces (2) from the laissez-faire model. The time constraint is still given by (3). Without significant loss of generality, and as e.g. in Piketty et al. (2014), suppose that the social welfare function puts zero weight on the utility of top bracket taxpayers. The optimal top tax rate τ^* is thus the one that maximizes tax revenue from those individuals.

In order for a confiscatory top marginal tax rate to be optimal, it is sufficient that the time constraint be binding for the individuals in the top tax bracket. This can be expressed as an assumption about the cutoff level of income $\overline{y}$, which has to be large enough.

Assumption (A): The cutoff level of income $\overline{y}$ is such that $\frac{\partial U(c^*, l^*)}{\partial l} > 0$ if $wl^* = \overline{y}$.

As implied by the analysis in the preceding section, the assumption that the individual

[12] An early overview of taxation issues concerning the rich is offered by Slemrod (1994). For a recent appraisal, see Diamond and Saez (2011). Bach et al. (2013) investigate the taxation of top incomes in Germany.

[13] See however Ales and Sleet (2016) and Scheuer and Werning (2017) who consider the role played by superstar effects in earnings determination.

[14] See Diamond (1998) and Saez (2001).

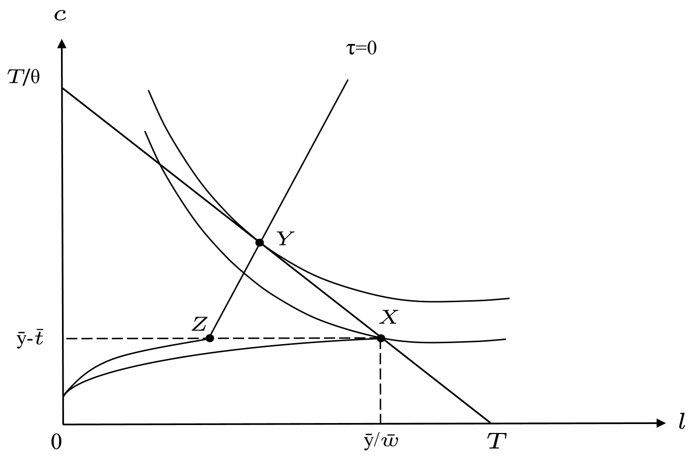

Figure 3: Top tax rate and individual optima.

at the cutoff level of income has a positive marginal utility of work ensures that the time constraint is binding for all top bracket taxpayers.

Proposition 2. *If (A) holds, a confiscatory top marginal tax rate is optimal.*

Proof.

Let $\overline{w}$ denote the wage rate of the individuals that optimally earn $\overline{y}$. Two cases have to be distinguished, depending on the budget constraint (6) being binding or not at the optimum. Suppose first that it is, as in the case depicted in Figure 3, where point X yields the optimal bundle for an individual with wage $\overline{w}$.

The top tax bracket includes all individuals whose wage is larger than $\overline{w}$. Fig. 3 also shows the budget constraint of an individual with a wage that is strictly larger than $\overline{w}$ under the assumption $\tau = 0$, in which case the individual optimally chooses point Y. Increasing τ makes this individual's budget line rotate clockwise around point Z until it reaches the horizontal position for $\tau = 1$. As τ increases from 0 to 1, the individual optimum moves along the time constraint from Y to X. Labor supply, earnings, and tax revenue are maximized at point X which corresponds to $\tau = 1$. Since the same reasoning applies to every individual in the top tax bracket and the optimal top marginal tax rate is the one that maximizes total tax revenue, $\tau^* = 1$.

Consider now the remaining case where the budget constraint for the individual with wage $\overline{w}$ is not binding at the optimum. This case is depicted in Figure 4 where the optimum is again denoted by X. Fig. 4 also shows the budget constraint of an individual with

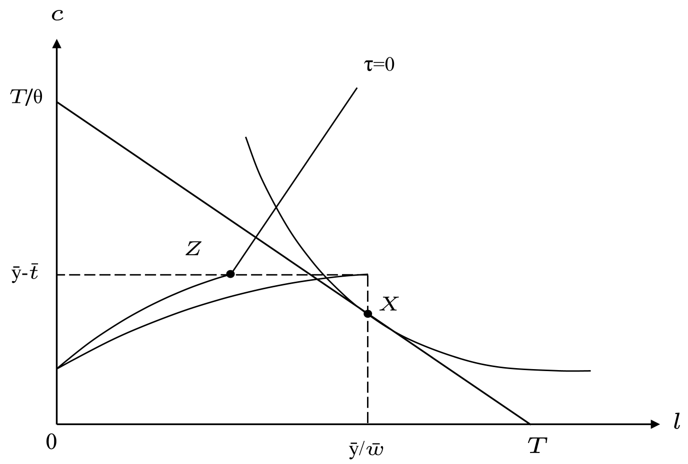

Figure 4: Top tax rate and individual optima.

a wage that is strictly larger than $\overline{w}$ under the assumption $\tau = 0$. Also this individual optimally chooses point X. Increasing τ makes this individual's budget line rotate clockwise around point Z until it reaches the horizontal position for $\tau = 1$. As τ increases from 0 to 1, the individual optimum remains fixed at X. Tax revenue is therefore maximized by $\tau = 1$ and since this is true for every individual in the top tax bracket, $\tau^* = 1$. QED

The current model uncovers forces that are absent from the traditional one and dramatically strengthen the case for redistribution. If the budget constraint of top earners binds - the case in Fig. 3 - increasing their marginal tax rate increases their labor supply and total output, thereby reversing the sign of their behavioral response as found in the standard labor supply model. Since in the current model top earners are short of time, they optimally select that length of work that leaves them precisely the time they need in order to spend their net earnings. A higher tax mechanically reduces consumption spending, which makes more time available for work. Hence, their earnings and tax payments increase in response to increased taxation. If instead the budget constraint of top earners is slack - the case in Fig. 4 - increasing the top marginal tax rate has no behavioral consequences. For those top earners, work is not instrumental in generating consumption opportunities, but a value in itself. Increasing the top tax rate simply reduces the waste associated with top earners' unspent disposable income and allows for a Pareto-improvement by redistributing income to those who have the time to consume it.[15]

To the best of my knowledge, there is no systematic direct evidence on the labor-supply curve of the super rich.[16] However, there is a literature based on tax returns that explores the response of top earners' *taxable* income to τ and usually finds small negative responses.[17] In light of Proposition 2, such empirical findings may be interpreted as suggesting that the group of taxpayers affected by the top marginal tax rate includes individuals whose time constraint is slack, i.e. $\overline{y}$ is not sufficiently large for assumption (A) to be descriptively accurate. Alternatively, (A) may be accurate and the observed negative responses of taxable income to τ be driven by increased incentives to avoid taxes and/or decreased incentives to engage in compensation bargaining - rather than by real supply-side effects.[18] Tax avoidance and bargaining incentives could arise also in the current model if the taxpayers' budget constraints bind. If the budget constraint is slack, taxpayers are satiated and have thus no incentive to engage in tax avoidance or compensation bargaining. However, they could find it optimal to engage in such activities if they desire income for status reasons, e.g. to rank high in Forbes lists.

3 Savings and bequests

The super-rich often have descendants they support by means of bequests. Instead of using income for own consumption, they make it available on the capital market where it can be used for investment. The capitalized savings are then transferred to their descendants. Such bequests differ from personal consumption activities in that larger transfers do not entail an additional use of time. Do the insights from the preceding section survive the inclusion of savings and bequests?

The main aspects of this issue can be gauged from a simple two-period model in which every top earner lives one period and has one offspring. In period 0, the top earner works, consumes and saves. Her savings are bequeathed to her offspring who works and consumes

[15] Therefore, the optimal top marginal tax rate is 100 percent even if the top earners receive a large weight in the social welfare function.

[16] There is however a study by Moffit and Wilhelm (2000) for the U.S. They find that the labor supply of the affluent is virtually vertical.

[17] See Saez et al. (2012) for a careful overview.

[18] Piketty et al. (2014) offer a model that includes those two channels. They show that $\tau^* = (1 + tae_2 + ae_3)/(1 + ae)$, where a is the Pareto coefficient of top incomes, e is the aggregate elasticity of taxable income in the top bracket with respect to the net-of-tax rate, e_2 is the tax avoidance elasticity component, t is the marginal tax rate at which sheltered income is taxed, and e_3 is the compensation bargaining elasticity component. In turn, e is the sum of e_2, e_3 and the standard elasticity of labor supply. While reliable estimates of the total elasticity e are available, decomposing it into its three components is difficult. Piketty et al. (2014) present some estimation results based on aggregate data and conclude that real supply-side effects play a minor role. The current model suggests that for some top earners $e_2 + e_3 > 1$.

in period 1. While every top earner receives a very high wage but no inheritance, the wage of her descendant need not be high and his inheritance is endogenously determined. Without significant loss of generality, I assume that the descendant earns a wage equal to zero. Every top earner is assumed to be altruistic with respect to her descendant and thus to maximize

$$U = v(c_0) + \alpha l_0 - \frac{\beta}{2} l_0^2 + \gamma \left[v(c_1) + \alpha l_1 - \frac{\beta}{2} l_1^2 \right], \tag{7}$$

where $v' > 0 > v''$ and $\gamma \in (0,1)$. Given her wage, the top earner chooses her labor supply l_0 and savings so as to maximize (7), taking her descendant's decisions in period 1 into account. This amounts to maximizing (7) under the intertemporal budget constraint

$$w l_0 \geq c_0 + \frac{c_1}{1+r}, \tag{8}$$

and the time constraints

$$T_i \geq \theta_i c_i + l_i, \tag{9}$$

where $i = 0, 1$.[19] The model of the preceding section obtains as a limiting case of this one if both γ and v'' go to zero. In analogy to that model, I posit $\alpha > v'(T_i/\theta_i)/\theta_i$, $i = 0, 1$.

Distinctive properties of labor supply arise in this model if the wage rate of the parent, w, becomes large enough, more precisely if it is larger than the lowest w such that both time constraints (9) are binding at the optimum. Let w' denote such a threshold wage level. For $w \geq w'$, the labor supply of such a top earner can be derived from the following two-step program. First, the consumption levels of the two generations (c_0, c_1) are chosen so as to maximize

$$U = v(c_0) + \alpha(T_0 - \theta_0 c_0) - \frac{\beta}{2}(T_0 - \theta_0 c_0)^2 + \gamma \left[v(c_1) + \alpha(T_1 - \theta_1 c_1) - \frac{\beta}{2}(T_1 - \theta_1 c_1)^2 \right], \tag{10}$$

subject to

$$w T_0 \geq (1 + \theta_0 w) c_0 + \frac{c_1}{1+r}. \tag{11}$$

Then, the labor supply of the top earner is determined as

$$l_0 = T_0 - \theta_0 c_0. \tag{12}$$

The solution to this program is denoted by $l_0^*(w)$.

[19] This model is formally equivalent to the usual two-period model with work in period 1 and retirement in period 2, augmented with time constraints.

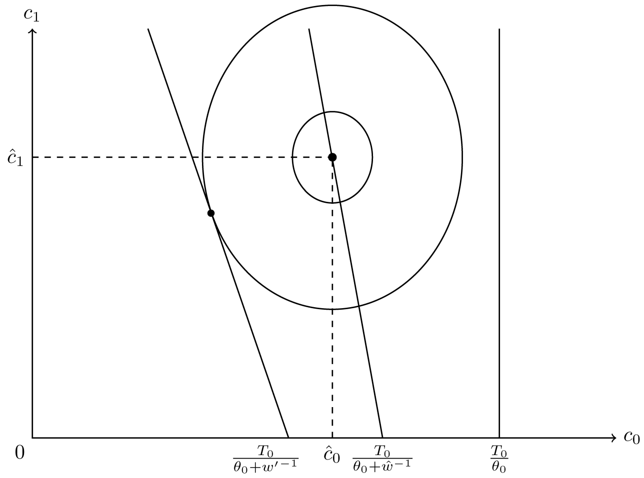

Figure 5: Optima at wage levels w' and $\widehat{w}$.

Proposition 3. *There exists a wage level $\widehat{w} > w'$, such that for any $w \geq \widehat{w}$ the labor supply of the top earners is a constant $\widehat{l_0} = T_0 - \theta_0 \widehat{c}_0$,* where $\widehat{c}_0$ is the unique solution to

$$\frac{v'(\widehat{c}_0)}{\theta_0} = \alpha - \beta(T_0 - \theta_0 \widehat{c}_0). \tag{13}$$

For $w > \widehat{w}$ the intertemporal budget constraint (8) is not binding at the optimum. Furthermore, $\widehat{l_0} < l_0^*(w)$ if $w \in [w', \widehat{w})$.

Proof.

As illustrated by Figure 5, the indifference curves of the utility function (10) are quasi-circles around the bliss point $(\widehat{c}_0, \widehat{c}_1)$ determined by

$$v'(\widehat{c}_i) = \alpha\theta_i - \beta\theta_i(T_i - \theta_i \widehat{c}_i), \tag{14}$$

for $i = 0, 1$. Define $\widehat{w}$ as the smallest w such that $(\widehat{c}_0, \widehat{c}_1)$ satisfies the budget constraint (11). Increasing w above $\widehat{w}$ shifts the budget line to the right and makes it converge to the vertical line defined by $c_0 = T_0/\theta_0$. Those wage increases have no effect on optimal consumption and labor supply, which implies $wl_0^* > c_0^* + \frac{c_1^*}{1+r}$.

If $w \in [w', \widehat{w})$, the optimum necessarily has $c_i^* < \widehat{c}_i$, $i = 0, 1$, because the indifference curve must be negatively sloped in order to be tangential to the budget constraint. Using (12), this implies $l_0^*(w) > \widehat{l_0}$. QED

An immediate corollary of Proposition 3 is that the marginal utility of work is positive for the top earners. This is apparent from (13) since its RHS is the marginal utility of work and its LHS is strictly positive. As implied by (14), at the optimum also their heirs receive a positive marginal utility from working.

Thus, the model with savings inherits the distinctive properties of labor supply of the static model in the preceding section: positive marginal utility from work, backward-bending labor-supply curve, and the possibility of rational underconsumption. This applies in a symmetric way to the determination of the optimal taxation of top earners. To be more precise, consider the tax schedule of the previous section with a top tax bracket that starts at a cutoff level of income $\overline{y}$. Denoting by $\tau \in [0, 1]$ the top marginal tax rate, the intertemporal budget constraint of individuals in the top tax bracket reads

$$wl_0(1-\tau) + \tau\overline{y} - \overline{t} \geq c_0 + \frac{c_1}{1+r},$$

which can be written as

$$w(1-\tau)T_0 + \tau\overline{y} - \overline{t} \geq [1 + \theta_0 w(1-\tau)]c_0 + \frac{c_1}{1+r}. \tag{15}$$

Following a similar line of reasoning as in the preceding section it is straightforward to show that if the cutoff income $\overline{y}$ is sufficiently high, the optimal top marginal tax rate is 100 percent. If the intertemporal budget constraint in not binding at the optimum, i.e.

$$\overline{y} > \overline{t} + (1 + \theta_0 w)\widehat{c}_0 + \frac{\widehat{c}_1}{1+r},$$

then, increasing τ does not affect consumption and labor supply, so that $\tau^* = 1$. If $\overline{y}$ is lower, so that the budget constraint may bind, but it is still high enough for the time constraints remaining binding, increasing τ may decrease c_0^* and thus increase l_0^*. This is necessarily so if θ_0 is sufficiently close to zero because in that case the increase in τ is similar to a pure negative income effect, as apparent from (15). Since the utility function (10) is separable, the good c_0 is normal and its consumption diminishes if τ is increased, hence l_0^* increases. Also in that case, $\tau^* = 1$. Disincentives to work may only set in if the increase in τ leads to such a strong substitution of c_1 by c_0 that the latter increases despite the decrease of net full income.

According to Proposition 3, the model with savings also inherits from the static model the possibility that some income may never be spent. If the parent correctly anticipates that simply trasferring her wealth to the offspring will provide the latter with so much disposable income that the offspring will not be able to entirely consume it, then there is no reason for the parent to bother about investing her wealth. Money may lay forgotten in some bank account and banknotes may be used to light cigars.

This model might offer a building block of a microfunded Keynesian theory of aggregate output determination. However, as compared to the static model, the condition for the budget constraint of a super-rich to be slack is more restrictive: it is only for wage levels such that also her descendant's bliss point of consumption is reached that some income is left unspent - and a shortfall of aggregate demand occurs. An even stronger qualification applies to the case of a model with an arbitrary number of generations. If each dynasty has G generations, where G can be infinite, the dynasty of a super-rich leaves some income unspent if and only if the bliss-point level of consumption can be reached for every member of her dynasty, i.e.

$$\sum_{i=0}^{G} \frac{w_i T_i}{(1+r)^i} > \sum_{i=0}^{G} \frac{(1+\theta_i w_i)\widehat{c}_i}{(1+r)^i},$$

where $\widehat{c}_i$ is implicitly defined by (14) for all i.

4 Charitable giving

Beyond personal consumption and bequests, donations constitute a significant category of expenditure of the super-rich. By way of an example, some super-rich recently started an initiative called "The Giving Pledge", promoting voluntary commitments by billionaires to dedicate more than half of their wealth to philanthropy.

One may argue that adding more zeros to a check for a donation requires a negligible amount of time and such a category of expenditure should therefore be excluded from the time constraint (3). In reality, as everybody personally acquainted with charitable giving knows, philanthropic engagement is a time consuming activity: any considerable additional donation comes along with a screening of potential recipients, a decision on the allocation of the money to be donated, and a monitoring of the use made of it, all activities that have to take into account the behavior of other actual and potential givers, and all activities that require a significant amount of time to be properly performed. To the extent that the quality of own donations matters to the giver - hopefully a realistic feature of actual giving - philanthropic expenditures should therefore affect the time budget of individuals in a similar fashion as the customized consumer goods discussed in section 2. The model in that section may thus capture truly dedicated philanthropy.

If quality concerns for donations do not arise, charitable giving may better be modeled as an income use that does not require time. And differently from the inheritances considered in the model of the preceding section, it is unlikely that philanthropic donations make their recipients' time constraints binding. In order to capture the role of such donations, it is thus helpful to revert to the static model of section 2 and modify it by introducing a warm-glow motive in the utility function, that is to posit:

$$U = c + \gamma \frac{g^{1-\sigma}}{1-\sigma} + \alpha l - \frac{\beta}{2} l^2, \tag{16}$$

where $g \geq 0$ is the amount of charitable giving and parameters γ and σ are strictly positive. In order to ensure an interior solution, I posit that γ is bounded from above by a strictly positive number $\overline{\gamma}$, that will be determined shortly. The utility function (16) is maximized under the budget constraint

$$c + g \leq wl \tag{17}$$

and the time constraint (3).

An immediate consequence of introducing a philanthropic motive in this way is that the budget constraint (17) must be binding at the optimum. This eliminates the possibility of rational underconsumption - a possibility that arose in the models of the two previous sections. By contrast, it does not preclude the possibility of a backward-bending labor supply, and thus the optimality of a confiscatory top marginal tax rate.

Proposition 4. *If $\sigma > 1$, there exists a wage rate w' such that for all $w > w'$, $dl^*/dw < 0$. Furthermore, $\lim_{w \to \infty} l^*(w) = \widehat{l}$.*

Proof.

Starting from a wage rate such that the time constraint is not binding, it is routine to demonstrate that increasing the wage rate increases c^* and l^* until at some wage w' the time constraint (3) becomes binding at the optimum. Hence, for $w > w'$, and assuming for the moment being an interior solution, the optimal labor supply and charitable giving obtain from maximizing

$$\mathcal{L} = wl - g + \gamma \frac{g^{1-\sigma}}{1-\sigma} + \alpha l - \frac{\beta}{2} l^2 + \lambda \left[T - \theta(wl - g) - l \right],$$

where $\lambda > 0$ is a Lagrange multiplier. Computing the first-order conditions and substituting out λ yields:

$$w + \alpha - \beta l^* = \left(\frac{1 - \gamma {g^*}^{-\sigma}}{\theta} \right) (1 + \theta w). \tag{18}$$

Using (3) to substitute out consumption from the budget constraint yields:

$$g^* = \left(w + \theta^{-1} \right) l^* - \frac{T}{\theta}. \tag{19}$$

Combining eqs. (18) and (19) and rearranging, one obtains

$$\alpha - \beta l^* = \frac{1}{\theta} - \frac{\gamma(1+\theta w)}{\theta}\left[\frac{(1+\theta w)l^* - T}{\theta}\right]^{-\sigma}, \tag{20}$$

which implicitly defines the optimal labor supply l^*. Using this equation, it is easy to demonstrate that $l^* \in (0, T)$ if $\gamma < [1+\theta(\beta T - \alpha)](w'T)^\sigma/(1+\theta w') \equiv \overline{\gamma}$. Hence, $wl^* > 0$, and by contradiction it is standard to prove that $g^* > 0$. It remains to be shown that the solution obtained from eqs. (20) and (19) implies $c^* > 0$. This follows from the binding time constraint and $l^* < T$.

Denote by $F(l^*, w)$ the RHS of (20). Using the implicit function theorem it is straightforward to show that $dl^*/dw < 0$ if and only if $F(l^*, w)$ is increasing in w. Computing its partial derivative yields

$$\frac{\partial F}{\partial w} = -\gamma {g^*}^{-\sigma}\left[1 - \left(\frac{1+\theta w}{\theta}\right)\frac{\sigma l^*}{g^*}\right].$$

Hence, $\partial F/\partial w > 0$ if and only if

$$g^* < \sigma\left(w + \theta^{-1}\right) l^*.$$

Substituting (19) into this inequality shows that $dl^*/dw < 0$ if and only if

$$\sigma > \frac{1}{1 + \frac{T}{\theta g^*}},$$

which is necessarily satisfied if $\sigma > 1$.

The asymptotic behavior of the labor supply follows from noting that

$$\lim_{w\to\infty} F(l^*, w) = \frac{1}{\theta}$$

if $\sigma > 1$. From this and (20) one has

$$\lim_{w\to\infty} l^*(w) = \frac{\alpha - \theta^{-1}}{\beta} = \widehat{l}, \tag{21}$$

as defined by (5). QED

At wages larger than w', the time constraint is binding at the optimum. Similarly to the model of section 2, further wage increases are accomodated by a decrease of labor effort and an increase of consumption activities so as to exhaust the time endowment. However, in the current model an increasing share of those additional earnings is spent on charitable giving, an activity which, by assumption, does not require time. From (19) and the budget constraint, the share of income devoted to charitable giving is

$$\frac{g^*}{wl^*} = \frac{w+\theta^{-1}}{w} - \frac{T}{\theta w l^*} = 1 - \left(\frac{T - l^*}{\theta w l^*}\right),$$

so that

$$\lim_{w\to\infty}\frac{g^*}{wl^*}=1$$

by (21). Since $\widehat{l}<\widetilde{l}$, an immediate corollary of Prop. 4 is that the marginal utility of labor is positive for the philanthropic super-rich if their wage rate is sufficiently high. Moreover, the backward bending of the labor-supply curve for high wages implies that the optimal top marginal tax rate can be 100 percent if the cutoff income level $\overline{y}$ is high enough. Then, increasing τ and thus reducing the net wage makes the super-rich consume less and work more. The role of the top marginal tax rate is similar to the one it played in the models of sections 2 and 3 in the case in which both constraints are binding.[20]

How restrictive is the (sufficient) condition $\sigma>1$? It may be noted that for ordinary earners, i.e. agents with a wage rate lower than w', $1/\sigma$ equals the price elasticity of charitable giving, in absolute terms. The most recent empirical studies find charitable giving to be rather price inelastic. For instance, in a natural experiment framework Fack and Landais (2010) find price elasticities in absolute value to be in a range between 0.2 and 0.6, which suggests that the condition $\sigma>1$ is one that is likely to be satisfied in practice.

5 Conclusion

I have revisited the standard model of labor supply under two additional assumptions that make it more suitable to analyze the behavior of the super-rich: consumption requires time and some limited amount of work is enjoyable. Whereas introducing each assumption without the other one does not produce novel insights, combining them together does. The working super-rich consume up to a point where their time constraint binds and optimally stop working at a point where their marginal utility of work is still positive. Their labor-supply curve is backward bending even in the absence of income effects. In such a situation, a top marginal tax rate on labor income of 100 percent can be optimal. Furthermore, some of the income accruing to the super-rich dynasties may optimally never be spent. In this way, the microeconomic behavior of the super-rich may trigger a lack of effective demand at the macroeconomic level which can be substantial if their share in total income is large.

If labor productivity keeps growing at a higher rate than the length of human life, over time an ever increasing share of the workforce may come to face a decision problem

[20] As usual, the desirability of taxation is subject to the qualification that the redistributive objective of the planner cannot be achieved more efficiently through private donations rather than by means of social transfers.

qualitatively similar to the one faced by today's working super-rich. For those future workers the key trade-off will not be the one between less leisure and more commodities but between less time for personally rewarding work and more time for consumption activities. The model developed in this paper suggests that in such a future economy the incentive costs of taxing above-average incomes may be substantially lower than today. This would loosen to a great extent the restrictions on political redistribution that are today imposed by efficiency considerations.

References

Ales, L. and C. Sleet (2016), Taxing top CEO incomes, *American Economic Review* 106, 3331-66.

Atkinson, A, and T. Piketty (2010), *Top Incomes - A Global Perspective*, Oxford: Oxford University Press.

Bach, S., Corneo, G. and V. Steiner (2013), Effective taxation of top incomes in Germany, *German Economic Review* 14, 115-137.

Baron, J. (1988), The employment relationship as a social relation, *Journal of the Japanese and International Economies* 2, 492-525.

Becker, G. (1965), A theory of the allocation of time, *Economic Journal* 75, 493-517.

Corneo, G. and R. Rob (2003), Working in public and private firms, *Journal of Public Economics* 87, 1335-1352.

Csikszentmihalyi, M. and J. LeFevre (1989), Optimal experience in work and leisure, *Journal of Personality and Social Psychology* 56, 815-822.

Deci, E. and R. Ryan (1980), The empirical exploration of intrinsic motivational processes, *Advances in Experimental Social Psychology* 13, 39-80.

Diamond, P. (1998), Optimal income taxation: An example with a U-shaped pattern of optimal marginal tax rates, *American Economic Review* 88, 83-95.

Diamond, P. and E. Saez (2011), The case for a progressive tax: From basic research to policy recommendations, *Journal of Economic Perspectives* 25, 165-190.

Fack, G. and C. Landais (2010), Are tax incentives for charitable giving efficient? Evidence from France, *American Economic Journal: Economic Policy* 2, 117-141.

Funk, P. (2015), Human capital, polarization, and Pareto-improving activating welfare, mimeo University of Cologne.

Grandmont, J.-M. and Y. Younes (1972), On the role of money and the existence of a monetary equilibrium, *Review of Economic Studies* 39, 355-372.

Kleven, H. J. (2004), Optimum taxation and the allocation of time, *Journal of Public Economics* 88, 545-557.

Kreps, D. (1997), Intrinsic motivation and extrinsic incentives, *American Economic Review* 87, 359-364.

Kumhof, M., Ranciere, R. and P. Winant (2015), Inequality, leverage and crises, *American Economic Review* 105, 1217-1245.

Mas-Colell, A., Whinston, M. and J. Green (1995), *Microeconomic Theory*, New York: Oxford University Press.

Moffitt, R. and M. Wilhelm (2000), Taxation and the labor supply decisions of the affluent, in Slemrod, J. (ed.), *Does Atlas Shrug? The Economic Consequences of Taxing* *the Rich*, New York: Harvard University Press and Russel Sage Foundation.

Piketty, T., Saez, E. and S. Stantcheva (2014), Optimal taxation of top labor incomes: A tale of three elasticities, *American Economic Journal: Economic Policy* 6, 230-271.

Saez, E. (2001), Using elasticities to derive optimal income tax rates, *Review of Economic Studies* 68, 205-229.

Saez, E. (2015), Striking it richer: The evolution of top incomes in the United States (Updated with 2014 preliminary estimates), mimeo UC Berkeley.

Saez, E., Slemrod, J. and S. Giertz (2012), The elasticity of taxable income with respect to marginal tax rates: A critical review, *Journal of Economic Literature* 50, 3-50.

Saez, E. and G. Zucman (2016), Wealth inequality in the United States since 1913: Evidence from capitalized income tax data, *Quarterly Journal of Economics* 131, 519-578.

Scheuer, F. and I. Werning (2017), The taxation of superstars, forthcoming *Quarterly Journal of Economics*.

Slemrod, J. (1994), On the high-income Laffer curve, in Slemrod, J. (ed.), *Tax Progressivity and Income Inequality*, Cambridge: Harvard University Press.

APPENDIX

A.1 - The changing nature of work in the basic model[21]

This appendix derives the complete labor supply curve of the model of section 2.1. As the wage rate increases from 0, first only the budget constraint (2) binds at the optimum, then both constraints bind, and finally only the time budget constraint (3) is binding at the optimum. Denote by w' the wage at which the first regime switch occurs and by $\widehat{w}$ the wage at which the second regime switch occurs. I am going to show that w' is the positive root of the quadratic equation

$$(\theta w + 1)(\alpha + w) - \beta T = 0, \tag{22}$$

and that

$$\widehat{w} = \frac{\beta T - \alpha + \theta^{-1}}{\alpha\theta - 1}. \tag{23}$$

The problem faced by the agent is to choose positive levels of c and l so as to maximize (1) subject to (2) and (3). Recall from the main text that we posit $T > \alpha/\beta$ and $\alpha > 1/\theta$. The Kuhn-Tucker conditions for the solution (which are necessary and sufficient) are that the negative gradient of the objective is in the cone spanned by the gradients of the binding constraints. Let $\lambda \geq 0$ denote the Lagrange multiplier associated with the budget constraint and let $\mu \geq 0$ denote the Lagrange multiplier associated with the time constraint. It is apparent that a solution where both constraints are slack is impossible. We are thus left with three cases to consider.

Case 1: both constraints bind.

In this case, the Kuhn-Tucker conditions are:

$$1 = \lambda + \mu\theta, \tag{24}$$

$$\beta l - \alpha = \lambda w - \mu. \tag{25}$$

From the two binding primal constraints, compute that

$$c^* = \frac{T}{\theta + w^{-1}},$$

$$l^* = \frac{T}{\theta w + 1}, \tag{26}$$

which corresponds to Eq. (4) in the main text.

[21] I am grateful to John Roemer for offering several of the computations presented here.

Because both contraints bind, both multipliers must be strictly positive. From (24), (25) and (26), one has:

$$\mu^* = \frac{(\theta w + 1)(\alpha + w) - \beta T}{(\theta w + 1)^2},$$
$$\lambda^* = 1 - \mu^* \theta.$$

The condition $\mu^* > 0 \wedge \lambda^* > 0$ is thus equivalent to:

$$(\theta w + 1)(\alpha + w) - \beta T > 0, \tag{27}$$

$$w\theta(\alpha\theta - 1) < \theta\beta T - \alpha\theta + 1. \tag{28}$$

The LHS of condition (27) is a quadratic function with roots given by

$$\frac{-(\alpha\theta + 1) \pm \sqrt{(\alpha\theta + 1)^2 + 4\theta(\beta T - \alpha)}}{2\theta}.$$

The positive root is

$$w' = \frac{-(\alpha\theta + 1) + \sqrt{(\alpha\theta + 1)^2 + 4\theta(\beta T - \alpha)}}{2\theta}. \tag{29}$$

It follows that condition (27) is satisfied if and only if

$$w > w'.$$

Rearranging terms in condition (28) and using (23) shows that (28) is satisfied if and only if

$$w < \widehat{w}.$$

So, case 1 obtains if and only if

$$w' < w < \widehat{w}.$$

For this case not to be vacuous, I need to show that $\widehat{w} > w'$. Suppose by way of contradiction that the opposite were true. From (23) and (29), this implies

$$\sqrt{(\alpha\theta + 1)^2 + 4\theta(\beta T - \alpha)} \geq 1 + \alpha^2\theta^2 + 2\theta(\beta T - \alpha).$$

Squaring both sides and rearranging terms yields:

$$\alpha\theta(2 + \alpha\theta) + 4\theta(\beta T - \alpha) + 0 \geq \alpha^2\theta^2(2 + \alpha^2\theta^2) + 4\theta(1 + \alpha^2\theta^2)(\beta T - \alpha) + 4\theta^2(\beta T - \alpha)^2.$$

Comparing term by term the two sides of this inequality shows that each term on the RHS is strictly larger than its counterpart on the LHS. Hence, we have a contradiction which proves that $\widehat{w} > w'$.

Case 2: only the budget constraint binds.

The Kuhn-Tucker conditions become:

$$1 = \lambda,$$

$$\beta l - \alpha = \lambda w.$$

Whence, $\lambda^* > 0$ and

$$l^* = \frac{w + \alpha}{\beta}, \tag{30}$$

and from the budget constraint,

$$c^* = \frac{w(w + \alpha)}{\beta}.$$

We have to check that the time constraint is slack. This reduces to the inequality:

$$(\theta w + 1)(\alpha + w) < \beta T.$$

This is satisfied if and only if condition (27) is not:

$$w < w'.$$

Case 3: only the time constraint binds.

The Kuhn-Tucker conditions become:

$$1 = \mu\theta,$$

$$\beta l - \alpha = -\mu.$$

Whence, $\mu^* > 0$ and

$$l^* = \frac{\alpha\theta - 1}{\beta\theta}, \tag{31}$$

which corresponds to Eq. (5) in the main text. From the time constraint, one has

$$c^* = \frac{\beta T - \alpha + \theta^{-1}}{\beta\theta}.$$

We have now to check that the budget constraint is slack. This reduces to the inequality:

$$T < \frac{(\alpha\theta - 1)(1 + \theta w)}{\beta\theta}.$$

Rearranging terms and using (23), this is equivalent to

$$w > \widehat{w}.$$

Reverting to case 1 above, we can compute the critical wage w_+ of Prop. 1, starting from which labor has a positive marginal utility at the optimum. From $\alpha - \beta l^* = 0$ and (26), that wage is

$$w_+ = \frac{\beta T - \alpha}{\alpha\theta}.$$

For this to be the wage mentioned in Prop. 1, one has to check that $w_+ \in (w', \widehat{w})$. This is easily demonstrated by deriving a contradiction if the opposite were true.

A.2 - The model with pure free time

This appendix shows that including pure free time - i.e. leisure wihout any consumption of commodities - in the utility function is immaterial for top earners as long as the marginal utility of free time is bounded from above by $1/\theta$.

Let $f \geq 0$ denote pure free time and replace the time constraint (3) with

$$T = \theta c + l + f. \tag{32}$$

The marginal utility from free time is strictly positive, nonincreasing, and bounded from above. Without loss of generality, assume that it is a constant δ so that the utility function (1) is replaced with

$$U = c + \alpha l - \frac{\beta}{2} l^2 + \delta f. \tag{33}$$

The problem is to maximize (33) subject to (2) and (32). The following claim is to be shown: if $\delta < 1/\theta$, there exists a wage rate $\underline{w} > 0$ such that $f^* = 0$ for all $w \geq \underline{w}$.

From the Lagrangean

$$\mathcal{L} = c + \alpha l - \frac{\beta}{2} l^2 + \delta f + \lambda(wl - c) + \mu(T - \theta c - l - f),$$

one obtains the following FOCs:

$$1 - \lambda^* - \theta\mu^* \leq 0, \tag{34}$$

$$\alpha - \beta l^* + w\lambda^* - \mu^* \leq 0, \tag{35}$$

$$\delta - \mu^* \leq 0. \tag{36}$$

Let $w \geq \underline{w}$ and assume by way of contradiction $f^*(w) > 0$. Then, by (36),

$$\mu^* = \delta.$$

Because of $\delta < 1/\theta < \alpha$, also $c^*(w) > 0$ and $l^*(w) > 0$, so that also (34) and (35) hold as equalities. Then, using (34) and (36) to substitute out the Lagrange multipliers from (35) reveals that

$$l^*(w) = \frac{\alpha - \delta + (1 - \theta\delta)w}{\beta}. \tag{37}$$

From (32) and the budget constraint (2) one has

$$f^* = T - (1 + \theta w)l^*. \tag{38}$$

Inserting (37) into (38) gives the optimal amount of free time as a function of the wage rate:

$$f^*(w) = T - \frac{(1 + \theta w)[\alpha - \delta + (1 - \theta\delta)w]}{\beta}.$$

The function f^* thus defined satisfies $f^*(0) > 0$, $df^*/dw < 0$, $d^2 f^*/dw^2 < 0$.

Now, define $\underline{w}$ as the positive root of

$$(1 + \theta w)[\alpha - \delta + (1 - \theta\delta)w] - \beta T = 0,$$

so that $f^*(\underline{w}) = 0$. Hence, if $w \geq \underline{w}$, $f^*(w) \leq 0$ a contradiction. This shows that at those wage rates, $f^* = 0$, $\mu^* > \delta$ and the model with pure free time is equivalent to the one in the main text in the case of a binding time constraint.